Illinois Journal of Mathematics
Volume 52, Number 1, Spring 2008, Pages 47–76
S 0019-2082

THE TOTAL ABSOLUTE CURVATURE OF OPEN CURVES IN E^3

KAZUYUKI ENOMOTO, JIN-ICHI ITOH, AND ROBERT SINCLAIR

Abstract. The total absolute curvature of open curves in E^3 is discussed. We study the curves which attain the infimum of the total absolute curvature in the set of curves with fixed end-points, end-directions, and length. We show that if the total absolute curvature of a sequence of curves in this set tends to the infimum, the limit curve must lie in a plane. Moreover, it is shown that the limit curve is either a subarc of a closed plane convex curve or a piecewise linear curve with at most three edges. The uniqueness of the curves minimizing the total absolute curvature is also discussed. This extends the results in [*Yokohama Math. J.* **48** (2000), 83–96], which deals with a similar problem for curves in E^2.

1. Introduction

Let Σ be a C^2 curve in the 3-dimensional Euclidean space E^3. The total absolute curvature $\tau(\Sigma)$ of Σ is the total integral of the curvature of Σ. If Σ is only piecewise C^2, we add the exterior angles at nonsmooth points to define $\tau(\Sigma)$. In particular, the total absolute curvature of a piecewise linear curve is just the sum of the exterior angles.

The study of the total absolute curvature of curves has a long history since Fenchel proved that the total absolute curvature of any closed curve in E^3 is not less than 2π and the infimum is attained by a plane convex curve [11]. Fenchel's theorem has been extended in various directions ([16], etc.), but it seems that most results are concerned with closed curves and not much has been done for open curves. In this paper, we study the total absolute curvature of open curves in E^3.

Received June 10, 2006; received in final form September 20, 2007.
2000 *Mathematics Subject Classification.* 53A04.

When we are given points p and q in E^3, unit tangent vectors X at p and Y at q and a positive constant L $(> |pq|)$, we denote by $\mathcal{C}(p, X, q, Y, L)$ the set of all piecewise C^2 curves in E^3 whose endpoints, end-directions, and length are p, q, X, Y, and L. We study the shape of a curve which minimizes the total absolute curvature in $\mathcal{C}(p, X, q, Y, L)$. Fenchel's theorem says that, if the curve is closed (i.e., $p = q$, $X = Y$), then $\inf\{\tau(\Sigma) : \Sigma \in \mathcal{C}(p, X, q, Y, L)\} = 2\pi$, which does not depend on L. For the open case, in contrast, $\inf\{\tau(\Sigma) : \Sigma \in \mathcal{C}(p, X, q, Y, L)\}$ usually depends on the length L. This can be seen even in the planar case (i.e., the case when Σ lies in a plane), as Remark 5.8 shows. The planar case is studied in our previous paper [8]. We also note that the problem becomes almost trivial if we do not fix the length, as we explain in Remark 5.9. Furthermore, in certain biological and engineering contexts, to be discussed below, it is natural to specify the length.

Another difference between the closed case and the open case is that for the open case $\inf\{\tau(\Sigma) : \Sigma \in \mathcal{C}(p, X, q, Y, L)\}$ is not necessarily attained by an element of $\mathcal{C}(p, X, q, Y, L)$; When the total absolute curvature of a sequence of curves in $\mathcal{C}(p, X, q, Y, L)$ tends to $\inf\{\tau(\Sigma) : \Sigma \in \mathcal{C}(p, X, q, Y, L)\}$, the limit curve Σ_0 may have end-directions different from X or Y, and then Σ_0 becomes only an element of $\mathcal{C}(p, q, L)$, the set of all piecewise C^2 curves whose endpoints and length are p, q and L. This phenomenon has already been seen in the planar case [8].

In the main theorem of the present paper (Theorem 5.3), we show that if $\{\Sigma_k : k = 1, 2, 3, \dots\}$ is a sequence of curves in $\mathcal{C}(p, X, q, Y, L)$ such that $\tau(\Sigma_k)$ tends to $\inf\{\tau(\Sigma) : \Sigma \in \mathcal{C}(p, X, q, Y, L)\}$ as $k \to \infty$, then the limit curve must lie in a plane. Moreover, it is shown that the limit curve is either a subarc of a closed plane convex curve or a piecewise linear curve with at most three edges (Theorem 5.3).

Since the limit curve may have different end-directions, we introduce the notion of the "extended total absolute curvature" $\tilde{\tau}$, which is defined, for curves in $\mathcal{C}(p, q, L)$, as the sum of $\tau(\Sigma)$, the angle between X and the initial tangent vector, and the angle between Y and the terminal tangent vector (Definition 2.2). If a sequence of curves $\{\Sigma_k\}$ in $\mathcal{C}(p, X, q, Y, L)$ converges to Σ_0, then $\tau(\Sigma_k)$ converges to $\tilde{\tau}(\Sigma_0)$.

We first study the simple case when the curve is only a piecewise linear curve with two edges, and determine the shape of the curve minimizing the extended total absolute curvature among all piecewise linear curves with two edges in $\mathcal{C}(p, q, L)$ (Theorem 3.8). Then we study in Section 4 the case when the curve is a piecewise linear curve with three edges. Using Theorem 3.8, we show that if $\tilde{\tau}$ is minimized by P_0 in the set of all piecewise linear curves with three edges, then P_0 must lie in a plane. Moreover, it is shown that P_0 is part of a convex quadrilateral or what we call a "Z-curve" or actually a piecewise linear curve with two edges (Theorem 4.3). Theorem 4.3 implies through an induction argument that if $\tilde{\tau}$ in the set of all piecewise linear curves with n edges is minimized by P_0, then P_0 lies in a plane and P_0 is part of a convex polygon or a Z-curve or a piecewise linear curve with two edges (Theorem 5.1). Through approximation of a piecewise smooth curve by a piecewise linear curve, we finally show that if a piecewise smooth curve Σ_0 minimizes $\tilde{\tau}$, then Σ_0 lies in a plane and Σ_0 is part of a closed plane convex curve or a Z-curve or a piecewise linear curve with two edges (Theorem 5.3).

The uniqueness of the curves minimizing the extended total absolute curvature is also discussed. When a piece of a closed plane convex curve minimizes $\tilde{\tau}$, including the case when the curve is closed, the shape of the curve is relatively flexible. However, when piecewise linear curves are the only possibility to minimize $\tilde{\tau}$, the shape of the curve is unique in most cases. We explain what are the exceptional cases and how many curves minimize $\tilde{\tau}$ in those cases (Theorem 3.8, Theorem 5.3). In the process, we study the cut locus of the "two-sided" disk in the unit 2-sphere, which is described in the Appendix. What we prove in the Appendix is used in the proof of Theorem 3.8. One may have an independent interest in the subject of the Appendix.

The present paper generalizes the results in [8], in which we study a similar problem for curves in E^2. This problem is also studied for curves in S^2 in [9] and [10].

Our problem is somehow similar to the problem of elastic curves, which can be described as the problem of finding a curve in $\mathcal{C}(p, X, q, Y, L)$ minimizing $\int_\Sigma k(s)^2\,ds$. However, the shape of the curve minimizing the functional and the method of solution seem very different between these two problems. (See, for example, [22].) Milnor [17] studied functionals such as $\int_\Sigma \sqrt{k(s)^2 + \tau(s)^2}\,ds$ and $\int_\Sigma |\tau(s)|\,ds$ for closed curves in E^3, where τ is the torsion. It may be interesting to study the problem of finding a curve in $\mathcal{C}(p, X, q, Y, L)$ minimizing those functionals.

Biology provides us with a strong motivation to explore geometric optimization problems of fixed-length open chains. The protein folding problem [6] is one of the greatest unsolved problems in molecular biology, with immediate implications for the understanding of neurodegenerative diseases ([7], [18]) and applications in drug design ([1], [14], [15]). A protein can be modeled as a chain of fixed length with a large number of links, usually in the hundreds. The concept of curvature can be defined and is of biological relevance [24]. Its final structure is determined by its sequence of amino acids, which are usually known. What is not known is the mechanism by which a protein folds into its correct shape. It is believed that the correct shape can be predicted as a solution of an energy minimization problem ([19], [23]). The point we wish to make here is that a mathematical understanding of geometric optimization problems of finite length open curves or chains will either directly help in gaining a handle on the protein folding problem itself or at least provide us with a language with which to describe the eventual solution. Much more mathematical work needs to be done in this direction (although much has already been done, see [21] for but one example). The reader may find [4] of interest, as a serious attempt to explain molecular biology and its challenges to a mathematical audience. The special case of finding an optimal configuration of an open chain with given endpoints is known as the loop closure problem [3]. The problem is also of practical importance in robotics [5], where we expect that our results will find application.

2. Preliminaries

Let Σ be a piecewise C^2 curve in the 3-dimensional Euclidean space E^3. Let ℓ be the length of Σ and $x(s)$ $(0 \le s \le \ell)$ be a parameterization of Σ by its arclength. Let $0 = s_0 < s_1 < \cdots < s_n = \ell$ be the subdivision of $[0, \ell]$ such that $\{x(s) : s_{i-1} < s < s_i\}$ is C^2 for each $i = 1, \ldots, n$. If Σ is C^2 at $x(s)$, the curvature $k(s)$ of Σ is defined by $k(s) = |d^2x/ds^2|$. Let $T = T(s) = dx/ds$ be the unit tangent vector of Σ. For each i, let θ_i $(0 \le \theta_i \le \pi)$ be the angle at $x(s_i)$ between $\lim_{s \to s_i - 0} T(s)$ and $\lim_{s \to s_i + 0} T(s)$.

Definition 2.1. The *total absolute curvature* $\tau(\Sigma)$ of Σ is defined by

$$\tau(\Sigma) = \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} k(s)\, ds + \sum_{i=1}^{n-1} \theta_i.$$

We now define several classes of curves. Here, p and q are points in E^3, X and Y are unit tangent vectors of E^3 at p and q, respectively, and L is a positive constant greater than $|pq|$. Let

$$\begin{aligned}
\mathcal{C}(p,q) &= \{\Sigma : x(s) \mid x(0) = p, x(\ell) = q\},\\
\mathcal{C}(p,q,L) &= \{\Sigma \in \mathcal{C}(p,q) \mid \ell = L\},\\
\mathcal{C}(p,X,q,Y) &= \{\Sigma \in \mathcal{C}(p,q) \mid T(0) = X, T(\ell) = Y\},\\
\mathcal{C}(p,X,q,Y,L) &= \mathcal{C}(p,q,L) \cap \mathcal{C}(p,X,q,Y).
\end{aligned}$$

Let n be a positive integer and let $\mathcal{P}_n$ be the set of all piecewise linear curves with n edges. For all $m < n$, we regard $\mathcal{P}_m$ as a subset of $\mathcal{P}_n$ by allowing angles between two edges to be zero. Let $\mathcal{P}_n(p,q,L) = \mathcal{P}_n \cap \mathcal{C}(p,q,L)$.

When $\{\Sigma_k : k = 1, 2, \ldots\}$ is a sequence of curves in $\mathcal{C}(p,X,q,Y)$, the limit curve $\lim_{k\to\infty} \Sigma_k$ may not be an element of $\mathcal{C}(p,X,q,Y)$, since the limit curve may not be tangent to X or Y (Figure 2.1). If $\Sigma \in \mathcal{C}(p,q)$ is the limit curve of $\{\Sigma_k \in \mathcal{C}(p,X,q,Y) : k = 1, 2, \ldots\}$, then we have

$$\lim_{k\to\infty} \tau(\Sigma_k) = \angle(X, T(0)) + \tau(\Sigma) + \angle(T(\ell), Y).$$

This leads us to the following notion of the extended total absolute curvature.

Definition 2.2. For $\Sigma \in \mathcal{C}(p,q)$, the *extended total absolute curvature* $\tilde{\tau}(\Sigma)$ is defined by

$$\tilde{\tau}(\Sigma) = \angle(X, T(0)) + \tau(\Sigma) + \angle(T(\ell), Y),$$

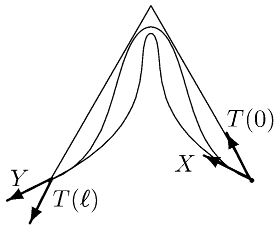


FIGURE 2.1. Illustration of the length condition and jumps of the tangents from the initial conditions to the limit tangents at the ends, as incorporated into the definition of extended total absolute curvature.

where $\angle(\cdot,\cdot)$ denotes the angle between two vectors with value in $[0,\pi]$.

