Curvature-Constrained Shortest Paths in a Convex Polygon

Pankaj K. Agarwal, Thérèse Biedl, Sylvain Lazard, Steve Robbins, Subhash Suri, Sue Whitesides

CURVATURE-CONSTRAINED SHORTEST PATHS IN A CONVEX POLYGON

PANKAJ K. AGARWAL[†], THERESE BIEDL[‡], SYLVAIN LAZARD[§], STEVE ROBBINS[¶], SUBHASH SURI[‖], AND SUE WHITESIDES[**]

Abstract. Let B be a point robot moving in the plane, whose path is constrained to have curvature at most 1, and let $\mathcal{P}$ be a convex polygon with n vertices. We study the collision-free, optimal path planning problem for B moving between two *configurations* inside $\mathcal{P}$ (a configuration specifies both a location and a direction of travel). We present an $O(n^2 \log n)$ time algorithm for determining whether a collision-free path exists for B between two given configurations. If such a path exists, the algorithm returns a shortest one. We provide a detailed classification of curvature-constrained shortest paths inside a convex polygon and prove several properties of them, which are interesting in their own right. For example, we prove that any such shortest path is comprised of at most eight segments, each of which is a circular arc of unit radius or a straight line segment. Some of the properties are quite general and shed some light on curvature-constrained shortest paths amid obstacles.

Key words. Nonholonomic motion planning, shortest paths, mobile robot, curvature constraint, bounded radius of curvature, Dubins path, computational geometry.

AMS subject classifications. 68U05

1. Introduction.

The *path-planning* problem, a central problem in robotics, involves planning a collision-free path for a robot moving amid obstacles, and has been widely studied (see, e.g., the book by Latombe [21] and the survey papers by Schwartz and Sharir [31] and Halperin, Kavraki and Latombe [17]). In the simplest form, given a moving point robot B, a set of obstacles, and a pair of configurations I and F specifying locations for B, we wish to find a continuous, collision-free path for B from I to F. This formulation, however, does not take into account the so-called *nonholonomic constraints* (for instance, bounds on acceleration or curvature), imposed on a robot by its physical limitations (see [21] for a more detailed discussion). Although there has been considerable recent work reported in the robotics and non-linear-control literatures on nonholonomic motion-planning problems (see [3, 5, 19, 20, 22, 24, 32, 42, 43] and references therein), relatively little theoretical work has been done in this important area from an algorithmic perspective.

[†]Center for Geometric Computing, Computer Science Department, Duke University, Box 90129, Durham, NC 27708–0129, USA; pankaj@cs.duke.edu; http://www.cs.duke.edu/∼pankaj/. Supported in part by National Science Foundation research grant EIA–9870734, EIA–997287, and CCR–9732787, by Army Research Office MURI grant DAAH04–96–1–0013, by a Sloan fellowship, and by a grant from the U.S.-Israeli Binational Science Foundation.

[‡]Department of Computer Science, University of Waterloo, Waterloo, ON, N2L 3G1, Canada; biedl@uwaterloo.ca. This research began during a postdoctoral tenure at McGill University.

[§]INRIA Lorraine - LORIA, 615 rue du jardin botanique, B.P. 101, 54602 Villers-les-Nancy Cedex, France. lazard@loria.fr. http://www.loria.fr/∼lazard/. This research began during a postdoctoral tenure at McGill University.

[¶]School of Computer Science, McGill University; stever@cs.mcgill.ca. Supported by an FCAR scholarship.

[‖]Department of Computer Science, University of California, Santa Barbara, CA 93106, USA; suri@cs.ucsb.edu; http://www.cs.ucsb.edu/∼suri/. Research partially supported by NSF Grant CCR-9501494.

[**]School of Computer Science, McGill University; sue@cs.mcgill.ca. Supported by NSERC and FCAR research grants.

In this paper, we study the path-planning problem for a point robot whose configurations are specified by giving both a location and a direction of travel. This means that any solution to the path-planning problem for given initial and final configurations I and F must respect the directions of travel specified by I and F as well as the locations they specify. Furthermore, we require the path of the robot to have curvature at most 1. This curvature constraint arises naturally when the point robot models a real-world robot with a minimum turning radius; see for example [21]. Recently, Reif and Wang [29] confirmed that the problem of deciding whether there exists a collision-free curvature-constrained path for B between two given configurations amid polygonal obstacles is NP-hard. As a first step toward understanding which environments admit efficient solution to this NP-hard problem, we study curvature-constrained path planning inside convex polygons.

We establish several new properties of shortest paths inside convex polygons and use them to characterize shortest paths in such environments. Using these properties and some geometric data structures [10], we present an efficient algorithm that, given initial and desired final configurations I and F in a convex polygon $\mathcal{P}$, determines whether there exists a curvature-constrained path from I to F that lies inside $\mathcal{P}$, and if so, computes a shortest one.

1.1. Previous results. Dubins [15] was perhaps the first to study curvature-constrained shortest paths. He proved that, in the absence of obstacles, a curvature-constrained shortest path from any start configuration to any final configuration consists of at most three segments, each of which is either a straight line segment or an arc of a circle of unit radius, assuming that the curvature of the path is upper bounded by 1. Using ideas from control theory, Boissonnat, Cérézo, and Leblond [5], and at the same time Sussmann and Tang [38], gave an alternative proof. Further characterizations of shortest paths, in an obstacle-free environment, have been made by Bui, Boissonnat, Laumond and Souères in [4, 8].

In the presence of polygonal obstacles, Jacobs and Canny [18] proved that if there exists a curvature-constrained path between two configurations, there exists in the closed free space a curvature-constrained shortest path between these two configurations. Fortune and Wilfong [16] gave a $2^{poly(n,m)}$ time algorithm that decides whether a path exists, without finding one, where n is the total number of vertices in the polygons defining the obstacles and m is the number of bits of precision with which all points are specified. Jacobs and Canny [18], Wang and Agarwal [41], and Sellen [33, 34] gave approximation algorithms for computing an *ε-robust* path. (Informally, a path is ε-robust if it may be deformed by ε in a certain way while remaining feasible.) For the restricted case of pairwise disjoint *moderate obstacles*, i.e., convex obstacles whose boundaries have curvature upper bounded by 1, Agarwal, Raghavan, and Tamaki [1] gave efficient approximation algorithms. Boissonnat and Lazard [6] gave an $O(n^2 \log n)$ time algorithm for computing an exact shortest path for the case when the edges of the pairwise disjoint moderate obstacles consist of n segments that are circular arcs of unit radius or line segments. Their algorithm can be used to compute an optimal curvature-constrained path inside a convex polygon in time $O(n^7)$. Recently, Ahn, Cheong, Matoušek, and Vigneron [2] characterized the region of all points that can be reached from a given configuration in a convex polygon. Wilfong [42, 43] studied a restricted problem in which the robot must stay on one of m line segments (thought of as "lanes"), except to turn between lanes. For a scene with n obstacle vertices, his algorithm preprocesses the scene in time $O(m^2(n^2 + \log m))$, following which queries are answered in time $O(m^2)$.

There has also been work on characterizing shortest curvature-constrained paths when B is allowed to make reversals, that is, to back up. Reeds and Shepp [28] were the first to compute the characterization of such shortest paths in an obstacle-free environment. Boissonnat *et al.* [5] and Sussmann *et al.* [38] presented an alternative proof using ideas from control theory. Souères and Laumond refined the characterization of shortest paths [36]. In the presence of obstacles, Desaulniers proved that a *shortest* path does not necessarily exist even when paths exist [13]. Finally, much work has been done on computing curvature-constrained paths, in the presence of obstacles, when reversals are allowed; see e.g. [3, 19, 20, 21, 22, 23, 24, 25, 32, 39, 40].

Other more general dynamic constraints have been considered in [3, 11, 12, 14, 26]. In particular, Sussmann [37] extended the characterization of curvature-constrained shortest paths to the 3-dimensional case.

1.2. Our model and results. Let B be a point robot and $\mathcal{P}$ a closed convex polygon with n vertices. For simplicity we assume that the edges of $\mathcal{P}$ are in general position: no two edges are parallel and no unit-radius circle (in $\mathbb{R}^2$) is tangent to three edges of $\mathcal{P}$. We believe that our techniques can be extended to carry through without this general position assumption, although the technical details would be daunting. A *configuration* X for B is a pair $(\textsc{loc}(X), \psi(X))$, where $\textsc{loc}(X)$ is a point in the plane representing the location of the robot and $\psi(X)$ is an angle between 0 and 2π representing its orientation. When the meaning is clear, we generally write X instead of $\textsc{loc}(X)$.

The image of a differentiable function $\Pi : [0, l] \to \mathbb{R}^2$ is called a *path*. We denote by Π both the function and the path it defines. We regard a path Π as oriented from $\Pi(0)$ to $\Pi(l)$. We assume a path Π is parameterized by its arc length, and we let $\|\Pi\|$ denote its length. We say that Π is a path from a configuration X to another configuration Y if $\Pi(0) = \textsc{loc}(X)$, $\Pi(l) = \textsc{loc}(Y)$, and the oriented angles (with respect to the positive x-axis) of $\Pi'(0)$ and $\Pi'(l)$ are $\psi(X)$ and $\psi(Y)$, respectively. A path is called *moderate* if its average curvature is at most 1 in every positive-length interval.[1] This implies that the curvature is defined almost everywhere and that it is at most 1 wherever it is defined. Indeed, the derivative of a moderate path Π satisfies a Lipschitz condition, and therefore the derivative is differentiable almost everywhere [30, Theorem 8.19].

Any curve that lies entirely within the closed polygon $\mathcal{P}$ is called *free*. A path is *feasible* if it is moderate and free. A path Π from a configuration X to another configuration Y is *optimal* if it is feasible and its length is minimum among all feasible paths from X to Y (if a feasible path from X to Y exists, then an optimal such path also exists [18]).

Throughout the paper, we say that we *compute* an object as a short way of saying that we compute such an object if one exists, and return a flag of non-existence otherwise.

Main results. Let $\mathcal{P}$ be an n-vertex convex polygon in the plane, and let I and F be two configurations inside $\mathcal{P}$.

(i) We give a classification of optimal paths from I to F.
(ii) We prove that an optimal path from I to F consists of at most eight maximal segments, each of which is either a line segment or a circular arc of unit radius.
(iii) We give an $O(n^2 \log n)$ time algorithm to determine whether a feasible path from I to F exists. If such a path exists, then the algorithm returns an

[1] The *average curvature* of a path Π in the interval $[s_1, s_2]$ is defined by $\|\Pi'(s_1) - \Pi'(s_2)\|/|s_1 - s_2|$.

optimal path from I to F. If there are only k edges of $\mathcal{P}$ at distance at most 6 from both I and F, then the running time of our algorithm becomes $O((n+k^2)\log n)$.

Note that Result (ii) above is actually quite surprising. Indeed, it means that the complexity of optimal paths inside a convex polygon is constant and does not depend on the number of edges of the polygon.

Note also that our algorithm for computing optimal paths is significantly faster than the algorithm implicit in the work of Boissonnat and Lazard [6] on computing an optimal path amid overlapping moderate obstacles, whose running time would be $O(n^7)$.

Some of the properties of moderate paths we prove are interesting in their own right. For example, one of these properties identifies a type of narrow region, called a "pocket", from which a moderate path cannot escape once it enters from outside. The conclusion will highlight this and another of these properties, which require technical definitions from later sections to describe in detail.

Our paper is organized as follows. In Section 2, we present basic definitions, notation, and useful known results. In Section 3, we give a classification of the optimal paths. In Sections 4 and 5, we describe our algorithm. We conclude in Section 6 with a discussion and some open problems.

2. Geometric Preliminaries. Given a configuration X, let L_X denote the oriented line passing through $\textsc{loc}(X)$ with orientation $\psi(X)$. A configuration X *belongs* to an oriented path (or curve) Π if $\textsc{loc}(X) \in \Pi$ and L_X is the oriented tangent line to Π at $\textsc{loc}(X)$. Note that a configuration X belongs to two oriented unit-radius circles. We will use $\mathcal{C}_X^+$ and $\mathcal{C}_X^-$ to denote the two circles of unit radius, oriented counterclockwise and clockwise, respectively, to which the configuration X belongs.

If X and Y are two points on a simple closed curve γ, then $\gamma^+[X,Y]$ (respectively $\gamma^-[X,Y]$) denotes the portion of γ from X to Y in the counterclockwise (respectively clockwise) direction, including X and Y; we will use $\gamma^+(X,Y), \gamma^-(X,Y)$ to denote portions excluding X, Y. Similarly, for a path Π and two configurations $X, Y \in \Pi$, we will use $\Pi[X,Y]$ to denote the portion of Π from X to Y.

Segments and Dubins paths. Let Π be a feasible path. We call a nonempty subpath of Π a *C-segment* (respectively *S-segment*) if it is a circular arc of unit radius (respectively line segment) and maximal, i.e., it is not strictly contained in a longer circular arc (respectively line segment) of the path. A *segment* is either a C-segment or an S-segment. A C-segment on a path Π is called a C^+-segment (respectively C^--segment) if Π induces a counterclockwise (respectively clockwise) orientation on it. Suppose Π consists of a C-segment, an S-segment, and a C-segment; then we will say that Π is of type CSC, or C_1SC_2, or $C_1S_2C_3$ if for ease of reference we want to number segments in order of their appearance in the sequence; superscripts $+$ and $-$ will be used to specify the orientations of C-segments of Π. Abusing the notation slightly, we often use the same symbol to denote both the type of a segment and the segment itself. Thus we may denote the first C-segment occurring in some path of type C_1SC_2 by C_1. The above notation can be generalized to an arbitrarily long sequence. Recall that throughout the paper, C-segments and S-segments have positive length. Dubins [15] proved the following result.

LEMMA 2.1 (Dubins [15]). *In an obstacle-free environment, an optimal path between any two configurations is of type CCC or CSC, or a substring thereof.*

We will refer to paths of type CCC or CSC or substrings thereof as *Dubins paths*. In the presence of obstacles, Jacobs and Canny [18] observed that any subpath of an

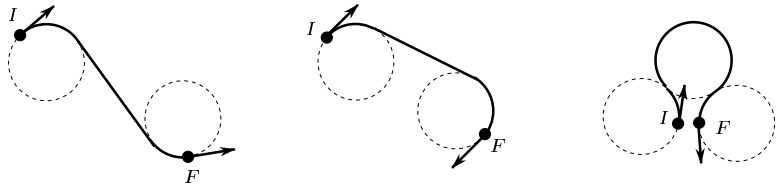


FIG. 2.1. *Different types of Dubins paths.*

optimal path that does not touch any obstacle except at the endpoints is a Dubins path. In particular, they proved the following.

LEMMA 2.2 (Jacobs and Canny [18]). *Let Ω be a closed polygonal environment, I an initial configuration, and F a final configuration. Then an optimal path from I to F in Ω consists of a sequence $\Pi_1 \cdots \Pi_k$ of feasible paths, where each Π_i is a Dubins path from a configuration X_{i-1} to a configuration X_i, such that $X_0 = I$, $X_k = F$, and, for $0 < i < k$,* $\text{LOC}(X_i) \in \partial\Omega$.

The above lemma implies that an optimal path in a closed polygonal environment consists of C- and S-segments. In the following, we will consider only those paths that are formed by C- and S-segments. We will refer to circles and circular arcs of unit radius simply as circles and circular arcs. Notationally we distinguish between a C-segment and its supporting circle, that is, the circle on which the circular arc lies, by using calligraphy font $\mathcal{C}$ for the latter. While the symbols C and $\mathcal{C}$ are similar in appearance, it will be clear from the context whether we are referring to a supporting circle (denoted $\mathcal{C}$) or to a circular arc (denoted C).

Terminal and nonterminal segments. A segment of a feasible path Π is called *terminal* if it is the first or the last segment of Π; otherwise it is called *nonterminal*. We apply the adjectives terminal and nonterminal to subpaths as well. If the first or last segment in Π is a C-segment, we will refer to it as a C_I-segment or a C_F-segment, respectively. Circles $\mathcal{C}_I^+$, $\mathcal{C}_I^-$, $\mathcal{C}_F^+$, and $\mathcal{C}_F^-$ are called *terminal circles* (see Figure 2.3).

The following lemma states some basic known properties of optimal paths; see [1, 15, 18].

LEMMA 2.3. *In an optimal path inside a convex polygon $\mathcal{P}$,*

(i) any nonterminal C-segment has length greater than π,

(ii) any nonterminal C-segment is tangent to $\partial\mathcal{P}$ or to a terminal circle in at least one point, and

(iii) no nonterminal subpath has type CCC.

Next we prove a property of a CS-subpath in an optimal path, which will be useful for our analysis.