Note that the definition of $\tilde{\tau}$ depends on the choice of X and Y, and $\tilde{\tau}(\Sigma)$ may be regarded as the total absolute curvature of $\Sigma\in\mathcal{C}(p,q)$ "as a curve in $\mathcal{C}(p,X,q,Y)$". If Σ happens to lie in $\mathcal{C}(p,X,q,Y)$, then we have $\tilde{\tau}(\Sigma)=\tau(\Sigma)$.

DEFINITION 2.3. An element Σ_0 of $\mathcal{C}(p,q,L)$ is called an *extremal curve in* $\mathcal{C}(p,X,q,Y,L)$ if

$$\tilde{\tau}(\Sigma_0)\le\tau(\Sigma)$$

holds for any $\Sigma\in\mathcal{C}(p,X,q,Y,L)$.

For $\Sigma\in\mathcal{C}(p,q,L)$, let T_Σ be the piecewise C^1 curve in S^2 which consists of $\bigcup_{i=1}^n\{T(s): s_{i-1}<s<s_i\}$, the geodesic arcs between $\lim_{s\to s_i-0}T(s)$ and $\lim_{s\to s_i+0}T(s)$ $(i=1,\dots,n)$, the geodesic arc between X and $T(0)$ and the geodesic arc between Y and $T(L)$. Then the length of T_Σ is equal to $\tilde{\tau}(\Sigma)$.

DEFINITION 2.4. A subarc of a closed plane convex curve (the boundary of a convex domain in a plane) is called a *plane convex arc*. $\{p,X,q,Y,L\}$ is said to *satisfy the convexity condition* if there exists a plane convex arc Σ in $\mathcal{C}(p,X,q,Y,L)$.

If $\{p,X,q,Y,L\}$ satisfies the convexity condition, then every plane convex arc Σ in $\mathcal{C}(p,X,q,Y,L)$ becomes an extremal curve in $\mathcal{C}(p,X,q,Y,L)$. If a sequence of plane convex arcs in $\mathcal{C}(p,X,q,Y,L)$ converges to a plane convex arc Σ_0 in $\mathcal{C}(p,q,L)$, then Σ_0 is an extremal curve in $\mathcal{C}(p,X,q,Y,L)$. For a plane convex arc Σ_0 the curve T_{Σ_0} becomes a subarc of a great circle.

REMARK 2.5. If $p=q$, the convexity condition is automatically satisfied by $\{p,X,q,Y,L\}$ for any X, Y and L.

3. Piecewise linear curves with two edges

In this section, we study the shape of a piecewise linear curve with two edges which attains $\inf\{\tilde{\tau}(P): P\in\mathcal{P}_2(p,q,L)\}$. Since $\mathcal{P}_2(p,q,L)$ is compact and $\tilde{\tau}$ is continuous in $\mathcal{P}_2(p,q,L)$, such a piecewise linear curve always exists.

Any element P of $\mathcal{P}_2(p,q,L)$ is expressed as

$$P = pp_1 \cup p_1 q$$

with $|pp_1| + |p_1 q| = L$. Then we have

$$\tilde{\tau}(P) = \angle(X, \overrightarrow{pp_1}) + \angle(\overrightarrow{pp_1}, \overrightarrow{p_1 q}) + \angle(\overrightarrow{p_1 q}, Y).$$

If we set

$$\xi = \frac{\overrightarrow{pp_1}}{|pp_1|},$$

then $\xi \in S^2$ is uniquely determined by P. Conversely, for any $\xi \in S^2$, there exists a unique point p_1 which satisfies

$$\frac{\overrightarrow{pp_1}}{|pp_1|} = \xi, \qquad |pp_1| + |p_1 q| = L.$$

By this, $\mathcal{P}_2(p,q,L)$ is identified with S^2. Thus, $\tilde{\tau}(P)$ on $\mathcal{P}_2(p,q,L)$ may be regarded as a function $\tilde{\tau}(\xi)$ on S^2. Let

$$f(\xi) = \frac{\overrightarrow{p_1 q}}{|p_1 q|}.$$

Then f defines a bijective map (Figure 3.1) on the unit sphere S^2.

If d denotes the distance in S^2, we have

$$\angle(X, \overrightarrow{pp_1}) = d(X,\xi), \qquad \angle(\overrightarrow{pp_1}, \overrightarrow{p_1 q}) = d(\xi, f(\xi)), \qquad \angle(\overrightarrow{p_1 q}, Y) = d(f(\xi), Y).$$

Hence, $\tilde{\tau}$ on $\mathcal{P}_2(p,q,L)$, as a function in S^2, is written as (Figure 3.2)

$$\tilde{\tau}(\xi) = d(X,\xi) + d(\xi, f(\xi)) + d(f(\xi), Y). \tag{3.1}$$

When $p \neq q$, we define a unit vector Z by

$$Z = \frac{\overrightarrow{pq}}{|pq|}.$$

Since ξ, $f(\xi)$ and Z lie in a plane as vectors in E^3, they lie on a great circle as points in S^2.

$\{p, X, q, Y, L\}$ satisfies the convexity condition if and only if we have either $p = q$ or $p \neq q$ and $\{X, f^{-1}(Y), Z, f(X), Y, -Z\}$ lies on a great circle in S^2 in this order (Figure 3.3).

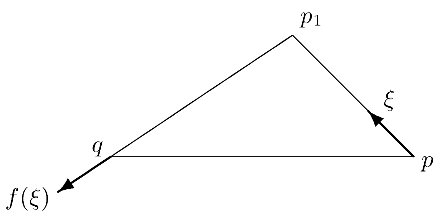


FIGURE 3.1. The map f.

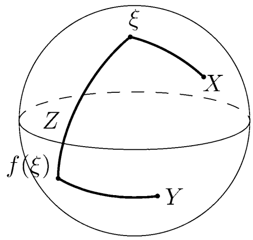


FIGURE 3.2. Illustration of the computation of $\tilde{\tau}$ in (3.1).

In the following argument, we assume $p \neq q$. Let $d_X = d(X, \cdot)$, $d_Y = d(Y, \cdot)$ and $d_Z = d(Z, \cdot)$. Then (3.1) is rewritten as

$$\tilde{\tau}(\xi) = d_X(\xi) + d_Z(\xi) + d_Z(f(\xi)) + d_Y(f(\xi)). \tag{3.2}$$

At any ξ with $\xi \neq \pm X$, d_X is differentiable and its gradient vector ∇d_X is defined. ∇d_X is identical with the unit tangent vector of the oriented geodesic from X to ξ. $\tilde{\tau}(\xi)$ is differentiable for all ξ with $\xi \neq \pm X$ and $\xi \neq f^{-1}(\pm Y)$. Note that $d_Z(\xi) + d_Z(f(\xi)) = d(\xi, f(\xi))$ is differentiable for all ξ.

Let T be a tangent vector of S^2 at ξ. Let f_* be the differential of f. To describe a property of f_*T, we take a positively oriented orthonormal frame $\{E_1, E_2\}$ of the tangent bundle of S^2 defined in $S^2 \backslash \{Z, -Z\}$ with $E_1 = \nabla d_Z$. In the following lemma, λ is a function defined by

$$\lambda = \frac{|pp_1|}{|p_1 q|}.$$

We set $D = |pq|$. Then we have

$$|pp_1| = \frac{L^2 - D^2}{2(L - D\cos d(Z,\xi))},$$

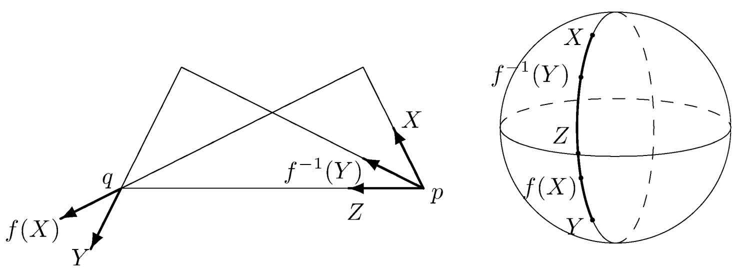


FIGURE 3.3. The convexity condition.

λ, as a function in ξ, is given by

$$\lambda(\xi) = \frac{|pp_1|}{L - |pp_1|} = \frac{L^2 - D^2}{L^2 + D^2 - 2LD\cos d(Z,\xi)}. \tag{3.3}$$

Lemma 3.1. *For any tangent vector T of S^2 at ξ, we have*

$$f_* T = \lambda(\xi)\bigl(-\langle T, E_1(\xi)\rangle E_1(f(\xi)) + \langle T, E_2(\xi)\rangle E_2(f(\xi))\bigr).$$

Proof. Let $\xi(t)$ be the point on the oriented geodesic $Z\xi$ such that $d(Z,\xi(t)) = t$. Then $\xi'(t) = E_1(\xi(t))$. Let $p_1(t)$ be the point in E^3 which corresponds to $\xi(t)$. Since $p_1(t)$ lies in one plane for all t, $f(\xi(t))$ moves along the geodesic through Z and ξ, which implies that $f_*(E_1(\xi(t)))$ is parallel to $E_1(f(\xi(t)))$. $d(Z, f(\xi(t))) = \angle(\overrightarrow{pq}, \overrightarrow{p_1(t)q})$ is related to $t = d(Z,\xi(t)) = \angle(\overrightarrow{pq}, \overrightarrow{pp_1(t)})$ as

$$d(Z, f(\xi(t))) = \arccos\left(\frac{2LD - (L^2 + D^2)\cos t}{L^2 + D^2 - 2LD\cos t}\right).$$

Thus, we have

$$\begin{aligned} f_*(E_1(\xi(t))) &= \frac{d}{dt} d(Z, f(\xi(t))) E_1(f(\xi(t))) \\ &= -\frac{L^2 - D^2}{L^2 + D^2 - 2LD\cos t} E_1(f(\xi(t))) \\ &= -\lambda(\xi(t)) E_1(f(\xi(t))). \end{aligned} \tag{3.4}$$

The circle centered at Z of radius $d(Z,\xi)$ is an integral curve of E_2. When ξ moves along this circle, $f(\xi)$ moves along the circle centered at Z of radius $d(Z, f(\xi))$, and we see that (Figure 3.4)

$$\begin{aligned} f_*(E_2(\xi)) &= \frac{\sin d(Z, f(\xi))}{\sin d(Z,\xi)} E_2(f(\xi)) \\ &= \frac{|pp_1|}{|p_1 q|} E_2(f(\xi)) \\ &= \lambda(\xi) E_2(f(\xi)). \end{aligned} \tag{3.5}$$

From (3.4) and (3.5), we have

$$\begin{aligned} f_* T &= \langle T, E_1(\xi)\rangle f_*(E_1(\xi)) + \langle T, E_2(\xi)\rangle f_*(E_2(\xi)) \\ &= \lambda(\xi)\bigl(-\langle T, E_1(\xi)\rangle E_1(f(\xi)) + \langle T, E_2(\xi)\rangle E_2(f(\xi))\bigr). \quad \square \end{aligned}$$

Lemma 3.2. *Suppose $\nabla\tilde{\tau} = 0$ at ξ_0. Then either (1) or (2) holds:*

(1) *$\{p, X, q, Y, L\}$ satisfies the convexity condition and ξ_0 is a point on the geodesic segment $Xf^{-1}(Y)$.*
(2) *$d(Z,\xi_0) = \arccos(D/L)$.*

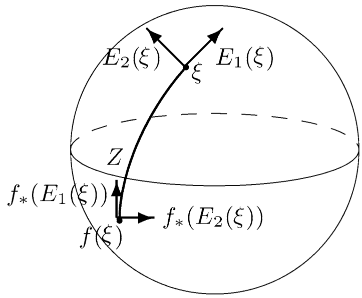


FIGURE 3.4. The geometry of (3.5).

Proof. Let T be a tangent vector of S^2 at ξ. By (3.2), we have

$$\begin{aligned} T\tilde{\tau} = &\langle \nabla d_X(\xi), T\rangle + \langle \nabla d_Z(\xi), T\rangle + \langle \nabla d_Z(f(\xi)), f_*T\rangle \\ &+ \langle \nabla d_Y(f(\xi)), f_*T\rangle. \end{aligned}$$

If we set $T = E_1(\xi) = \nabla d_Z(\xi)$, using Lemma 3.1, we have

$$E_1\tilde{\tau} = \langle \nabla d_X(\xi), E_1(\xi)\rangle + 1 + \lambda(\xi)\big(-1 - \langle \nabla d_Y(f(\xi)), E_1(f(\xi))\rangle\big). \tag{3.6}$$

If we set $T = E_2(\xi)$, we have

$$E_2\tilde{\tau} = \langle \nabla d_X(\xi), E_2(\xi)\rangle + \lambda(\xi)\langle \nabla d_Y(f(\xi)), E_2(f(\xi))\rangle. \tag{3.7}$$

If $\nabla\tilde{\tau} = 0$ at ξ_0, then (3.6) gives

$$\langle \nabla d_X(\xi_0), E_1(\xi_0)\rangle + 1 = \lambda(\xi_0)\big(1 + \langle \nabla d_Y(f(\xi_0)), E_1(f(\xi_0))\rangle\big) \tag{3.8}$$

and (3.7) gives

$$\langle \nabla d_X(\xi_0), E_2(\xi_0)\rangle = -\lambda(\xi_0)\langle \nabla d_Y(f(\xi_0)), E_2(f(\xi_0))\rangle. \tag{3.9}$$

Combining (3.8) and (3.9) with

$$\langle \nabla d_X(\xi_0), E_1(\xi_0)\rangle^2 + \langle \nabla d_X(\xi_0), E_2(\xi_0)\rangle^2 = |\nabla d_X(\xi_0)|^2 = 1 \tag{3.10}$$

and

$$\begin{aligned} &\langle \nabla d_Y(f(\xi_0)), E_1(f(\xi_0))\rangle^2 + \langle \nabla d_Y(f(\xi_0)), E_2(f(\xi_0))\rangle^2 \\ &\quad = |\nabla d_Y(f(\xi_0))|^2 = 1, \end{aligned} \tag{3.11}$$

we obtain

$$\big(\langle \nabla d_X(\xi_0), E_1(\xi_0)\rangle + 1\big)\big(1 - \lambda(\xi_0)\big) = 0. \tag{3.12}$$

This implies that either

$$\nabla d_X(\xi_0) = -\nabla d_Z(\xi_0) \tag{3.13}$$

or

$$\lambda(\xi_0) = 1 \tag{3.14}$$

holds.

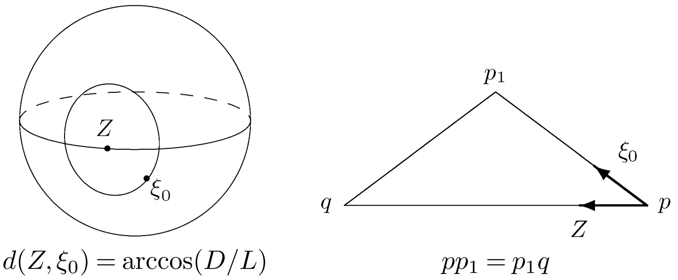


FIGURE 3.5. The geometry of Lemma 3.2.

If (3.13) holds, then (3.8) gives

$$\nabla d_Y(f(\xi_0)) = -\nabla d_Z(f(\xi_0)). \tag{3.15}$$

(3.13) and (3.15) imply that we have $\{X, \xi_0, Z, f(\xi_0), Y\}$ on a great circle in this order. Since $d(Z,X) \geq d(Z,\xi_0)$, we have $d(Z,f(X)) \leq d(Z,f(\xi_0))$. Hence, if such ξ_0 exists, we have $d(Z,f(X)) \leq d(Z,Y)$, or equivalently, $d(Z,X) \geq d(Z,f^{-1}(Y))$. Thus, $\{p,X,q,Y,L\}$ satisfies the convexity condition.

If (3.14) holds, we have $d(Z,\xi_0) = \arccos(D/L)$. □

REMARK 3.3.

(1) If $\{p,X,q,Y,L\}$ satisfies the convexity condition, we have $\nabla\tilde{\tau} = 0$ at any point on the geodesic segment $Xf^{-1}(Y)$.
(2) In the case (2) in Lemma 3.2, ξ_0 lies on the circle of radius $\arccos(D/L)$ centered at Z. Then we have $|pp_1| = |p_1q|$ by (3.3) and (3.14) (Figure 3.5).

Since $\tilde{\tau}(\xi)$ is continuous in S^2, it attains its minimum at some ξ_0. If $\tilde{\tau}$ is differentiable at ξ_0, ξ_0 has the property described in Lemma 3.2. If $\tilde{\tau}$ is not differentiable at ξ_0, by (3.1), we see that either $\xi_0 = \pm X$ or $\xi_0 = f^{-1}(\pm Y)$. Among these points $\xi_0 = -X$ and $\xi_0 = f^{-1}(-Y)$ are not possible to be the minimal points of $\tilde{\tau}$, as we see in the following lemma.

LEMMA 3.4. *Neither $\xi = -X$ nor $\xi = f^{-1}(-Y)$ can attain* $\min\tilde{\tau}$.

Proof. We will show that $\tilde{\tau}(-X) > \tilde{\tau}(X)$.

Let $P = pp_1 \cup p_1q$ be the element of $\mathcal{P}_2(p,q,L)$ with $\overrightarrow{pp_1}/|pp_1| = X$ and $P' = pp_1' \cup p_1'q$ be the one with $\overrightarrow{pp_1'}/|pp_1'| = -X$. Then we have (Figure 3.6)

$$\begin{aligned}\tilde{\tau}(X) &= \angle(\overrightarrow{pp_1},\overrightarrow{p_1q}) + \angle(\overrightarrow{p_1q},Y)\\ &= \angle(\overrightarrow{pp_1},\overrightarrow{pq}) + \angle(\overrightarrow{pq},\overrightarrow{p_1q}) + \angle(\overrightarrow{p_1q},Y)\\ &\leq \angle(X,\overrightarrow{pq}) + \angle(\overrightarrow{pq},\overrightarrow{p_1q}) + \angle(\overrightarrow{p_1q},\overrightarrow{pq}) + \angle(\overrightarrow{pq},Y)\end{aligned}$$

while

$$\tilde{\tau}(-X) = \pi + \angle(\overrightarrow{pp_1'},\overrightarrow{p_1'q}) + \angle(\overrightarrow{p_1'q},Y)$$

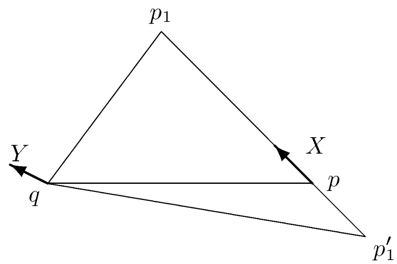


FIGURE 3.6. Computations in the proof of Lemma 3.4.

$$
\begin{aligned}
&= \pi + \angle(\overrightarrow{pp_1'}, \overrightarrow{pq}) + \angle(\overrightarrow{pq}, \overrightarrow{p_1'q}) + \angle(\overrightarrow{p_1'q}, Y) \\
&\geq 2\pi - \angle(X, \overrightarrow{pq}) + \angle(\overrightarrow{pq}, Y).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\tilde{\tau}(-X) - \tilde{\tau}(X) &\geq 2\pi - 2\angle(X, \overrightarrow{pq}) - 2\angle(\overrightarrow{pq}, \overrightarrow{p_1q}) \\
&= 2\pi - 2\angle(\overrightarrow{pp_1}, \overrightarrow{p_1q}) \\
&> 0.
\end{aligned}
$$

By a similar argument, one can show that $\tilde{\tau}(f^{-1}(-Y)) > \tilde{\tau}(f^{-1}(Y))$. $\square$

LEMMA 3.5 (Figure 3.7).

(1) *If* $d(X, Z) < \arccos(D/L)$, *then* $\xi = X$ *cannot attain* $\min \tilde{\tau}$, *unless* $\{p, X, q, Y, L\}$ *satisfies the convexity condition.*
(2) *If* $d(Y, Z) < \arccos(D/L)$, *then* $\xi = f^{-1}(Y)$ *cannot attain* $\min \tilde{\tau}$, *unless* $\{p, X, q, Y, L\}$ *satisfies the convexity condition.*

Proof. Let $\{E_1, E_2\}$ be a positively oriented orthonormal frame of the tangent bundle of S^2 with $E_1 = \nabla d_Z$. For any unit tangent vector T at X, we have $|f_* T| = \lambda(X)$ by Lemma 3.1. We choose T so that

$$f_* T = -\lambda(X) \nabla d_Y(f(X)).$$

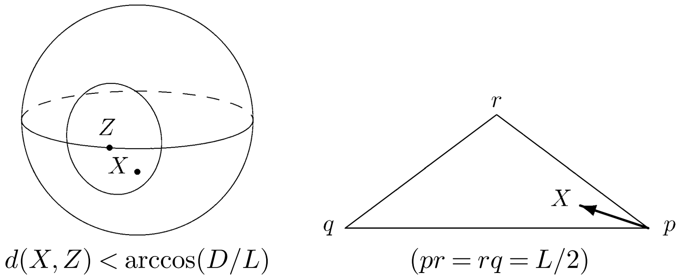


FIGURE 3.7. Illustration of the conditions of Lemma 3.5.

If we write T as
$$T = \cos\omega E_1(X) + \sin\omega E_2(X),$$
then f_*T becomes
$$f_*T = -\lambda(X)\big(\cos\omega E_1(f(X)) + \sin\omega E_2(f(X))\big).$$
$\tilde\tau(\xi)$ is not differentiable at $\xi = X$ because $d_X(\xi)$ is not differentiable there. However, the directional derivative of $d_X(\xi)$ does exist even at X, and we have $Td_X = 1$. Hence,
$$\begin{aligned} T\tilde\tau &= 1 + \langle \nabla d_Z, T\rangle + \langle \nabla d_Z(f(X)), f_*T\rangle + \langle \nabla d_Y(f(X)), f_*T\rangle \\ &= 1 + \cos\omega - \lambda(X)\cos\omega - \lambda(X) \\ &= (1+\cos\omega)\big(1-\lambda(X)\big). \end{aligned}$$
Since $d(X,Z) < \arccos(D/L)$, we have $\lambda(X) > 1$. We also have $\omega \neq \pi$, unless $\{p, X, q, Y, L\}$ satisfies the convexity condition. Thus, we have $T\tilde\tau < 0$, which implies that $\tilde\tau$ is not minimal at $\xi = X$. The proof for $\xi = f^{-1}(Y)$ is similar. □

LEMMA 3.6 (Figure 3.8). *If $d(X,Z) > d(Y,Z) \geq \arccos(D/L)$ holds, then we have $\tilde\tau(X) \leq \tilde\tau(f^{-1}(Y))$. The equality holds if and only if $\{p, X, q, Y, L\}$ satisfies the convexity condition.*

Proof. We have
$$\begin{aligned} \tilde\tau(X) &= d(X, f(X)) + d(f(X), Y) \\ &= d(X,Z) + d(Z, f(X)) + d(f(X), Y) \end{aligned} \tag{3.16}$$
and
$$\begin{aligned} \tilde\tau(f^{-1}(Y)) &= d(X, f^{-1}(Y)) + d(f^{-1}(Y), Y) \\ &= d(X, f^{-1}(Y)) + d(f^{-1}(Y), Z) + d(Z, Y). \end{aligned} \tag{3.17}$$
Set
$$R = d(X,Z), \qquad r = d(Z,Y), \qquad R_1 = d(Z, f(X)), \qquad r_1 = d(f^{-1}(Y), Z)$$

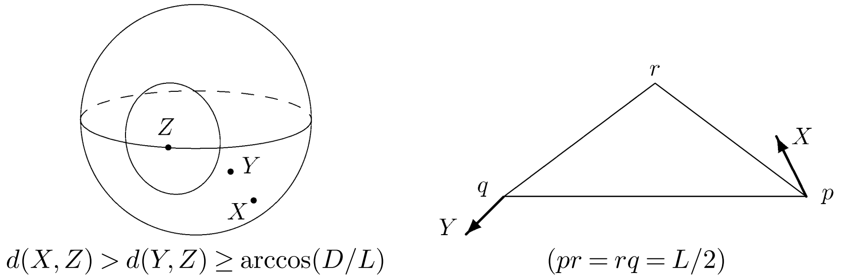
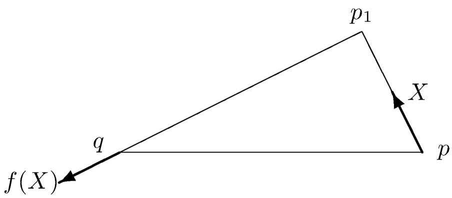


FIGURE 3.8. Illustration of the conditions of Lemma 3.6.



$\angle(X, \overrightarrow{pq}) = R$, $\angle(f(X), \overrightarrow{pq}) = R_1$.