LEMMA 2.4. *Let Π be an optimal path that contains a subpath of type CS where the C-segment is not terminal. Then the C-segment is tangent to $\partial\mathcal{P}$, and the length of C between its first point and its last point of tangency with $\partial\mathcal{P}$ is greater than π.*

Proof. Let Π be an optimal path containing a subpath Π' of type $C_1C_2S_3$ or $S_1C_2S_3$. See Figure 2.2. Let X be the common endpoint of the first and second segments of Π', let Y be the last point of tangency of C_2 with $\partial\mathcal{P}$ along Π, and let X' be the point antipodal to X on C_2 ($X' \in C_2$ because the length of C_2 is greater than π by Lemma 2.3(i)).

Dubins [15] proved that the perturbations shown in Figure 2.2 (transforming paths $C_1C_2S_3$ and $S_1C_2S_3$ into paths of type $CCCS$ and $SCCS$ by reducing the length of the first segment by any arbitrarily small value) shorten the paths in an obstacle-

free environment. These shortenings can be done unless an obstacle obstructs the shortcut, i.e., unless $\partial\mathcal{P}$ is tangent to C_2 after X' (along Π). Thus if $\|\Pi[X,Y]\| \leqslant \pi$, then Π can be shortened in $\mathcal{P}$, which contradicts the optimality of Π. □

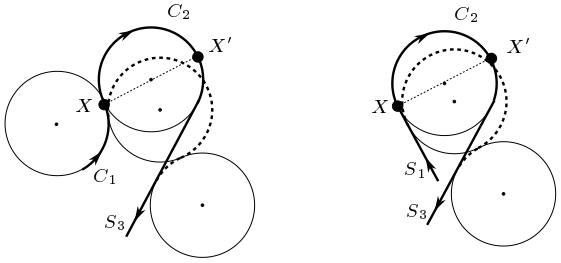


FIG. 2.2. *Length-reducing perturbations.*

Anchored segments. We define here the notion of anchored segments in a way that is similar to the notion introduced in [1]. A C-segment or circle is called *anchored* if it has at least two points of tangency with $\partial\mathcal{P}$ and the terminal circles. The terminal circles are not considered anchored. An anchored C-segment is denoted by $\bar{\bar{C}}$. By our general-position assumption on $\mathcal{P}$, there are a finite number of anchored circles. A C-segment with at least one point of tangency with $\partial\mathcal{P}$ is denoted by $\bar{C}$.

An anchored C-segment or circle is *$\mathcal{PP}$-anchored* if it is tangent to $\partial\mathcal{P}$ at two points, and *$\mathcal{PC}$-anchored* if it is tangent to $\partial\mathcal{P}$ at one point and tangent to a terminal circle at another point; see Figure 2.3.

A circular arc is called *long* if its length is greater than π; otherwise it is called *short*. A $\mathcal{PP}$-anchored C-segment is called *strongly $\mathcal{PP}$-anchored* if it contains the long arc defined by the points of tangency of its supporting circle with $\partial\mathcal{P}$ (see Figure 2.4b). Similarly, a $\mathcal{PC}$-anchored C-segment is called *strongly $\mathcal{PC}$-anchored* if it contains the long arc between a point of tangency of its supporting circle $\mathcal{C}$ with $\partial\mathcal{P}$ and a point of tangency of $\mathcal{C}$ with a terminal circle (see Figure 3.1a).

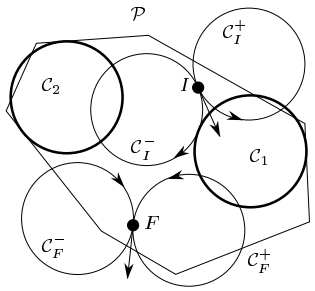


FIG. 2.3. *$\mathcal{PC}$-anchored ($\mathcal{C}_1$) and $\mathcal{PP}$-anchored ($\mathcal{C}_2$) circles.*

Closed moderate paths. We state here the following results about closed moderate paths.

PROPOSITION 2.5. *The region bounded by a simple moderate path Π whose initial and final locations coincide (the initial and final orientations may differ) contains at least one disk of unit radius. Moreover, if the initial and final orientations also coincide and if Π is not a circle of unit radius, then the region bounded by Π contains at least two distinct (but possibly overlapping) disks of unit radius.*

The second statement in the proposition above was proved by Pestov and Ionin [27] for the restricted class of closed C^2 continuous curves (i.e., twice differentiable curves with continuous second derivative) whose curvatures are bounded everywhere. At the time we submitted this paper, Ahn *et al.* [2] presented a simple generalization of Pestov and Ionin's result, which yields the first statement of Proposition 2.5; furthermore, it appears straightforward to generalize their proof to obtain the second statement. Rather than carrying out that generalization here, we include in Appendix A the proof we gave in the original version of this paper.

Pockets. We introduce here the notion of pockets. Let $\mathcal{C}$ be a circle intersecting $\partial\mathcal{P}$ at two or more points, and let X, Y be two consecutive[2] intersection points of $\partial\mathcal{P}$ with $\mathcal{C}$ so that the short arc of $\mathcal{C}$ joining X and Y lies inside $\mathcal{P}$. If $\mathcal{C}^+[X,Y]$ is the short arc and the turning angle[3] of $\partial\mathcal{P}^+(X,Y)$ is less than π, then the closed region bounded by $\partial\mathcal{P}^+[X,Y]$ and $\mathcal{C}^+[X,Y]$ is called a *pocket* (see Figure 2.4) and is denoted by $\Lambda_{\mathcal{C}}[X,Y]$. Similarly we define the pocket $\Lambda_{\mathcal{C}}[X,Y]$ for the case in which $\mathcal{C}^-[X,Y]$ is the shorter arc. We will mostly be interested in pockets for which $\mathcal{C}$ is tangent to $\partial\mathcal{P}$ at X.

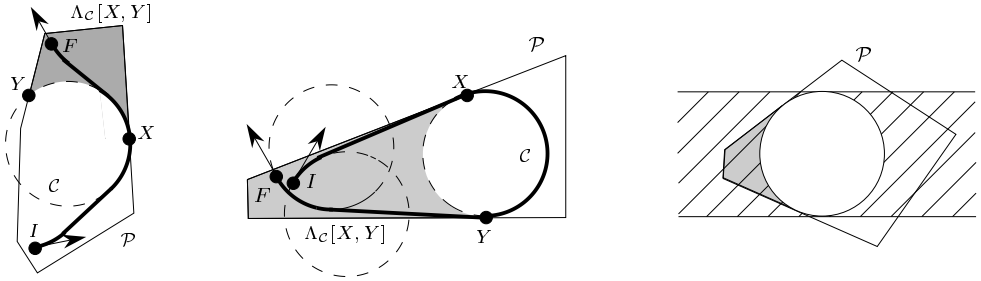


FIG. 2.4. *Pockets.*

We first prove the following simple property of pockets.

LEMMA 2.6. *A pocket cannot contain a unit circle.*

Proof. Any pocket is contained in an open strip of width 2 (see Figure 2.4c) which does not contain any circle of unit radius. Thus there is no room for a unit circle in a pocket. □

We prove the following lemma, which will be crucial for characterizing the optimal paths containing a strongly anchored C-segment.

LEMMA 2.7. *If a feasible path enters the interior of a pocket, then it cannot escape the pocket.*

Proof. For a contradiction, let Π be a feasible path that enters the interior of a pocket $\Lambda_{\mathcal{C}}$ at X and escapes it at Y. See Figure 2.5. Let $\mathcal{C}$ denote the circle defining the pocket $\Lambda_{\mathcal{C}}$, and let $\mathcal{D}$ be the closed disk whose boundary is $\mathcal{C}$. If $\mathcal{C}$ intersects $\mathcal{C}_X^+$ in exactly two points, let them be denoted by X and X^+. If $\mathcal{C} = \mathcal{C}_X^+$, let X^+ denote the point on $\mathcal{C}$ antipodal to X. If the intersection of the two circles consists of exactly one point, denote this point by $X = X^+$. We define X^-, Y^+ and Y^- similarly. Let O_X^+, O_X^-, and O be the centers of the circles $\mathcal{C}_X^+, \mathcal{C}_X^-$, and $\mathcal{C}$, respectively.

[2]Since $\mathcal{C}$ and $\partial\mathcal{P}$ are both convex, two consecutive intersection points of $\partial\mathcal{P}$ and $\mathcal{C}$ are consecutive on both $\partial\mathcal{P}$ and $\mathcal{C}$.

[3]The *turning angle* of a convex polygonal chain is $\sum_i(\pi - \theta_i)$, where θ_i is the interior angle at vertex i.

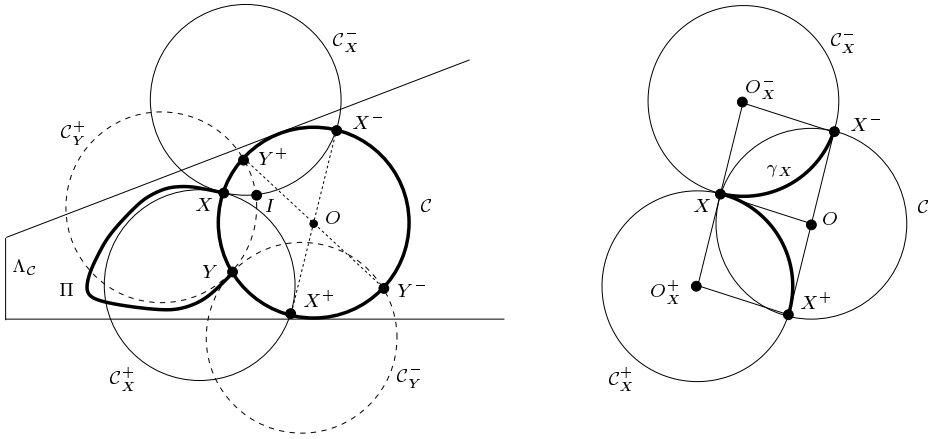


FIG. 2.5. *Illustration of the proof of Lemma 2.7.*

We first prove that the segments X^+X^- and Y^+Y^- are diameters of $\mathcal{C}$. See Figure 2.5. The quadrilaterals $O_X^+XOX^+$ and $O_X^-XOX^-$ are parallelograms. (These parallelograms flatten to line segments when $\mathcal{C}$ is equal to $\mathcal{C}_X^+$ or $\mathcal{C}_X^-$, but then X^+X^- is a diameter of $\mathcal{C}$ by definition.) Since the two parallelograms share the edge XO, and the edges O_X^+X and XO_X^- are collinear, the two parallelograms are congruent. Therefore, O is the middle point of the segment X^+X^-. Similarly, Y^+Y^- is also a diameter of $\mathcal{C}$.

Let γ_X be the union of the arcs of $\mathcal{C}_X^+$ and $\mathcal{C}_X^-$ in $\mathcal{D}$, i.e., $\gamma_X = (\mathcal{C}_X^+ \cup \mathcal{C}_X^-) \cap \mathcal{D}$, and let γ_Y be defined similarly. The set γ_X is either equal to $\mathcal{C}$ or consists of two unit radius circular arcs XX^- and XX^+; similarly for γ_Y. The set γ_X belongs to $\mathcal{D}$ and segments X^+X^- and Y^+Y^- are diameters of $\mathcal{C}$; thus γ_X separates (not necessarily strictly) Y^+ and Y^- in $\mathcal{D}$. Since γ_Y also belongs to $\mathcal{D}$, γ_X and γ_Y intersect, by the Jordan curve theorem.

First, note that the region $\Lambda_\mathcal{C} \cup \mathcal{D}$ cannot contain any unit radius disk except $\mathcal{D}$; this is because O is the only point in $\Lambda_\mathcal{C} \cup D$ that is distance at least 1 from the boundary of $\Lambda_\mathcal{C} \cup D$.

Second, note that Π must be simple. Otherwise, Π would contain a simple, moderate subpath Π', lying in $\Lambda_\mathcal{C}$, with equal initial and final locations. This contradicts Proposition 2.5, since by Lemma 2.6, a pocket cannot contain a unit disk.

Third, locations X and Y cannot be equal. Otherwise, Π would be simple, moderate and lying in $\Lambda_\mathcal{C}$, with equal initial and final locations. As before, this contradicts Proposition 2.5.

Now we describe how to obtain a path Π' that leads to a contradiction. Consider first the case in which γ_X and γ_Y each consist of two unit radius circular arcs XX^+, XX^- and YY^+, YY^-, respectively. Let I be the intersection point between γ_X and γ_Y that is the closest to X on γ_X (see Figure 2.5). This ensures that the circular arcs $IX \subset \gamma_X$ and $YI \subset \gamma_Y$ intersect only at I. Let Π' be the concatenation of the circular arc $IX \subset \gamma_X$, the path Π, and the circular arc $YI \subset \gamma_Y$. Path Π' is simple because Π is simple and the arcs IX and IY only intersect at I. Thus Π' is simple, moderate and contained in the region $\Lambda_\mathcal{C} \cup \mathcal{D}$, which does not contain any unit radius disk except $\mathcal{D}$. Thus, by Proposition 2.5, Π' encloses $\mathcal{D}$. It follows that the circular arcs IX and IY defining Π' cannot intersect the interior of $\mathcal{D}$. Thus the arcs IX and IY are reduced to a point $I = X = Y$, a contradiction.

Consider now the case in which exactly one of γ_X, γ_Y is equal to $\mathcal{C}$. Assume without loss of generality that $\gamma_X = \mathcal{C}$ and γ_Y consists of two unit radius circular arcs YY^+ and YY^-. See, for example, Figure 2.6a, b. Since $\gamma_X = \mathcal{C}$, circle $\mathcal{C}$ is equal to $\mathcal{C}_X^+$ or $\mathcal{C}_X^-$; assume without loss of generality that $\mathcal{C} = \mathcal{C}_X^+$. Let I be the point in $\gamma_X \cap \gamma_Y = \{Y^-, Y, Y^+\}$ such that $\mathcal{C}^+[I, X]$ has minimum length. It follows that the arc $\mathcal{C}^+[I, X]$ intersects the arc $YI \subset \gamma_Y$ only at I. As before, we define Π' as the concatenation of $\mathcal{C}^+[I, X]$, Π and YI, and again, Π' is simple, moderate and must enclose $\mathcal{D}$. It follows that the circular arc YI defining Π' cannot intersect the interior of $\mathcal{D}$, and thus $I = Y$. Thus path Π' is the concatenation of $\mathcal{C}^+[I, X]$ and Π. Path Π lies in $\Lambda_{\mathcal{C}}$, and $\Lambda_{\mathcal{C}} \cap \mathcal{D}$ is an arc of $\mathcal{C}$ of length smaller than π. Thus, since Π' must enclose $\mathcal{D}$, the arc $\mathcal{C}^+[I, X]$ is the long arc of $\mathcal{C}$ joining I and X. Since Y^+Y^- is a diameter of $\mathcal{C}$, the arc $\mathcal{C}^+[I, X]$ strictly contains Y^+ or Y^-, contradicting the definition of I.

The remaining case is when both γ_X and γ_Y are equal to $\mathcal{C}$. Let Π' be the concatenation of Π and the arc XY of $\mathcal{C}$ joining X to Y such that, if possible, Π' is moderate everywhere (see, for example, Figure 2.6c), or if not possible, XY is the short arc (see Figure 2.6d). In the first case, Π' encloses at least two unit radius disks by Proposition 2.5, contradicting the fact that $\Lambda_{\mathcal{C}} \cup \mathcal{D}$ contains Π' and only one unit radius disk $\mathcal{D}$. In the second case, Π' encloses one unit radius disk by Proposition 2.5, and lies in $\Lambda_{\mathcal{C}}$ because the short arc of $\mathcal{C}$ joining X and Y lies in $\Lambda_{\mathcal{C}}$. Thus $\Lambda_{\mathcal{C}}$ contains a unit disk, contradicting Lemma 2.6. □

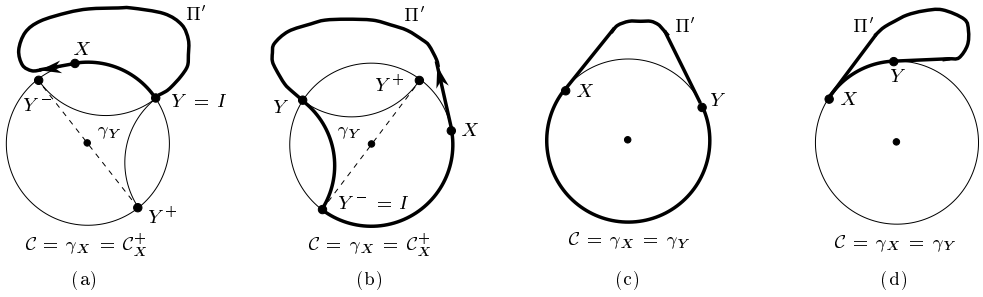


FIG. 2.6. *Illustration of the proof of Lemma 2.7.*

3. Classification of Optimal Paths. The goal of this section is to prove the first of our main results, namely a detailed characterization of optimal paths in convex polygons. We show that any optimal path is of type $C_I C S C C S C C_F$ or a subsequence of this form. However, not every subsequence of the above sequence can form an optimal path. The following theorem gives a more refined description of optimal path types. Recall that a segment has non-zero length by definition. In the following, we use $\cdot$ to denote a subpath of zero length.