FIGURE 3.9. The geometry of (3.18).

and

$$\theta = \angle XZf^{-1}(Y).$$

Using the plane trigonometry for $\triangle pp_1q$ with $\overrightarrow{pp_1}/|pp_1| = X$, we have (Figure 3.9)

$$\begin{aligned} \sin R_1 &= \frac{(L^2 - D^2)\sin R}{L^2 + D^2 - 2LD\cos R}, \\ \cos R_1 &= \frac{-(L^2 + D^2)\cos R + 2LD}{L^2 + D^2 - 2LD\cos R}. \end{aligned} \tag{3.18}$$

For $\triangle pp_1q$ with $\overrightarrow{p_1q}/|p_1q| = Y$, we have

$$\begin{aligned} \sin r_1 &= \frac{(L^2 - D^2)\sin r}{L^2 + D^2 - 2LD\cos r}, \\ \cos r_1 &= \frac{-(L^2 + D^2)\cos r + 2LD}{L^2 + D^2 - 2LD\cos r}. \end{aligned} \tag{3.19}$$

Using the sphere trigonometry for $\triangle Xf^{-1}(Y)Z$, we have

$$\cos(d(X, f^{-1}(Y))) = \cos R\cos r_1 + \cos\theta\sin R\sin r_1. \tag{3.20}$$

For $\triangle f(X)YZ$, we have

$$\cos(d(f(X), Y)) = \cos r\cos R_1 + \cos\theta\sin r\sin R_1. \tag{3.21}$$

Now we set

$$\delta = d(X, f^{-1}(Y)) - d(f(X), Y)$$

and regard it as a function in θ. By (3.20) and (3.21), δ is explicitly written as

$$\begin{aligned} \delta(\theta) = {} & \arccos(\cos R\cos r_1 + \cos\theta\sin R\sin r_1) \\ & - \arccos(\cos r\cos R_1 + \cos\theta\sin r\sin R_1). \end{aligned} \tag{3.22}$$

Then we have

$$\begin{aligned} \frac{d\delta}{d\theta} = {} & \sin\theta\sin R\sin r_1\big(1 - (\cos R\cos r_1 + \cos\theta\sin R\sin r_1)^2\big)^{-1/2} \\ & - \sin\theta\sin r\sin R_1\big(1 - (\cos r\cos R_1 + \cos\theta\sin r\sin R_1)^2\big)^{-1/2}. \end{aligned} \tag{3.23}$$

$d\delta/d\theta = 0$ holds if and only if we have either $\sin\theta = 0$ or

$$\begin{aligned}\sin R \sin r_1 \bigl(1-(\cos R\cos r_1+\cos\theta\sin R\sin r_1)^2\bigr)^{-1/2} \\ = \sin r\sin R_1\bigl(1-(\cos r\cos R_1+\cos\theta\sin r\sin R_1)^2\bigr)^{-1/2}.\end{aligned} \tag{3.24}$$

Using (3.18) and (3.19), we can rewrite (3.24) as

$$\begin{aligned}(1+\cos R\cos r+\cos\theta\sin R\sin r)\left(\frac{D}{L}\right)^2 - 2(\cos R+\cos r)\frac{D}{L} \\ + (1+\cos R\cos r-\cos\theta\sin R\sin r) = 0.\end{aligned} \tag{3.25}$$

Since

$$\begin{aligned}(\cos R+\cos r)^2 - (1+\cos R\cos r+\cos\theta\sin R\sin r) \\ \times (1+\cos R\cos r-\cos\theta\sin R\sin r) = -\sin^2\theta\sin^2 R\sin^2 r,\end{aligned}$$

(3.25) holds if and only if $\sin\theta = 0$. Hence $d\delta/d\theta = 0$ holds if and only if $\sin\theta = 0$. If $\theta = 0$, $\delta = R - r_1 + R_1 - r < 0$. If $\theta = \pi$, $\delta = R + r_1 - R_1 - r > 0$. Thus, we see that the minimum of δ is $R - r_1 + R_1 - r$, and hence

$$\begin{aligned}d(X, f^{-1}(Y)) - d(f(X), Y) \\ \geq d(X,Z) - d(f^{-1}(Y), Z) + d(Z, f(X)) - d(Z,Y).\end{aligned} \tag{3.26}$$

(3.26) gives

$$\begin{aligned}d(X,Z) + d(Z, f(X)) + d(f(X), Y) \\ \leq d(X, f^{-1}(Y)) + d(f^{-1}(Y), Z) + d(Z,Y),\end{aligned} \tag{3.27}$$

which means $\tilde{\tau}(X) \leq \tilde{\tau}(f^{-1}(Y))$. The equality holds if and only if $\theta = 0$, or equivalently, $\{p, X, q, Y, L\}$ satisfies the convexity condition. □

Lemma 3.7. *If either $d(X,Z) > \arccos(D/L)$ or $d(Y,Z) > \arccos(D/L)$ holds, then ξ with $d_Z(\xi) = \arccos(D/L)$ cannot attain $\min\tilde{\tau}$, with the only exception that $\{p, X, q, Y, L\}$ satisfies the convexity condition and ξ is the point on the geodesic segment $Xf^{-1}(Y)$ with $d_Z(\xi) = \arccos(D/L)$.*

Proof. We assume that $d(X,Z) > \arccos(D/L)$. Let ξ be any point in S^2 with $d_Z(\xi) = \arccos(D/L)$. Then we may apply (3.27) with $Y = f(\xi)$ to have

$$\begin{aligned}d(X,Z) + d(Z, f(X)) + d(f(X), f(\xi)) \\ \leq d(X,\xi) + d(\xi, Z) + d(Z, f(\xi)).\end{aligned} \tag{3.28}$$

By (3.28), we have

$$\begin{aligned}\tilde{\tau}(X) &= d(X,Z) + d(Z, f(X)) + d(f(X), Y) \\ &\leq d(X,\xi) + d(\xi,Z) + d(Z, f(\xi)) - d(f(X), f(\xi)) + d(f(X), Y) \\ &\leq d(X,\xi) + d(\xi,Z) + d(Z, f(\xi)) + d(f(\xi), Y) \\ &= \tilde{\tau}(\xi),\end{aligned}$$

where the equalities hold if and only if $\{p, X, q, Y, L\}$ satisfies the convexity condition and ξ is a point on $Xf^{-1}(Y)$. This shows that if the convexity condition is not satisfied, ξ with $d_Z(\xi) = \arccos(D/L)$ cannot attain $\min \tilde{\tau}$. The proof for the case when $d(Y, Z) > \arccos(D/L)$ is similar. □

We now state the main result of this section.

THEOREM 3.8. *Suppose that $P_0 = pp_1 \cup p_1 q$ is an extremal curve in $\mathcal{P}_2(p, X, q, Y, L)$.*

(1) *If $\{p, X, q, Y, L\}$ satisfies the convexity condition, then P_0 lies in a plane containing X and Y. In this case, p_1 can be any point such that $\overrightarrow{pp_1}/|pp_1|$ lies on the geodesic segment $Xf^{-1}(Y)$ in S^2.*

(2) *If $p \neq q$ and if $\angle(X, \overrightarrow{pq}) < \arccos(D/L)$ and $\angle(Y, \overrightarrow{pq}) < \arccos(D/L)$, then p_1 is a point which satisfies $|pp_1| = |p_1 q| = L/2$. Moreover, the following holds:*
  - (2-1) *If $\angle(X, \overrightarrow{pq}) \neq \angle(Y, \overrightarrow{pq})$, then p_1 is unique.*
  - (2-2) *If $X = Y = \overrightarrow{pq}/|pq|$, then p_1 can be any point with $|pp_1| = |p_1 q| = L/2$.*
  - (2-3) *If $\angle(X, \overrightarrow{pq}) = \angle(Y, \overrightarrow{pq}) = \alpha \neq 0$ and if $\angle(X, Y) \geq \arccos((L^2 - D^2 - 2D^2 \sin^2 \alpha \tan^2 \alpha)/(L^2 - D^2))$, then p_1 is unique.*
  - (2-4) *If $\angle(X, \overrightarrow{pq}) = \angle(Y, \overrightarrow{pq}) = \alpha \neq 0$ and if $\angle(X, Y) < \arccos((L^2 - D^2 - 2D^2 \sin^2 \alpha \tan^2 \alpha)/(L^2 - D^2))$, then there are exactly two choices for p_1.*

(3) *If $\{p, X, q, Y, L\}$ does not satisfy the convexity condition and if $\angle(X, \overrightarrow{pq}) \geq \arccos(D/L)$ and $\angle(X, \overrightarrow{pq}) > \angle(Y, \overrightarrow{pq})$, then p_1 is the unique point determined by $\overrightarrow{pp_1}/|pp_1| = X$.*

(4) *If $\{p, X, q, Y, L\}$ does not satisfy the convexity condition and if $\angle(Y, \overrightarrow{pq}) \geq \arccos(D/L)$ and $\angle(Y, \overrightarrow{pq}) > \angle(X, \overrightarrow{pq})$, then p_1 is the unique point determined by $\overrightarrow{p_1 q}/|p_1 q| = Y$.*

(5) *If $\{p, X, q, Y, L\}$ does not satisfy the convexity condition and $\angle(X, \overrightarrow{pq}) = \angle(Y, \overrightarrow{pq}) \geq \arccos(D/L)$, then p_1 is either the point determined by $\overrightarrow{pp_1}/|pp_1| = X$ or the point determined by $\overrightarrow{p_1 q}/|p_1 q| = Y$.*

Proof (Figure 3.10). Suppose $p = q$. Then for any $\xi \in S^2$ we have $f(\xi) = -\xi$, and hence by (3.1),

$$
\begin{aligned}
\tilde{\tau}(\xi) &= d(X, \xi) + \pi + d(-\xi, Y) \\
&= d(X, \xi) + \pi + \pi - d(\xi, Y) \\
&\geq 2\pi - d(X, Y).
\end{aligned}
$$

Here, the equality holds if and only if ξ lies on the geodesic segment $Xf^{-1}(Y)$. This proves (1) for the case when $p = q$.

Suppose $p \neq q$ and $\{p, X, q, Y, L\}$ satisfies the convexity condition. Then either $d(X, Z) \geq \arccos(D/L)$ or $d(Z, Y) \geq \arccos(D/L)$ holds, where $Z = \overrightarrow{pq}/|pq|$. If $d(X, Z) > \arccos(D/L)$ or $d(Z, Y) > \arccos(D/L)$ holds, (1) follows

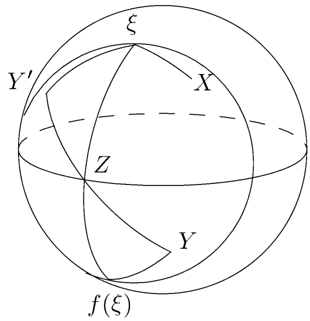


FIGURE 3.10. Illustration of the proof of Theorem 3.8.

from Lemma 3.7. Otherwise, we must have $d(X,Z) = d(Z,Y) = \arccos(D/L)$. Then we see that $\tilde{\tau}$ becomes minimal if and only if $\xi = X = f^{-1}(Y)$, which proves (1) for the case when $p \neq q$.

To prove (2) let $\xi_0 = \overrightarrow{pp_1}/|pp_1|$. Under the assumption in (2), $\{p, X, q, Y, L\}$ cannot satisfy the convexity condition. By Lemmas 3.4 and 3.5, we see that $\nabla\tilde{\tau} = 0$ at ξ_0. By Lemma 3.2, we have $d(Z,\xi_0) = \arccos(D/L)$, or equivalently, $|pp_1| = |p_1q| = L/2$. If we restrict the function $\tilde{\tau}(\xi)$ to the small circle $d_Z(\xi) = \arccos(D/L)$, it is written as

$$\tilde{\tau}(\xi) = d(X,\xi) + d(f(\xi), Y) + 2\arccos(D/L) \tag{3.29}$$

by (3.1). Hence, ξ_0 minimizes $d(X,\xi) + d(f(\xi), Y)$ among all ξ on the small circle $d_Z(\xi) = \arccos(D/L)$. Let Y' be the point on S^2 which is symmetric to Y with respect to Z. Then we have

$$d(X,\xi) + d(f(\xi), Y) = d(X,\xi) + d(\xi, Y'). \tag{3.30}$$

Both X and Y' are points in the disk D of radius $\arccos(D/L)$ centered at Z. If we regard D as a "two-sided" disk and regard X as a point on the "top" and Y' as a point on the "bottom", ξ_0 is the point where the shortest path from X to Y' in D meets the circle $d_Z(\xi) = \arccos(D/L)$. Now we use the proposition in the Appendix. If $X = Y = Z$ $(= Y')$, by (2) of the proposition, we can find the shortest path from X to Y' in all directions at X, which proves (2-2). In all other cases, there exist only one or two shortest paths from X to Y', and we have choices of the corresponding number for p_1. There are two choices for p_1 if and only if $Y' \in C_0$, where C_0 is the curve described in the Appendix. The condition that $Y' \in C_0$ is equivalent to the condition that

$$d(X,Z) = d(Y,Z), \quad 0 \leq \angle XZY < \pi - 2\theta_0, \tag{3.31}$$

where θ_0 is as given in the Appendix. (3.31) is equivalent to

$$\angle(X,\overrightarrow{pq}) = \angle(Y,\overrightarrow{pq}), \quad \angle(X,Y) < 2\arcsin\left(\frac{\sin\alpha\tan\alpha}{\tan R}\right), \tag{3.32}$$

where $\alpha = \angle(X,\overrightarrow{pq})$ and $R = \arccos(D/L)$. Now, (2) follows since we have

$$2\arcsin\left(\frac{\sin\alpha\tan\alpha}{\tan R}\right) = \arccos\left(\frac{L^2 - D^2 - 2d^2\sin^2\alpha\tan^2\alpha}{L^2 - D^2}\right).$$

In (3), $\{p,X,q,Y,L\}$ does not satisfy the convexity condition and we have $d(X,Z) \geq \arccos(D/L)$ and $d(X,Z) > d(Z,Y)$. If $d(X,Z) > \arccos(D/L)$, by Lemma 3.7, $\tilde{\tau}$ is not differentiable at ξ_0, and hence by Lemma 3.4, we have $\xi_0 = X$ or $\xi_0 = f^{-1}(Y)$. Then by Lemma 3.6, we see that $\xi_0 = X$. If $d(X,Z) = \arccos(D/L)$, ξ_0 may be a point with $d(\xi_0,Z) = \arccos(D/L)$. By (3.29) and (3.30), ξ_0 minimizes $d(X,\xi) + d(\xi,Y')$ for ξ with $d(\xi,Z) = \arccos(D/L)$. Hence, we must have $\xi_0 = X$, since $d(X,\xi)+d(\xi,Y') \geq d(X,Y')$ and the equality holds if and only if $\xi = X$. This proves (3). The proof for (4) is similar to that for (3).

In (5), $\{p,X,q,Y,L\}$ does not satisfy the convexity condition and we have $d(X,Z) = d(Z,Y) \geq \arccos(D/L)$. If $d(X,Z) = d(Z,Y) > \arccos(D/L)$, by Lemma 3.7 and Lemma 3.4, we have $\xi_0 = X$ or $\xi_0 = f^{-1}(Y)$. Since $\tilde{\tau}(X) = \tilde{\tau}(f^{-1}(Y))$ in this case, both of them minimize $\tilde{\tau}$. If $d(X,Z) = d(Z,Y) = \arccos(D/L)$, ξ_0 may be a point with $d(\xi_0,Z) = \arccos(D/L)$. By (3.29) and (3.30), ξ_0 minimizes $d(X,\xi) + d(\xi,Y')$ for ξ with $d(\xi,Z) = \arccos(D/L)$. Hence, both $\xi_0 = X$ and $\xi_0 = f^{-1}(Y)$ minimize $\tilde{\tau}$ since $d(X,\xi) + d(\xi,Y') \geq d(X,Y')$ and the equality holds if and only if $\xi = X$ or $\xi = Y'$. This proves (5). □

Remark 3.9. If we regard the closed case ($p = q$) as the limit ($|pq| \to 0$) of the open case ($p \neq q$) with $\overrightarrow{pq}/|pq| = -Y$, (1) for $p = q$ is derived from (1) for $p \neq q$.

4. Piecewise linear curves with three edges

In this section, we study the shape of a piecewise linear curve which attains $\inf\{\tilde{\tau}(P) : P \in \mathcal{P}_3(p,q,L)\}$. Since $\mathcal{P}_3(p,q,L)$ is compact and $\tilde{\tau}$ is continuous, such a piecewise linear curve always exists. Suppose that $P_0 = pp_1 \cup p_1p_2 \cup p_2q$ is an extremal curve in $\mathcal{P}_3(p,X,q,Y,L)$.

Let $P_0^1 = pp_1 \cup p_1p_2$ and $P_0^2 = p_1p_2 \cup p_2q$. Let $X_1 = \overrightarrow{pp_1}/|pp_1|$, $X_2 = \overrightarrow{p_1p_2}/|p_1p_2|$, $X_3 = \overrightarrow{p_2q}/|p_2q|$, $L_1 = |pp_1| + |p_1p_2|$ and $L_2 = |p_1p_2| + |p_2q|$. Let f_k ($k = 1,2$) be the map defined as f in Section 3 with $L = L_k$ and $\{p,q\} = \{p,p_2\}$ for f_1, $\{p,q\} = \{p_1,q\}$ for f_2. P_0^1 must be an extremal curve in $\mathcal{P}_2(p,X,p_2,X_3,L_1)$, since if not, the replacement of P_0^1 by an extremal curve in $\mathcal{P}_2(p,X,p_2,X_3,L_1)$ would produce an element of $\mathcal{P}_3(p,q,L)$ whose $\tilde{\tau}$ is smaller than $\tilde{\tau}(P_0)$, which contradicts the minimality of $\tilde{\tau}(P_0)$. By Theorem 3.8, one of the following holds:

(a1) $\{p, X, p_2, X_3, L_1\}$ satisfies the convexity condition and X_1, as a point in S^2, lies on the geodesic segment $Xf_1^{-1}(X_3)$.
(b1) $\{p, X, p_2, X_3, L_1\}$ does not satisfy the convexity condition and $|pp_1| = |p_1p_2|$.
(c1) $\{p, X, p_2, X_3, L_1\}$ does not satisfy the convexity condition and $X_1 = X$.
(d1) $\{p, X, p_2, X_3, L_1\}$ does not satisfy the convexity condition and $X_2 = X_3$.

Note that if (d1) occurs, then P_0 becomes a piecewise linear curve with two edges, and if (a1) occurs, then P_0 is a plane convex arc.

Similarly, P_0^2 must be an extremal curve in $\mathcal{P}_2(p_1, X_1, q, Y, L_2)$, and one of the following holds:

(a2) $\{p_1, X_1, q, Y, L_2\}$ satisfies the convexity condition and X_2, as a point in S^2, lies on the geodesic segment $X_1f_2^{-1}(Y)$.
(b2) $\{p_1, X_1, q, Y, L_2\}$ does not satisfy the convexity condition and $|p_1p_2| = |p_2q|$.
(c2) $\{p_1, X_1, q, Y, L_2\}$ does not satisfy the convexity condition and $X_2 = X_1$.
(d2) $\{p_1, X_1, q, Y, L_2\}$ does not satisfy the convexity condition and $X_3 = Y$.

If (c2) occurs, then P_0 becomes a piecewise linear curve with two edges, and if (a2) occurs, then P_0 is a plane convex arc.

Definition 4.1. A piecewise linear curve with three edges $P = pp_1 \cup p_1p_2 \cup p_2q$ is called a *Z-curve* (Figure 4.1) if P satisfies $\overrightarrow{pp_1}/|pp_1| = \overrightarrow{p_2q}/|p_2q|$.

We note that any Z-curve lies in a plane.

Lemma 4.2. *Suppose $X = Y$. If a Z-curve $P = pp_1 \cup p_1p_2 \cup p_2q$ with $\overrightarrow{pp_1}/|pp_1| = X$ and $\overrightarrow{p_2q}/|p_2q| = Y$ becomes an extremal curve in $\mathcal{P}_3(p, X, q, Y, L)$, then we have $\angle(X, \overrightarrow{pq}) \geq \arccos(D/L)$.*

Proof. Let $P' = pp_1' \cup p_1'q$ be the element of $\mathcal{P}_2(p, q, L)$ which satisfies $\overrightarrow{pp_1'}/|pp_1'| = X$. Since $X = Y$, we have $\tilde{\tau}(P) = \tilde{\tau}(P')$. If P is an extremal curve in $\mathcal{P}_3(p, X, q, Y, L)$, then P' must be an extremal curve in $\mathcal{P}_2(p, X, q, Y, L)$. By Theorem 3.8, we have $\angle(X, \overrightarrow{pq}) \geq \arccos(D/L)$. □

The main result in this section is stated as follows.

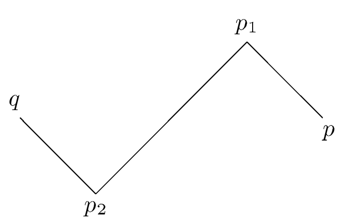


Figure 4.1. A Z-curve.

THEOREM 4.3. *Suppose that* $P_0 = pp_1 \cup p_1p_2 \cup p_2q$ *is an extremal curve in* $\mathcal{P}_3(p, X, q, Y, L)$*. Then* P_0 *must lie in a plane. Moreover,* P_0 *is actually an element of* $\mathcal{P}_2(p, q, L)$ *except for the following two cases:*

(1) $\{p, X, q, Y, L\}$ *satisfies the convexity condition and* P_0 *is a plane convex arc. (In this case,* p_1 *and* p_2 *can be any points such that* X*,* $\overrightarrow{pp_1}/|pp_1|$*,* $\overrightarrow{p_1p_2}/|p_1p_2|$*,* $\overrightarrow{p_2q}/|p_2q|$*,* Y*,* $\overrightarrow{qp}/|qp|$ *lie on a geodesic in* S^2 *in this order.)*

(2) $p \neq q$*,* $X = Y$*,* $\angle(X, \overrightarrow{pq}) \geq \arccos(D/L)$ *and* P_0 *is a Z-curve tangent to* X *at* p *and to* Y *at* q*.*

Proof. We define unit vectors Z_k $(k = 1, 2)$ by $Z_1 = \overrightarrow{pp_2}/|pp_2|$ and $Z_2 = \overrightarrow{p_1q}/|p_1q|$. In the following argument, $\angle_p(X_1, Z_1)$, for example, denotes the angle between X_1 and Z_1 at p. We also set $\gamma_1 = \arccos(d_1/L_1)$ and $\gamma_2 = \arccos(d_2/L_2)$, where $d_1 = |pp_2|$ and $d_2 = |p_1q|$.

Suppose (b1) and (b2) simultaneously hold for P_0. Since (b1) holds, we have

$$\begin{aligned} &\angle_p(X, Z_1) \leq \gamma_1, \\ &\angle_{p_2}(X_3, Z_1) \leq \gamma_1, \\ &\angle_p(X_1, Z_1) = \angle_{p_2}(X_2, Z_1) = \gamma_1. \end{aligned} \tag{4.1}$$

Since (b2) holds, we have

$$\begin{aligned} &\angle_{p_1}(X_1, Z_2) \leq \gamma_2, \\ &\angle_q(Y, Z_2) \leq \gamma_2, \\ &\angle_{p_1}(X_2, Z_2) = \angle_q(X_3, Z_2) = \gamma_2. \end{aligned} \tag{4.2}$$

Using (4.1) and (4.2), we obtain

$$\begin{aligned} \angle_{p_2}(X_3, Z_1) &\geq \angle_{p_2}(X_2, X_3) - \angle_{p_2}(X_2, Z_1) \\ &= 2\gamma_2 - \angle_{p_2}(X_2, Z_1) \\ &\geq \gamma_2 + \angle_{p_1}(X_1, Z_2) - \angle_{p_2}(X_2, Z_1) \\ &= \angle_{p_1}(X_2, Z_2) + \angle_{p_1}(X_1, Z_2) - \angle_{p_2}(X_2, Z_1) \\ &\geq \angle_{p_1}(X_1, X_2) - \angle_{p_2}(X_2, Z_1) \\ &= 2\gamma_1 - \angle_{p_2}(X_2, Z_1) \\ &= \gamma_1. \end{aligned} \tag{4.3}$$

From (4.1) and (4.3), we have

$$\angle_{p_2}(X_3, Z_1) = \gamma_1 \tag{4.4}$$

and all equalities in (4.3) must hold, i.e.,

$$\angle_{p_1}(X_1, Z_2) + \angle_{p_1}(X_2, Z_2) = \angle_{p_1}(X_1, X_2), \tag{4.5}$$

$$\angle_{p_2}(X_2, Z_1) + \angle_{p_2}(X_3, Z_1) = \angle_{p_2}(X_2, X_3), \tag{4.6}$$

$$\angle_{p_1}(X_1, Z_2) = \gamma_2. \tag{4.7}$$

(4.5) and (4.6) imply that P_0 lies in a plane. If $\angle_p(X, Z_1) < \gamma_1$, (4.4) implies that p_1p_2 is tangent to X_3, which means that P_0 is an element of $\mathcal{P}_2(p, q, L)$. Hence, if P_0 is not an element of $\mathcal{P}_2(p, q, L)$, by (4.1), we must have $\angle_p(X, Z_1) = \gamma_1$ and pp_1 is tangent to X. By a similar reason, p_2q is tangent to Y. Combining (4.2), (4.5), and (4.7), we obtain

$$\begin{aligned} 2\gamma_2 &= \angle_{p_1}(X_1, Z_2) + \angle_{p_1}(X_2, Z_2) \\ &= \angle_{p_1}(X_1, X_2) \\ &= 2\gamma_1. \end{aligned} \tag{4.8}$$

Since $\{p, X, p_2, X_3, L_1\}$ does not satisfy the convexity condition, (4.8) implies that $X = Y$, and P is a Z-curve which is tangent to X at p and tangent to Y at q.