THEOREM 3.1.

An optimal path Π inside $\mathcal{P}$ either is a Dubins path or is one of the types listed below. Except in case (B.i), all the C-segments labeled $\bar{\bar{C}}$ are strongly anchored.

(A) If Π has no nonterminal CC subpath, then Π is one of the following types:

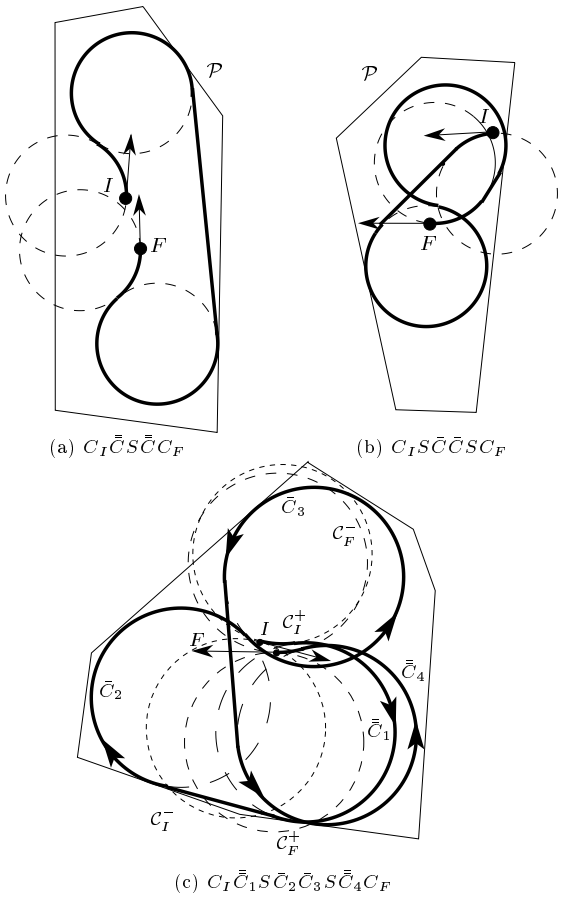
(A.i) $\Pi_I S \bar{\bar{C}} S \Pi_F$, where $\Pi_I \in \{C_I, \cdot\}$ and $\Pi_F \in \{C_F, \cdot\}$ (see Figure 2.4b), and

(A.ii) $\Pi_I S \Pi_F$, where $\Pi_I \in \{C_I \bar{\bar{C}}, C_I, \cdot\}$ and $\Pi_F \in \{\bar{\bar{C}} C_F, C_F, \cdot\}$ (see Figure 3.1a).

(B) If Π has a nonterminal CC subpath, then Π is one of the following types:

(B.i) $C_I C \bar{\bar{C}} C_F$ or $C_I \bar{\bar{C}} C C_F$,

(B.ii) $\Pi_I S\bar{\bar{C}}CC_F$ *or* $C_I C\bar{\bar{C}}S\Pi_F$, *where* $\Pi_I \in \{C_I, \cdot\}$ *and* $\Pi_F \in \{C_F, \cdot\}$, *and*
(B.iii) $\Pi_I \bar{C}\bar{C}\Pi_F$, *where* $\Pi_I \in \{C_I\bar{\bar{C}}S, C_I S, C_I, S\}$ *and* $\Pi_F \in \{S\bar{\bar{C}}C_F, SC_F, C_F, S\}$ *(see Figures 3.1b, c).*



(a) $C_I\bar{\bar{C}}S\bar{\bar{C}}C_F$ (b) $C_I S\bar{C}\bar{C}SC_F$

(c) $C_I\bar{\bar{C}}_1 S\bar{C}_2\bar{C}_3 S\bar{\bar{C}}_4 C_F$

FIG. 3.1. *Examples of shortest paths.*

REMARK 3.2. We informally checked by a case analysis on figures that for the polygon and configurations I and F shown in Figure 3.1c, no path of type (A) or (B) is feasible except paths of type $C_I CS\bar{C}\bar{C}SCC_F$. This leads us to believe that the type $C_I CS\bar{C}\bar{C}SCC_F$, having eight segments, does occur as an optimal path type.

The proof of Theorem 3.1 is based on the following lemmas.

LEMMA 3.3 (Agarwal, Raghavan and Tamaki [1]). *An optimal path has at most one nonterminal CC subpath. Moreover, any nonterminal C-segment that precedes (respectively follows) a* C_1C_2 *subpath is oriented the same way as* C_1 *(respectively* C_2*).*

Next we state a lemma whose proof we postpone until Section 3.1.

LEMMA 3.4. *(i) If an optimal path has a subpath of type SCS, then the C-segment in that subpath is strongly* $\mathcal{PP}$*-anchored.*
(ii) If an optimal path has a subpath of type $C_1C_2C_3S$ *(or* $SC_3C_2C_1$*) so that the C-segment* C_2 *does not touch* $\partial\mathcal{P}$*, then* C_3 *is strongly* $\mathcal{PP}$*-anchored (see Figure 3.5).*

We next characterize some optimal paths that contain a strongly anchored C-segment.

LEMMA 3.5. *(i) If an optimal path* Π *contains a strongly* $\mathcal{PP}$*-anchored* C*-segment* $\bar{\bar{C}}$*, then* $\Pi = \Pi_I \bar{\bar{C}} \Pi_F$*, where* Π_I *and* Π_F *lie in some pockets defined by* $\bar{\bar{\mathcal{C}}}$*.*
(ii) If an optimal path Π *contains a strongly* $\mathcal{PC}$*-anchored* C*-segment* $\bar{\bar{C}}$ *whose supporting circle is not free, then* Π *is of type* $C_I \bar{\bar{C}} \Pi_F$ *or* $\Pi_I \bar{\bar{C}} C_F$*, where* Π_I *and* Π_F *lie in some pockets defined by* $\bar{\bar{\mathcal{C}}}$*.*

Proof. In case (i), the segment $\bar{\bar{C}}$ is strongly $\mathcal{PP}$-anchored, so it is tangent to $\partial\mathcal{P}$ at two points. Let X be the first point of tangency encountered along $\bar{\bar{C}}$, and let Y be the second point of tangency. See Figure 3.2a. By the general-position assumption, $\bar{\bar{C}}$ does not touch $\partial\mathcal{P}$ at any point other than X and Y. Without loss of generality, let $\bar{\bar{C}}$ be oriented counterclockwise. Let $\bar{\bar{\mathcal{C}}}$ be its supporting circle and define X' as the first point of intersection between $\bar{\bar{\mathcal{C}}}$ and $\partial\mathcal{P}$ encountered moving clockwise along $\bar{\bar{\mathcal{C}}}$ from X. Notice that X' belongs to $\bar{\bar{\mathcal{C}}}^-(X, Y]$ since the intersection between $\bar{\bar{\mathcal{C}}}$ and $\partial\mathcal{P}$ contains Y.

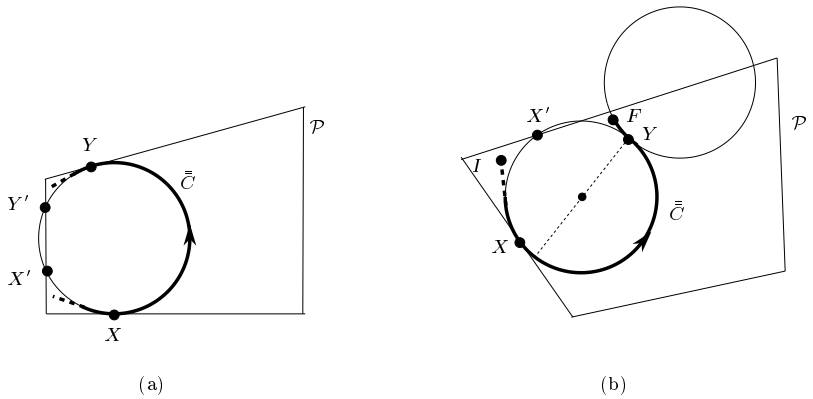


FIG. 3.2. *Illustration of the proof of Lemma 3.5. In (a), an optimal path containing a strongly* $\mathcal{PP}$*-anchored* C*-segment must start and end in a pocket.*

Segment $\bar{\bar{C}}$ is strongly anchored, so $\|\Pi[X, Y]\| = \|\bar{\bar{\mathcal{C}}}^+[X, Y]\| > \pi$. Point X' is on the short arc $\bar{\bar{\mathcal{C}}}^+[Y, X]$, so $\|\bar{\bar{\mathcal{C}}}^+[X', X]\| < \pi$. The turning angle of $\partial\mathcal{P}^+[X, Y]$ is equal to $\psi(Y) - \psi(X) \pmod{2\pi}$ which, in turn, is equal to $\|\Pi[X, Y]\| > \pi$. The total turning angle around $\partial\mathcal{P}$ is 2π, so the turning angle of $\partial\mathcal{P}^+[X', X] < \pi$. Finally, $\bar{\bar{\mathcal{C}}}$ is tangent to $\partial\mathcal{P}$ at X so $\bar{\bar{\mathcal{C}}}^+[X', X]$ lies inside the polygon. Thus arc $\bar{\bar{\mathcal{C}}}^+[X', X]$ forms the pocket $\Lambda_{\bar{\bar{\mathcal{C}}}}[X', X]$.