Suppose (c1) and (b2) simultaneously hold for P_0. Since (c1) holds, we have

$$\begin{aligned} &\angle_p(X, Z_1) \geq \gamma_1, \\ &\angle_p(X, Z_1) \geq \angle_{p_2}(X_3, Z_1). \end{aligned} \tag{4.9}$$

Using (4.2) and (4.9), we obtain

$$\begin{aligned} \angle_{p_2}(X_3, Z_1) &\geq \angle_{p_2}(X_2, X_3) - \angle_{p_2}(X_2, Z_1) \\ &= 2\gamma_2 - \angle_{p_2}(X_2, Z_1) \\ &\geq \gamma_2 + \angle_{p_1}(X_1, Z_2) - \angle_{p_2}(X_2, Z_1) \\ &= \gamma_2 + \angle_{p_1}(X_1, Z_2) - \angle_{p_1}(X_1, X_2) + \angle_p(X_1, Z_1) \\ &\geq \gamma_2 + \angle_{p_1}(X_1, Z_2) - \angle_{p_1}(X_2, Z_2) - \angle_{p_1}(X_1, Z_2) \\ &\quad + \angle_p(X_1, Z_1) \\ &= \angle_p(X_1, Z_1) \\ &= \angle_p(X, Z_1). \end{aligned} \tag{4.10}$$

From (4.9) and (4.10), we have

$$\angle_{p_2}(X_3, Z_1) = \angle_p(X, Z_1) \tag{4.11}$$

and all equalities in (4.10) must hold, i.e.,

$$\angle_{p_1}(X_1, Z_2) + \angle_{p_1}(X_2, Z_2) = \angle_{p_1}(X_1, X_2), \tag{4.12}$$

$$\angle_{p_2}(X_2, Z_1) + \angle_{p_2}(X_3, Z_1) = \angle_{p_2}(X_2, X_3), \tag{4.13}$$

$$\angle_{p_1}(X_1, Z_2) = \gamma_2. \tag{4.14}$$

(4.12) and (4.13) imply that P lies in a plane. If $\angle_q(Y, Z_2) < \gamma_2$, (4.7) implies that p_1p_2 is tangent to X_1, which means that P_0 is an element of $\mathcal{P}_2(p, q, L)$. Hence, if P_0 is not an element of $\mathcal{P}_2(p, q, L)$, by (4.2), we must have $\angle_p(Y, Z_2) = \gamma_2$ and p_2q is tangent to Y. Since $X = X_1$, (4.11) implies that pp_1 is parallel to p_2q. Now we see that $X = Y$ and P_0 is a Z-curve which is tangent to X at p and tangent to Y at q.

By a similar reasoning, we see that if (b1) and (d2) simultaneously hold for P_0, then P_0 must be a Z-curve which is tangent to X at p and tangent to Y at q.

Suppose (c1) and (d2) simultaneously hold for P_0. Since (d2) holds, we have

$$\begin{aligned}&\angle_q(Y, Z_2) \geq \gamma_2,\\ &\angle_q(Y, Z_2) \geq \angle_{p_1}(X_1, Z_2).\end{aligned} \tag{4.15}$$

Using (4.9) and (4.15), we obtain

$$\begin{aligned}\angle_{p_1}(X_1, Z_2) &\geq \angle_{p_1}(X_1, X_2) - \angle_{p_1}(X_2, Z_2)\\ &= \angle_{p_1}(X_1, X_2) - \angle_{p_2}(X_2, X_3) + \angle_q(X_3, Z_2)\\ &= \angle_{p_1}(X_1, X_2) - \angle_{p_2}(X_2, X_3) + \angle_q(Y, Z_2)\\ &\geq \angle_{p_1}(X_1, X_2) - \angle_{p_2}(X_3, Z_1) - \angle_{p_2}(X_2, Z_1)\\ &\quad + \angle_q(Y, Z_2)\\ &= \angle_p(X_1, Z_1) + \angle_{p_2}(X_2, Z_1) - \angle_{p_2}(X_3, Z_1)\\ &\quad - \angle_{p_2}(X_2, Z_1) + \angle_q(Y, Z_2)\\ &= \angle_p(X_1, Z_1) - \angle_{p_2}(X_3, Z_1) + \angle_q(Y, Z_2)\\ &= \angle_p(X, Z_1) - \angle_{p_2}(X_3, Z_1) + \angle_q(Y, Z_2)\\ &\geq \angle_q(Y, Z_2).\end{aligned} \tag{4.16}$$

From (4.15) and (4.16), we have

$$\angle_{p_1}(X_1, Z_2) = \angle_q(Y, Z_2) \tag{4.17}$$

and all equalities in (4.16) must hold, i.e.,

$$\angle_{p_1}(X_1, Z_2) + \angle_{p_1}(X_2, Z_2) = \angle_{p_1}(X_1, X_2), \tag{4.18}$$

$$\angle_{p_2}(X_2, Z_1) + \angle_{p_2}(X_3, Z_1) = \angle_{p_2}(X_2, X_3), \tag{4.19}$$

$$\angle_p(X, Z_1) = \angle_{p_2}(X_3, Z_1). \tag{4.20}$$

(4.18) and (4.19) imply that P lies in a plane. Since $X = X_1$ and $\{p, X, p_2, X_3, L_1\}$ does not satisfy the convexity condition, (4.20) implies that pp_1 is parallel to p_2q. Since pp_1 is tangent to X and p_2q is tangent to Y, we see that $X = Y$ and P_0 is a Z-curve which is tangent to X at p and tangent to Y at q.

Now we see that one of (a1), (d1), (a2), and (c2) must hold for P_0 unless P_0 is a Z-curve tangent to X at p and to Y at q. Any of (a1), (d1), (a2), and (c2) implies that P_0 lies in a plane.

Having (d1) or (c2) means that P_0 is actually an element of $\mathcal{P}_2(p, q, L)$.

Suppose that (a1) and (b2) hold simultaneously for P_0. By (a1), P_0 is a convex arc and we have $\angle_{p_1}(X_1, Z_2) > \angle_{p_1}(X_2, Z_2) = \gamma_2$, while (b2) gives $\angle_{p_1}(X_1, Z_2) \leq \gamma_2$. Hence, (a1) and (b2) cannot simultaneously hold for P_0. Similarly, (b1) and (a2) cannot simultaneously hold for P_0.

(a1) and (d2) cannot simultaneously hold for P_0, because if this happens, $Y = X_3$ holds and $\{p_1, X_1, q, Y, L_2\}$ satisfies the convexity condition, which contradicts (d2). Similarly, (c1) and (a2) cannot simultaneously hold for P_0.

The remaining case is when (a1) and (a2) simultaneously hold for P_0. To prove that $\{p, X, q, Y, L\}$ satisfies the convexity condition, we take p_2' so that $\overrightarrow{p_1p_2'}|p_1p_2'| = X_1$. Then $pp_2' \cup p_2'q$ becomes an element of $\mathcal{P}_2(p, q, L)$ whose $\tilde{\tau}$ is equal to $\tilde{\tau}(P_0)$. Since we can move p_1 continuously preserving $\tilde{\tau}$, this construction gives infinitely many elements of $\mathcal{P}_2(p, q, L)$ which are extremal curves in $\mathcal{P}_2(p, X, q, Y, L)$. Such a thing cannot happen unless $\{p, X, q, Y, L\}$ satisfies the convexity condition, by Theorem 3.8. Now we prove that P_0 is a plane convex arc. (a1) implies that we have $\{X, X_1, X_2, X_3\}$ on a great circle in S^2 in this order. (a2) implies that we have $\{X_1, X_2, X_3, Y\}$ on the same great circle in this order. Since $\{p, X, q, Y, L\}$ satisfies the convexity condition, we have $\{X, Z, Y, -Z\}$ on the same great circle in this order. Thus, on this great circle, we have $\{X_1, X_2, X_3, -Z\}$ in this order. This implies that the quadrilateral pp_1p_2q is convex, and hence P_0 is a plane convex arc.

Now we have checked all possible combinations between {(a1), (b1), (c1), (d1)} and {(a2), (b2), (c2), (d2)}, and see we have only three possibilities for P_0:

(i) $\{p, X, q, Y, L\}$ satisfies the convexity condition and P_0 is a plane convex arc.
(ii) $X = Y$ and P_0 is a Z-curve with $\overrightarrow{pp_1}/|pp_1| = X$ and $\overrightarrow{p_2q}/|p_2q| = Y$.
(iii) P_0 is actually an element of $\mathcal{P}_2(p, q, L)$.

If $X = Y$, then $\{p, X, q, Y, L\}$ does not satisfy the convexity condition. If $X = Y$ and $L < \arccos(D/L)$, then we see from Lemma 4.2 that P_0 cannot be a Z-curve, and hence P_0 must be an element of $\mathcal{P}_2(p, q, L)$. If $X = Y$ and $L \geq \arccos(D/L)$, then we see from Theorem 3.8(5) that, among all elements of $\mathcal{P}_2(p, q, L)$, $pp_1 \cup p_1q$ with $\overrightarrow{pp_1}/|pp_1| = X$ and $pp_1 \cup p_1q$ with $\overrightarrow{p_1q}/|p_1q| = Y$ minimize $\tilde{\tau}$. Since $X = Y$, any Z-curve tangent to X at p and tangent to Y at q has the same value of $\tilde{\tau}$ as these curves. Thus we see that Z-curves tangent to X at p and tangent to Y at q and the elements of $\mathcal{P}_2(p, q, L)$ given in Theorem 3.8(5) are extremal curves in this case. □

5. Piecewise linear curves with n edges and piecewise smooth curves

In this section, we study the shape of a piecewise linear curve which attains $\inf\{\tilde{\tau}(P) : P \in \mathcal{P}_n(p, q, L)\}$. Since $\mathcal{P}_n(p, q, L)$ is compact and $\tilde{\tau}$ is continuous, such a piecewise linear curve always exists.

THEOREM 5.1. *Suppose that $P_0 \in \mathcal{P}_n(p, q, L)$ is an extremal curve in $\mathcal{P}_n(p, X, q, Y, L)$. Then P_0 must lie in a plane. Moreover, P_0 is actually an element of $\mathcal{P}_2(p, q, L)$, except for the following two cases:*

(1) $\{p, X, q, Y, L\}$ *satisfies the convexity condition and* P_0 *is a plane convex arc.*
(2) $p \neq q$, $X = Y$, $\angle(X, \overrightarrow{pq}) \geq \arccos(D/L)$ *and* P_0 *is a Z-curve tangent to* X *at* p *and to* Y *at* q.