The pocket $\Lambda_{\bar{\bar{\mathcal{C}}}}[X', X]$ contains $\Pi[I, X]$, and thus Π_I, by Lemma 2.7. Indeed, otherwise, the path Π_I contains $\bar{\bar{\mathcal{C}}}^+[X', X]$ and $\partial\mathcal{P}$ is tangent to $\bar{\bar{\mathcal{C}}}$ at X'. This contradicts either the optimality of Π (if $X' = Y$), or the fact that X is the first point of tangency between $\bar{\bar{\mathcal{C}}}$ and $\partial\mathcal{P}$, encountered along Π (if $X' \neq Y$).

Similarly, Π_F is contained in a pocket $\Lambda_{\bar{\bar{\mathcal{C}}}}[Y, Y']$, where Y' is the first point of intersection between $\bar{\bar{\mathcal{C}}}$ and $\partial\mathcal{P}$ encountered moving counterclockwise along $\bar{\bar{\mathcal{C}}}$ from Y.

For the case (ii), we assume that $\Pi = \Pi_I \bar{\bar{C}} C_F$; the case in which $\Pi = C_I \bar{\bar{C}} \Pi_F$ is symmetrical. Let X be the first point along $\bar{\bar{C}}$ that is tangent to $\partial\mathcal{P}$, and let Y be the last point of $\bar{\bar{C}}$. See Figure 3.2b. Then the proof is exactly the same as in (i), once noted that X' always exists on $\bar{\bar{\mathcal{C}}}^-(X, Y]$ because $\bar{\bar{\mathcal{C}}}$ is not free by assumption. □

LEMMA 3.6. *Let* Π *be an optimal path that contains a subpath of type* $C_1 S \bar{\bar{C}}$ *or* $\bar{\bar{C}} S C_1$*, where* $\bar{\bar{C}}$ *is either a strongly* $\mathcal{PP}$*-anchored* C*-segment, or a strongly* $\mathcal{PC}$*-anchored* C*-segment whose supporting circle is not free. Then* C_1 *is terminal.*

Proof. We consider the case in which Π contains a subpath of type $C_1S\bar{\bar{C}}$; the case of a subpath of type $\bar{\bar{C}}SC_1$ is symmetrical. Suppose that C_1 is not terminal. If $\bar{\bar{C}}$ is strongly $\mathcal{PP}$-anchored or is a strongly $\mathcal{PC}$-anchored C-segment whose supporting circle is not free, then $\bar{\bar{\mathcal{C}}}$ defines one or more pockets. By Lemma 3.5, C_1 belongs to one of the pockets defined by $\bar{\bar{\mathcal{C}}}$, say $\Lambda_{\bar{\bar{\mathcal{C}}}}$.

We claim that the circle $\mathcal{C}_1$ supporting the C-segment C_1 is not free. Indeed, otherwise, there would exist a feasible path that enters $\Lambda_{\bar{\bar{\mathcal{C}}}}$ on $\mathcal{C}_1$ (since no circle of unit radius is entirely contained in a pocket by Lemma 2.6) and escapes the pocket on $\bar{\bar{C}}$, contradicting Lemma 2.7.

Since C_1 is not terminal, its length is greater than π, by Lemma 2.3(i). The length of $\bar{\bar{C}}$ is also greater than π because anchored circles are not terminal by definition.

Suppose first that the C-segments C_1 and $\bar{\bar{C}}$ have the same orientation along Π. Since the lengths of C_1 and $\bar{\bar{C}}$ are greater than π, the convex hull of C_1 and $\bar{\bar{C}}$ contains $\mathcal{C}_1$. By convexity of $\mathcal{P}$, the convex hull of C_1 and $\bar{\bar{C}}$, and thus $\mathcal{C}_1$, lies inside $\mathcal{P}$, which contradicts the above claim that $\mathcal{C}_1$ is not free.

Suppose now that the C-segments C_1 and $\bar{\bar{C}}$ have opposite orientation along Π (see Figure 3.3). Let A and B be the first and last point of the S-segment, respectively, in the subpath $C_1S\bar{\bar{C}}$. Let U be the last point of tangency along Π between C_1 and $\partial\mathcal{P}$, and let U' be the point antipodal to U on $\mathcal{C}_1$. By Lemma 2.4, U' belongs to C_1. Note that the arc on C_1 joining U' and A is longer than π. Let ℓ be the line tangent to $\mathcal{C}_1$ at U; ℓ contains an edge of $\mathcal{P}$.

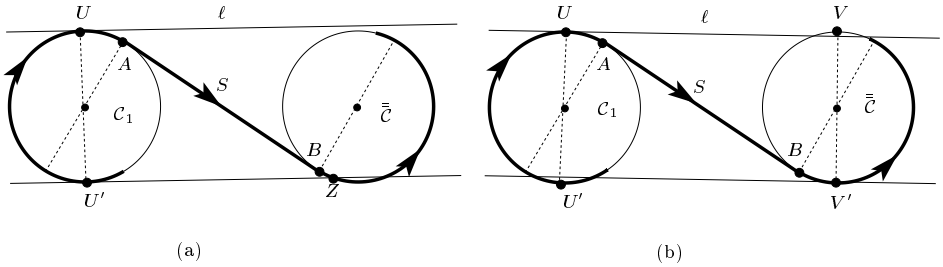


FIG. 3.3. *Illustration of the proof of Lemma 3.6.*

If ℓ does not intersect $\bar{\bar{\mathcal{C}}}$, then the line tangent to C_1 at U' (which is parallel to ℓ) intersects the subpath $S\bar{\bar{C}}$ at a point Z (see Figure 3.3a). It follows that the convex hull of $\Pi[U', Z]$ contains $\mathcal{C}_1$. Thus $\mathcal{C}_1$ is free, again contradicting the claim that $\mathcal{C}_1$ is not free. On the other hand, if ℓ intersects $\bar{\bar{\mathcal{C}}}$, then let $V, V' \in \bar{\bar{\mathcal{C}}}$ such that $\overrightarrow{VV'} = \overrightarrow{UU'}$ (see Figure 3.3b). Points B and V lie on opposite sides of ℓ, and $\mathcal{P}$ lies in one of the half-planes bounded by ℓ. Since $B \in \Pi \subset \mathcal{P}$, we conclude that $V \notin \mathcal{P}$. Recall that $\|\bar{\bar{C}}\| \geqslant \pi$; therefore $V' \in \bar{\bar{C}}$. The point on $\bar{\bar{C}}$ that is antipodal to the last point of $\bar{\bar{C}}$ lies on $\Pi[B, V']$. Thus, by Lemma 2.4 applied on the reversed path of Π, $\bar{\bar{C}}$ is tangent to $\partial\mathcal{P}$ at a point on $\Pi[B, V']$. Since the line tangent to $\bar{\bar{C}}$ at V' separates U' and A, any line tangent to $\Pi[B, V']$ separates U' and A. Thus $U' \notin \mathcal{P}$, which in turn implies that $U' \notin C_1$, thereby contradicting the claim that $U' \in C_1$. □

We now prove Theorem 3.1.

Proof of Theorem 3.1: The proof proceeds by considering how a nonterminal C-segment may appear in Π. If there is no nonterminal C-segment in Π, then Π is of type C_ISC_F or a substring thereof, i.e., Π is a Dubins path.