Proof. If $n \leq 3$, this theorem follows from Theorem 4.3. Hence, it suffices to show that if $n \geq 4$ and $P_0 \in \mathcal{P}_n(p, q, L)$ is not an element of $\mathcal{P}_{n-1}(p, q, L)$, then P_0 must be a plane convex arc. We write P_0 as $P_0 = p_0p_1 \cup p_1p_2 \cup \cdots \cup p_{n-2}p_{n-1} \cup p_{n-1}p_n$ with $p_0 = p$, $p_n = q$. Here, we fix some notations. Let

$$P_0^k = p_{k-1}p_k \cup p_kp_{k+1} \cup p_{k+1}p_{k+2} \quad (k = 1, \ldots, n-2),$$
$$L_k = |p_{k-1}p_k| + |p_kp_{k+1}| + |p_{k+1}p_{k+2}| \quad (k = 1, \ldots, n-2),$$
$$X_k = \overrightarrow{p_{k-1}p_k}/|p_{k-1}p_k| \quad (k = 1, \ldots, n), \qquad X_0 = X, \qquad X_{n+1} = Y.$$

Since P_0 is an extremal curve in $\mathcal{P}_n(p, X, q, Y, L)$, for every k with $1 \leq k \leq n-2$, P_0^k must be an extremal curve in $\mathcal{P}_3(p_{k-1}, X_{k-1}, p_{k+2}, X_{k+3}, L_k)$. By assumption, P_0^k is not an element of $\mathcal{P}_2(p_{k-1}, p_{k+2}, L_k)$. Then by Theorem 4.3, we have two possibilities:

(i) $\{p_{k-1}, X_{k-1}, p_{k+2}, X_{k+3}, L_k\}$ satisfies the convexity condition and we have $\{X_{k-1}, X_k, X_{k+1}, X_{k+2}, X_{k+3}\}$ on a great circle in this order.
(ii) $X_{k-1} = X_{k+3}$ and P_0^k is a Z-curve with $X_k = X_{k-1}$ and $X_{k+2} = X_{k+3}$.

Since $n \geq 4$, if (ii) occurs for some k, then P_0 becomes an element of $\mathcal{P}_{n-1}(p, q, L)$, which contradicts our assumption. Thus, we must have (i) for all k. Then $\{X_{k-1}, \ldots, X_{k+3}\}$, as points in S^2, lie on a great circle in this order. Hence, $\{X, X_1, \ldots, X_n, Y\}$ lie on a great circle in this order, which implies that P_0 lies in a plane. To prove that P_0 is a plane convex arc, it suffices to show that we have $\{X_1, \ldots, X_n, -Z\}$ on a great circle in this order. As we see in the proof for Theorem 4.3, there exists p'_k such that $p_{k-1}p'_k \cup p'_kp_{k+2}$ becomes an extremal curve in $\mathcal{P}_3(p_{k-1}, X_{k-1}, p_{k+2}, X_{k+3}, L_k)$. Moreover, there are infinitely many choices for p'_k. After replacing P_0^k by $p_{k-1}p'_k \cup p'_kp_{k+2}$, the curve is still an extremal curve in $\mathcal{P}_n(p, X, q, Y, L)$. Hence, if it is not an element of $\mathcal{P}_2(p, X, q, Y, L)$, we can find an extremal curve in $\mathcal{P}_n(p, X, q, Y, L)$ with fewer vertices by a similar replacement. We may repeat this until we obtain an element of $\mathcal{P}_2(p, X, q, Y, L)$ which is an extremal curve in $\mathcal{P}_n(p, X, q, Y, L)$. This curve is an extremal curve in $\mathcal{P}_2(p, X, q, Y, L)$ as well and there are infinitely many choices for such a curve. Such a thing cannot happen in a plane unless $\{p, X, q, Y, L\}$ satisfies the convexity condition. This implies that we have $\{X, Z, Y, -Z\}$ on a great circle in this order, and hence we have $\{X_1, \ldots, X_n, -Z\}$ on a great circle in this order. □

Remark 5.2. If $\{p, X, q, Y, L\}$ satisfies the convexity condition and if $P_0 = pp_1 \cup p_1p_2 \cup \cdots \cup p_{n-2}p_{n-1} \cup p_{n-1}q$ is an extremal curve in $\mathcal{P}_n(p, X, q, Y, L)$, P_0 must be a plane convex arc. Moreover, as we see in the proof, $p_1, \ldots, p_{n-1}$ can be any points which satisfy the following conditions:

(i) $|pp_1| + |p_1p_2| + \cdots + |p_{n-2}p_{n-1}| + |p_{n-1}q| = L$.
(ii) $\{X, \overrightarrow{pp_1}/|pp_1|, \overrightarrow{p_1p_2}/|p_1p_2|, \ldots, \overrightarrow{p_{n-1}q}/|p_{n-1}q|, Y, \overrightarrow{qp}/|qp|\}$ lies on a great circle in this order.

Now we state our main result of the present paper.

THEOREM 5.3. *Suppose that* Σ_0 *an extremal curve in* $\mathcal{C}(p, X, q, Y, L)$. *Then* Σ_0 *must lie in a plane. Moreover,* Σ_0 *is actually an element of* $\mathcal{P}_2(p, q, L)$ *except for the following two cases:*

(1) $\{p, X, q, Y, L\}$ *satisfies the convexity condition and* Σ_0 *is a plane convex arc.*
(2) $p \neq q$, $X = Y$, $\angle(X, \overrightarrow{pq}) \geq \arccos(D/L)$ *and* Σ_0 *is a Z-curve tangent to* X *at* p *and to* Y *at* q.

Proof. Suppose that Σ_0 in $\mathcal{C}(p, q, L)$ satisfies $\tilde{\tau}(\Sigma_0) = \inf\{\tilde{\tau}(\Sigma) : \Sigma \in \mathcal{C}(p, q, L)\}$ but does not lie in any plane. If $x(s)$ $(0 \leq s \leq L)$ is a parameterization of Σ_0 by arclength, we can construct a piecewise linear curve with n edges by connecting $x((i-1)L/n)$ and $x(iL/n)$ for $i = 1, \ldots, n$. By taking a point p_i in a neighborhood of $x(iL/n)$ for $i = 1, \ldots, n-1$, we can construct another piecewise linear curve P_n whose length is L. Since Σ_0 does not lie in a plane, P_n does not lie in a plane for some n. Then P_n cannot be an extremal curve in $\mathcal{P}_n(p, X, q, Y, L)$ by Theorem 5.1 and we must have

$$\tilde{\tau}(P_n) > \inf\{\tilde{\tau}(P) : P \in \mathcal{P}_n(p, q, L)\} + \delta \tag{5.1}$$

for some $\delta > 0$. For any positive constant $\varepsilon < \delta$, if we take n sufficiently large, we have

$$\tilde{\tau}(P_n) < \tilde{\tau}(\Sigma_0) + \varepsilon. \tag{5.2}$$

It follows from (5.1) and (5.2) that

$$\begin{aligned} \inf\{\tilde{\tau}(P) : P \in \mathcal{P}_n(p, q, L)\} &< \tilde{\tau}(P_n) - \delta \\ &< \tilde{\tau}(\Sigma_0) + \varepsilon - \delta \\ &< \tilde{\tau}(\Sigma_0) \\ &= \inf\{\tilde{\tau}(\Sigma) : \Sigma \in \mathcal{C}(p, q, L)\}. \end{aligned}$$

This is a contradiction, since we must have

$$\inf\{\tilde{\tau}(P) : P \in \mathcal{P}_n(p, q, L)\} \geq \inf\{\tilde{\tau}(\Sigma) : \Sigma \in \mathcal{C}(p, q, L)\}.$$

Thus Σ_0 must lie in a plane.

Suppose that none of (1)–(3) holds. Then none of (1)–(3) in Theorem 5.1 holds for P_n, an element of $\mathcal{P}_n(p, q, L)$ which is defined from Σ_0 as above. Since P_n is not an extremal curve in $\mathcal{P}_n(p, X, q, Y, L)$, we must have (5.1) again for some $\delta > 0$. We also have (5.2) again for any positive constant $\varepsilon < \delta$, if we take n sufficiently large. This gives a contradiction again. Thus, (1)–(3) must hold for Σ_0. □

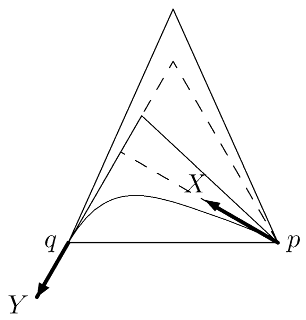


FIGURE 5.1. The shapes of extremal curves as a function of their length.

COROLLARY 5.4. *For any* $\{p, X, q, Y, L\}$*, there exists an element of* $\mathcal{P}_2(p, q, L)$ *which is an extremal curve in* $\mathcal{C}(p, X, q, Y, L)$.

COROLLARY 5.5. *If* X*,* Y *and* $\overrightarrow{pq}$ *do not lie in a plane, then every extremal curve in* $\mathcal{C}(p, X, q, Y, L)$ *is an element of* $\mathcal{P}_2(p, q, L)$.

REMARK 5.6. Theorem 5.3 generalizes the results in [8], where a similar problem is considered for curves in E^2.

REMARK 5.7. One may apply Theorem 5.3 for closed curves by setting $p = q$ and $X = Y$. As a result, one can derive the classical Fenchel's theorem from Theorem 5.3.

REMARK 5.8. Even if we fix p, q, X, and Y, the shape of the extremal curves in $\mathcal{C}(p, X, q, Y, L)$ changes by the length L (Figure 5.1). For example, let $p = (1, 0, 0)$, $q = (-1, 0, 0)$, $X = (\cos(5\pi/6), \sin(5\pi/6), 0)$ and $Y = (\cos(4\pi/3), \sin(4\pi/3), 0)$. If $2 < L < 1+\sqrt{3}$, $\{p, X, q, Y, L\}$ satisfies the convexity condition and there are infinitely many convex arcs in $\mathcal{C}(p, q, L)$ which become extremal curves in $\mathcal{C}(p, X, q, Y, L)$. If $L = 1+\sqrt{3}$, $\{p, X, q, Y, L\}$ satisfies the convexity condition but $pp_1 \cup p_1 q$ with $\overrightarrow{pp_1}/|pp_1| = X$ and $\overrightarrow{p_1 q}/|p_1 a| = Y$ is the only extremal curve in $\mathcal{C}(p, X, q, Y, L)$. If $1 + \sqrt{3} < L \leq 4$, $\{p, X, q, Y, L\}$ satisfies the condition in Theorem 3.8(3) and $pp_1 \cup p_1 q$ with $\overrightarrow{pp_1}/|pp_1| = X$ is the only extremal curve in $\mathcal{C}(p, X, q, Y, L)$. If $L > 4$, $\{p, X, q, Y, L\}$ satisfies the condition in Theorem 3.8(2-1) and $pp_1 \cup p_1 q$ with $p_1 = (0, \sqrt{L^2 - 4}/2, 0)$ ($|pp_1| = |p_1 q|$) is the only extremal curve in $\mathcal{C}(p, X, q, Y, L)$.

REMARK 5.9. If we do not fix the length, the problem becomes almost trivial. Let us denote by $\mathcal{C}(p, X, q, Y)$ the set of all piecewise C^2 curves in E^3 whose endpoints and end-directions are p, q, X, Y. For any $\Sigma \in \mathcal{C}(p, X, q, Y)$, let $\bar{\Sigma}$ be the closed curve defined by $\Sigma \cup qp$. Then we have $\tau(\bar{\Sigma}) = \tau(\Sigma) + \angle(Y, \overrightarrow{qp}/|qp|) + \angle(\overrightarrow{qp}/|qp|, X)$. Since $\tau(\bar{\Sigma}) \geq 2\pi$, we have $\tau(\Sigma) \geq 2\pi - \angle(Y, \overrightarrow{qp}/|qp|) - \angle(\overrightarrow{qp}/|qp|, X) = \angle(X, \overrightarrow{pq}/|pq|) + \angle(\overrightarrow{pq}/|pq|, Y)$. Thus, we see that $\inf\{\tau(\Sigma) : \Sigma \in \mathcal{C}(p, X, q, Y)\} = \angle(X, \overrightarrow{pq}/|pq|) + \angle(\overrightarrow{pq}/|pq|, Y)$ and the infimum is attained by $\Sigma = pq$.

Appendix. Cut locus of a two-sided disk in the unit 2-sphere

Let C be a small circle of radius R ($R < \pi/2$) in the unit 2-sphere. Let D be the "two-sided" round disk bounded by C. We denote the top of D by D_1 and the bottom by D_2. Let $\pi : D \to D_1$ be the natural projection. Let d be the distance on D which is naturally induced from the standard metric of the unit 2-sphere. For $x \in D_1$ and $y \in D_2$, we have

$$d(x, y) = \min\{d(x, \xi) + d(\xi, y); \xi \in C\}.$$

We will see how many different points on C can attain $d(x, y)$. If there exist more than one points as such, y is a cut point of x. Let Z_i ($i = 1, 2$) be the center of D_i. We now assume that $x \neq Z_1$. For $\eta \in D_1 \backslash \{Z_1\}$ let $\theta(\eta)$ be the oriented angle ($0 \leq \theta(\eta) < 2\pi$) from the geodesic $Z_1 x$ to $Z_1 \eta$. Let ξ_0 be the point on C with $0 < \theta(\xi_0) < \pi/2$ such that the geodesic $x\xi_0$ intersects $Z_1 x$ perpendicularly at x. Let $\theta_0 = \theta(\xi_0)$. θ_0 is explicitly given by

$$\theta_0 = \arccos \frac{\tan \alpha}{\tan R},$$

where $\alpha = d(x, Z_1)$. We extend θ to a function in D by defining $\theta(y)$ for $y \in D_2$ as $\theta(\pi(y))$. Let (Figure A.1)

$$C_0 = \{y \in D_2 : d(y, Z_2) = \alpha, 2\theta_0 < \theta(y) < 2\pi - 2\theta_0\}.$$

The main result in this appendix is stated as follows.

PROPOSITION.

(1) *Suppose $x \in D_1$ and $x \neq Z_1$. If $y \in C_0$, two geodesics from x to y realize $d(x, y)$. If $y \notin C_0$, only one geodesic from x to y realizes $d(x, y)$.*

(2) *Suppose $x = Z_1$. If $y = Z_2$, every geodesic $x\xi \cup \xi y$ ($\xi \in C$) realizes $d(x, y)$. If $y \neq Z_2$, only one geodesic from x to y realizes $d(x, y)$.*

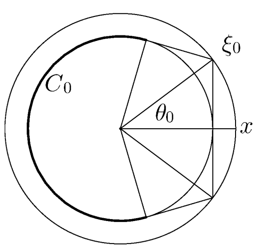


FIGURE A.1. The set C_0.

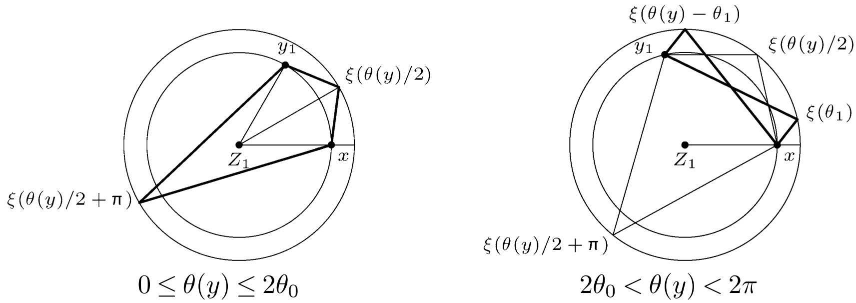


FIGURE A.2. The two cases of the proof of the proposition.

Proof. Let $x \in D_1$ and $y \in D_2$. Since C is parameterized by θ, we regard $d(x,\xi)+d(\xi,y)$ for $\xi \in C$ as a function in θ, and denote it by $f(\theta)$. Setting $y_1 = \pi(y)$, $f(\theta)$ is written as

$$f(\theta) = d(x,\xi(\theta)) + d(\xi(\theta), y_1).$$

Then

$$f'(\theta) > 0 \quad (=0, <0)$$

if and only if

$$\angle(\nabla d_x(\xi(\theta)), \nabla d_{Z_1}(\xi(\theta))) - \angle(\nabla d_{y_1}(\xi(\theta)), \nabla d_{Z_1}(\xi(\theta))) > 0 \quad (=0, <0),$$

respectively. Suppose $x \neq Z_1$ and let y be a point which satisfies $d(y,Z_2) = \alpha$. Then we have $f'(\theta(y)/2) = 0$ and $f'(\theta(y)/2+\pi) = 0$. These are the only critical points of $f(\theta)$ if $0 \le \theta(y) \le 2\theta_0$ or $2\pi - 2\theta_0 \le \theta(y) \le 2\pi$, while there are two more critical points θ_1 $(0 < \theta_1 < \theta(y)/2)$ and $\theta(y) - \theta_1$ if $2\theta_0 < \theta(y) < 2\pi - 2\theta_0$ (Figure A.2). $f(\theta)$ becomes minimal at $\theta = \theta(y)/2$ if $0 \le \theta(y) \le 2\theta_0$ or $2\pi - 2\theta_0 \le \theta(y) \le 2\pi$, while $f(\theta)$ becomes minimal at $\theta = \theta_1$ and $\theta = \theta(y) - \theta_1$ if $2\theta_0 < \theta(y) < 2\pi - 2\theta_0$.

This means that for every $y \in C_0$ there are two distinct minimizing geodesics from x to y. Let

$$\bar{C}_0 = \{y \in D_2 : d(y,Z_2) = \alpha, 2\theta_0 \le \theta(y) \le 2\pi - 2\theta_0\}.$$

Then we see that every point η in D lies on a minimizing geodesic from x to some point y in $\bar{C}_0$. If $\eta \neq y$, the subarc of the geodesic from x to y is the only minimizing geodesic from x to η. In fact, if there is another minimizing geodesic from x to η, we can find a shortcut to construct a path from x to y shorter than the original minimizing geodesic, which is a contradiction. Hence, for every $\eta \in D \setminus \bar{C}_0$ there exists only one minimizing geodesic from x to η. If $y \in \bar{C}_0 \setminus C_0$, $\theta(y) = 2\theta_0, 2\pi - 2\theta_0$ and there exists only one minimizing geodesic from x to y. Thus, there are two distinct minimizing geodesics from x to y if $y \in C_0$, while there is only one minimizing geodesic from x to y if $y \notin C_0$. This completes the proof for (1).

FIGURE A.3. A "lens surface" with an approximation to the cut locus [12] from a point below the equator on the far side.

If $x = Z_1$ and $y = Z_2$, then $f(\theta)$ becomes a constant $2R$. This means that any geodesic $Z_1\xi \cup \xi Z_2$ ($\xi \in C$) is minimizing. Any $y \neq Z_2$ lies on some minimizing geodesic from Z_1 to Z_2, and the subarc of that geodesic is the only minimizing geodesic from Z_1 to y. This proves (2). □

REMARK. The problem of determining the structure of the cut loci of the 2-sided disk in S^2 may be studied from a slightly different point of view. If we flip D_2 and attach D_1 and D_2 along the small circle C, then the resulting surface (which may be called a "lens surface", Figure A.2) becomes a closed convex surface of revolution (which is not smooth along the small circle). The cut locus of $x \in D_1$ should appear in this surface in the same way as in our 2-sided disk. Put simply, our proposition claims that the cut locus of a point in the "southern hemisphere" appears as a subarc of a circle of constant latitude in the northern hemisphere.

The reader is invited to compare such a construction of two smooth surfaces sewn with a distribution of curvature along the boundary with the case of billiards. A beautiful discussion of the conjugate loci of circular billiard tables is to be found in [2].

REMARK. The problem of determining the structure of the cut loci of the 2-sided disk in S^2 may be studied from a slightly different point of view. If we attach two round disks D_1 and D_2 along the small circle C so that the resulting surface becomes a closed convex surface of revolution (which is not smooth along the small circle), the cut locus of $x \in D_1$ should appear in this surface in the same way as in our 2-sided disk. If one smooths the surface along the small circle, one obtains a closed surface of revolution with positive Gaussian curvature. The structure of cut loci is studied on some smooth surfaces of revolution [20], or general ellipsoids [13], and the behaviors of geodesics in our case are quite similar to their case. The results in their papers also suggest that the cut locus of x is a subarc of the small circle $\{y \in D_2 : d(y, Z_2) = d(x, Z_1)\}$.

Acknowledgments. The authors express their thanks to Minoru Tanaka for his valuable suggestions. We are also very grateful to the referee for his many valuable comments.

References

[1] P. A. Alexander, Y. He, Y. Chen, J. Orban and P. N. Bryan, *The design and characterization of two proteins with 88% sequence identity but different structure and function*, Proc. Natl. Acad. Sci. USA **104** (2007), 11963–11968.

[2] R. L. Bishop, *Circular billiard tables, conjugate loci, and a cardioid*, Regul. Chaotic Dyn. **8** (2003), 83–95. MR 1963970

[3] A. A. Canutescu and R. L. Dunbrack, Jr., *Cyclic coordinate descent: A robotics algorithm for protein loop closure*, Protein Sci. **12** (2003), 963–972.

[4] A. Carbone and M. Gromov, *Mathematical slices of molecular biology*, Gaz. Math. **88**, suppl. (2001), 1–80. MR 1834502

[5] G. S. Chirikjian and J. W. Burdick, *Kinematically optimal hyper-redundant manipulator configurations*, IEEE Trans. Robot. Autom. **11** (1995), 794–806.

[6] V. Daggett and A. Fersht, *The present view of the mechanism of protein folding*, Nat. Rev. Mol. Cell Biol. **4** (2003), 497–502.

[7] C. M. Dobson, *Protein folding and misfolding*, Nature **426** (2003), 884–890.

[8] K. Enomoto, *The total absolute curvature of open plane curves of fixed length*, Yokohama Math. J. **48** (2000), 83–96. MR 1788833

[9] K. Enomoto and J. Itoh, *The total absolute curvature of nonclosed curves in S^2*, Results Math. **45** (2004), 21–34. MR 2044010

[10] K. Enomoto and J. Itoh, *The total absolute curvature of nonclosed curves in S^2 (II)*, Results Math. **45** (2004), 230–240. MR 2078450

[11] W. Fenchel, *Über Krümmung und Windung geschlossener Raumkurven*, Math. Ann. **101** (1929), 238–252. MR 1512528

[12] J. Itoh and R. Sinclair, *Thaw: A tool for approximating cut loci on a triangulation of a surface*, Experiment. Math. **13** (2004), 309–325. MR 2103329

[13] J. Itoh and K. Kiyohara, *The cut loci and conjugate loci on ellipsoids*, Manuscripta Math. **114** (2004), 247–264. MR 2067796

[14] W. Jin, O. Kambara, H. Sasakawa, A. Tamura and S. Takada, *De novo design of foldable proteins with smooth folding tunnel: Automated negative design and experimental verification*, Structure **11** (2003), 581–590.

[15] B. Kuhlman, G. Dantas, G. C. Ireton, G. Vanini, B. L. Stoddard and D. Baker, *Design of a novel globular protein fold with atomic-level accuracy*, Science **302** (2003), 1364–1368.

[16] J. Milnor, *On the total curvature of knots*, Ann. Math. **52** (1953), 248–257. MR 0037509

[17] J. Milnor, *On total curvatures of closed space curves*, Math. Scand. **1** (1953), 289–296. MR 0059030

[18] S. B. Prusiner, *Shattuck lecture—neurodegenerative diseases and prions*, N. Engl. J. Med. **344** (2001), 1516–1526.

[19] J. W. H. Schymkowitz, F. Rousseau and L. Serrano, *Surfing on protein folding energy landscapes*, Proc. Natl. Acad. Sci. USA **99** (2002), 15846–15848.

[20] R. Sinclair and M. Tanaka, *The cut locus of a two-sphere of revolution and Toponogov's comparison theorem*, Tohoku Math. J. **59** (2007), 379–399. MR 2365347

[21] E. L. Starostin, *On the writhing number of a non-closed curve*, Physical and numerical models in knot theory including applications to the life sciences (J. A. Calvo et al., eds.), Series on Knots and Everything, vol. 36, World Scientific Publishing, Singapore, 2005, pp. 525–546, Chapter 26. MR 2197956

[22] C. Truesdell, *The influence of elasticity on analysis; the classical heritage*, Bull. Amer. Math. Soc. **9** (1983), 293–310. MR 0714991

[23] P. G. Wolynes, *Energy landscapes and solved protein-folding problems*, Philos. Trans. R. Soc. Lond. Ser. A **363** (2005), 453–467.

[24] J. M. Zimmerman and L. J. Maher III, *Solution measurement of DNA curvature in papillomavirus E2 binding sites*, Nucl. Acids Res. **31** (2003), 5134–5139.

Kazuyuki Enomoto, Faculty of Industrial Science and Technology, Tokyo University of Science, Oshamambe, Hokkaido 049-3514, Japan

E-mail address: enomoto@rs.kagu.tus.ac.jp

Jin-ichi Itoh, Faculty of Education, Kumamoto University, Kumamoto, Kumamoto 860-8555, Japan

E-mail address: j-itoh@gpo.kumamoto-u.ac.jp

Robert Sinclair, Mathematical Biology Unit, Okinawa Institute of Science and Technology, Okinawa Industrial Technology Center Annex, 12-2 Suzaki, Uruma, Okinawa 904-2234, Japan

E-mail address: sinclair@oist.jp