Assume now that there is a nonterminal C-segment in Π. Then such a segment belongs to a subpath of Π of type either SCS or CC. Suppose Π contains a subpath of type SCS. By Lemma 3.4, the C-segment in SCS must be strongly $\mathcal{PP}$-anchored. Thus, by Lemma 3.6, Π is of type $C_I S\bar{C}SC_F$, or substrings (containing $S\bar{C}S$) thereof. In other words, Π is of type (A.i).

If Π contains a nonterminal C-segment but not a subpath of type SCS, we know it must contain a subpath of type CC. There are two cases to consider, depending on whether the CC subpath is terminal.

Case 1: Π does not contain any nonterminal subpath of type CC. Thus one of the C-segments in any CC subpath must be a terminal segment. Either Π is of type C_IC_F or C_ICC_F, or any nonterminal C-segment of Π is also adjacent to an S-segment. Π must then be of type C_ICSCC_F, or any substring thereof containing S and a terminal CC. By Lemma 2.4, the nonterminal C-segments are strongly $\mathcal{PC}$-anchored. All these types of paths are thus either Dubins paths or of type (A.ii).

Case 2: Π contains a nonterminal subpath of type CC. By Lemma 3.3, it is the unique nonterminal CC subpath in Π. Thus Π has the form $\Pi_ICC\Pi_F$, where Π_I, Π_F do not contain a nonterminal subpath of type CC. Thus any nonterminal C-segment in Π_I must be followed by an S-segment; otherwise Π_ICC will contain a nonterminal CCC subpath, which contradicts the optimality of Π (Lemma 2.3(iii)). Furthermore, since we have no SCS subpath in Π, a nonterminal C segment must be preceded by a terminal C-segment. This means $\Pi_I = C_ICS$ or a subsequence of it. The subsequence cannot be empty, for otherwise the middle CC subpath of Π would in fact be terminal; nor can it be simply CC, as noted above. Thus $\Pi_I \in \{C_ICS, C_IS, C_I, S\}$. Similarly, $\Pi_F \in \{SCC_F, SC_F, C_F, S\}$.

If $\Pi_I = C_ICS$ or $\Pi_F = SCC_F$, then the nonterminal C-segment in Π_I or Π_F is strongly anchored, by Lemma 2.4.

If both Π_I and Π_F contain an S-segment, then the nonterminal CC subpath in Π is preceded and followed by an S-segment. Thus both C-segments of the nonterminal CC subpath in Π touch $\partial\mathcal{P}$. Indeed, otherwise Π contains a subpath of type SCC or CCS that does not touch $\partial\mathcal{P}$, which contradicts Lemma 2.2. Hence if both Π_I and Π_F contain an S, then Π is of type (B.iii).

If neither Π_I nor Π_F contains an S-segment, then the path is of type C_ICCC_F. By Lemma 2.2, one of the nonterminal C-segments must touch $\partial\mathcal{P}$. This C-segment is also tangent to a terminal circle and is therefore $\mathcal{PC}$-anchored. Thus the path is of type (B.i). Note that if both nonterminal C-segments touch $\partial\mathcal{P}$, then the path is of type $C_I\bar{C}\bar{C}C_F$, which can be considered as type (B.i) or (B.iii).

The last case to consider is when exactly one of Π_I or Π_F contains an S-segment. Say $\Pi_I = C_I$ and $\Pi_F \neq C_F$. The path has form $C_IC_1C_2\Pi_F$, where Π_F starts with an S-segment. We know that C_2 must touch $\partial\mathcal{P}$ by Lemma 2.3(ii). If C_1 also touches $\partial\mathcal{P}$, then the path Π is of type (B.iii). Otherwise, if C_1 does not touch $\partial\mathcal{P}$, then by Lemma 3.4(ii), C_2 must be strongly $\mathcal{PP}$-anchored. Lemma 3.6 then restricts the path Π to be of type (B.ii). Similarly, if $\Pi_I \neq C_I$ and $\Pi_F = C_F$, the path Π is of type (B.ii) or (B.iii). □

3.1. Proof of Lemma 3.4.

3.1.1. Proof of Lemma 3.4(i). We prove here Lemma 3.4(i) which states:

If an optimal path has a subpath of type SCS, then the C-segment in that subpath is strongly $\mathcal{PP}$-anchored.

Consider an optimal path $\Pi = S_1C_3S_5$ from configuration I to configuration F. We show that the path Π can be shortened unless the C-segment C_3 is strongly $\mathcal{PP}$-anchored. Assume without loss of generality that the C_3 is oriented counterclockwise, and refer to Figure 3.4.

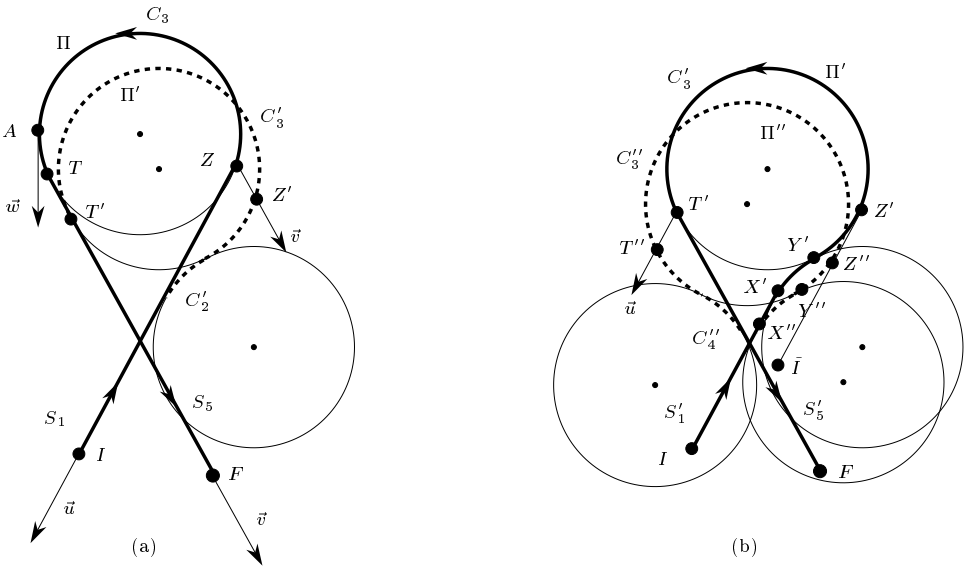


FIG. 3.4. *For the proof of Lemma 3.4(i).*

First perturbation. We first apply the perturbation shown in Figure 3.4a, which transforms the path $\Pi = S_1C_3S_5$ into a path $\Pi'(\varepsilon_1) = S_1'C_2'C_3'S_5'$ by reducing the length of S_5 by a small value ε_1 (i.e., $\|S_5'\| = \|S_5\| - \varepsilon_1$), such that the new segment C_2' is oriented clockwise and is smaller than π (segment C_3' is naturally oriented counterclockwise). For simplicity, in what follows, we denote $\Pi'(\varepsilon_1)$ by Π'.

CLAIM 3.7 (Dubins [15]). *In an obstacle-free environment, path Π' is shorter than Π, for any $\varepsilon_1 > 0$ small enough.*

This directly yields that, if the optimal path $S_1C_3S_5$ is tangent to $\partial\mathcal{P}$ at the junction between C_3 and S_5, then the perturbed path Π' has shortened in $\mathcal{P}$, unless C_3 is strongly $\mathcal{PP}$-anchored. By symmetry, we get the same result if the path Π is tangent to $\partial\mathcal{P}$ at the junction between S_1 and C_3. Therefore, we can assume in the following that neither S_1 nor S_5 is touching $\partial\mathcal{P}$.

Second perturbation. We now transform (see Figure 3.4b) the path $\Pi' = S_1'C_2'C_3'S_5'$ into a path $\Pi''(\varepsilon_1, \varepsilon_2) = S_1''C_2''C_3''C_4''S_5''$ by reducing the length of S_1' by a small value ε_2 (i.e., $\|S_1''\| = \|S_1'\| - \varepsilon_2$), such that the length of C_2' does not change (i.e., $\|C_2''\| = \|C_2'\|$) and the new segment C_4'' is oriented clockwise and is smaller than π (segments C_2'' and C_3'' are naturally oriented clockwise and counterclockwise respectively). For simplicity, we denote $\Pi''(\varepsilon_1, \varepsilon_2)$ by Π''.

CLAIM 3.8. *In an obstacle-free environment, path Π'' is shorter than Π', for any $\varepsilon_2 > 0$ small enough.*

Proof. Refer to Figure 3.4b. Let Y' be the first point of C_3' and Z' be the first point on C_3' such that the tangent to C_3 at Z' is parallel to the line segment S_1. Point Z' exists on C_3' because, by construction, $\|C_2'\| < \pi$ and $\|C_3'\| > \pi$. Furthermore, $\|C_2'\| = \|C_3'[Y', Z']\|$. We define similarly points Y'' and Z'' on C_3''. Then it follows from

$\|C_2'\| = \|C_2''\|$ that $\|C_3'[Y', Z']\| = \|C_3''[Y'', Z'']\|$. Thus if X' and X'' denote the first point of the segments C_2' and C_2'' respectively, we get $\|\Pi'[X', Z']\| = \|\Pi''[X'', Z'']\|$.

Now let $\tilde{I}$ be the point defined by $\overrightarrow{\tilde{I}Z'} = \overrightarrow{IX'}$ (or $\overrightarrow{I\tilde{I}} = \overrightarrow{X'Z'} = \overrightarrow{X''Z''}$). Let $\tilde{\Pi}'$ be the path from $\tilde{I}$ to F that is the concatenation of the line segment $\tilde{I}Z'$ and $\Pi'[Z', F]$. Similarly, let $\tilde{\Pi}''$ be the path from $\tilde{I}$ to F that is the concatenation of the line segment $\tilde{I}Z''$ and $\Pi''[Z'', F]$. We get, by construction, that $\|\tilde{\Pi}'\| = \|\Pi'\| - \|\Pi'[X', Z']\|$ and $\|\tilde{\Pi}''\| = \|\Pi''\| - \|\Pi''[X'', Z'']\|$. Since $\|\Pi'[X', Z']\| = \|\Pi''[X'', Z'']\|$, $\|\tilde{\Pi}'\| - \|\tilde{\Pi}''\| = \|\Pi'\| - \|\Pi''\|$. But we know that $\tilde{\Pi}''$ is shorter than $\tilde{\Pi}'$ because $\tilde{\Pi}''$ is obtained by the Dubins' length-reducing perturbation, shown in Figure 3.4a, applied on the reversed path of $\tilde{\Pi}'$ which is of type SCS. Thus Π'' is shorter than Π' in an obstacle-free environment, for any $\varepsilon_2 > 0$ small enough. □

CLAIM 3.9. *If the C-segment C_3 in the optimal path $\Pi = S_1C_3S_5$ is not strongly $\mathcal{PP}$-anchored, then we can choose ε_1 and ε_2 arbitrarily small such that Π'' is free in $\mathcal{P}$.*

Proof. Let $\vec{u}$ be the unit vector whose direction is opposite to the direction of S_1, and let $\vec{v}$ be the unit vector whose direction is the same as the direction of S_5. See Figure 3.4a.

In the two perturbations described above, the lengths of C_3 and C_3' increase. More precisely, the translated copy of C_3 by vector $\varepsilon_1\vec{v}$ is part of C_3', and the translated copy of C_3' by vector $\varepsilon_2\vec{u}$ is part of C_3''. Thus if Z'' and T'' are the translated copies of the endpoints of C_3 by vector $\varepsilon_1\vec{v}+\varepsilon_2\vec{u}$, Π'' is the concatenation of $\Pi''[I, Z'']$, $\Pi''[Z'', T'']$, and $\Pi''[T'', F]$, where $\Pi''[Z'', T'']$ is the the translated copy of C_3 by vector $\varepsilon_1\vec{v}+\varepsilon_2\vec{u}$. On the other hand, any arbitrarily small neighborhood around S_1 (respectively S_5) contains $\Pi''[I, Z'']$ (respectively $\Pi''[T'', F]$), for any ε_1 and ε_2 small enough. Thus, since we assumed that neither S_1 nor S_5 touches $\partial\mathcal{P}$, neither $\Pi''[I, Z'']$ nor $\Pi''[T'', F]$ touches $\partial\mathcal{P}$, for any ε_1 and ε_2 small enough. Thus it is sufficient to show that, if C_3 is not strongly $\mathcal{PP}$-anchored, then we can choose ε_1 and ε_2 arbitrarily small such that the translated copy of C_3 by vector $\varepsilon_1\vec{v} + \varepsilon_2\vec{u}$ is free.

Let A be the last point of tangency on C_3 with $\partial\mathcal{P}$. Let $\vec{w}$ be the unit vector tangent to C_3 at A (with direction corresponding to the orientation of C_3). If C_3 is not strongly $\mathcal{PP}$-anchored, then, for any $\mu > 0$ small enough, the translated copy of C_3 by $\mu\vec{w}$ is free. On the other hand, for λ_1 and λ_2 such that $\vec{w} = \lambda_1\vec{v} + \lambda_2\vec{u}$, the λ_1 and λ_2 are non-negative, because by Lemma 2.4, the distance on C_3 between the first point of C_3 and A is greater than π. Therefore, if C_3 is not strongly $\mathcal{PP}$-anchored, then, for any $\mu > 0$ small enough, the path Π'' defined with $\varepsilon_1 = \mu\lambda_1$ and $\varepsilon_2 = \mu\lambda_2$ is free in $\mathcal{P}$. □

By Claims 3.7, 3.8, and 3.9, if the C-segment C_3 in the optimal path $\Pi = S_1C_3S_5$ is not strongly $\mathcal{PP}$-anchored, then we can choose ε_1 and ε_2 arbitrarily small so that Π'' is free and shorter than Π. This concludes the proof of Lemma 3.4(i).

3.1.2. Proof of Lemma 3.4(ii). We prove Lemma 3.4(ii) which states:

> *If an optimal path has a subpath of type $C_1C_2C_3S$ (or $SC_3C_2C_1$) so that the C-segment C_2 does not touch $\partial\mathcal{P}$, then C_3 is strongly $\mathcal{PP}$-anchored.*

Consider an optimal path $\Pi = C_1C_2C_3S$ from configuration I to configuration F so that the C-segment C_2 does not touch $\partial\mathcal{P}$. We prove that the C-segment C_3 is strongly $\mathcal{PP}$-anchored. Assume without loss of generality that C_1 is oriented counterclockwise. We assume for a contradiction that C_3 is not strongly $\mathcal{PP}$-anchored and show that the path Π can be shortened in $\mathcal{P}$.

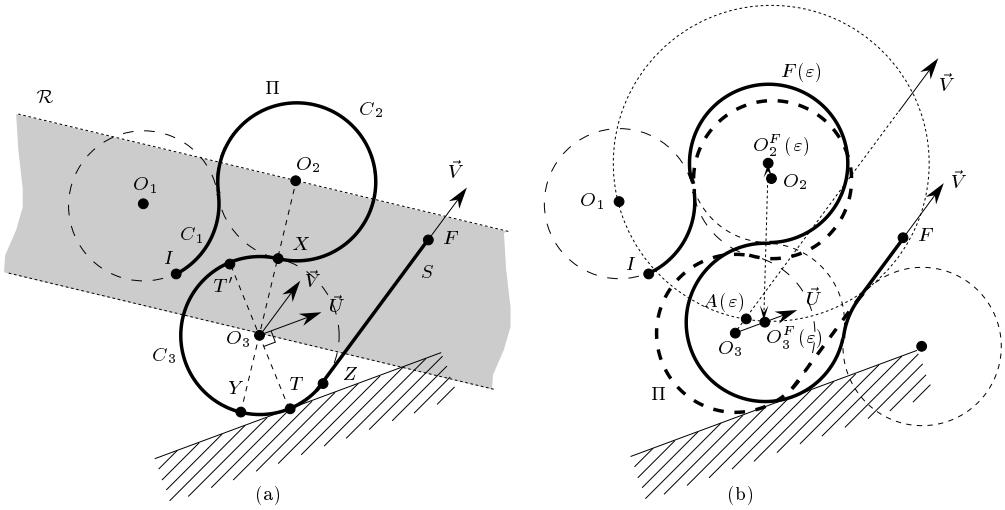


FIG. 3.5. *(a): Path* Π. *(b): A shorter path* $F(\varepsilon)$.

Notation. Let O_i, $i = 1, 2, 3$, denote the center of the circle supporting the C-segment C_i of Π. See Figure 3.5a. Let X denote the point of tangency between the C-segments C_2 and C_3, and let Y denote the antipodal point of X on C_3 ($\|C_3\| > \pi$, by Lemma 2.3). Let Z be the point of tangency between the C-segment C_3 and the S-segment. Let T be the last point of tangency with $\partial\mathcal{P}$ on the (oriented) circular arc $\Pi[Y, Z]$; such a point exists by Lemma 2.4. Let T' be the antipodal point of T on C_3. By definition, the circular arc $\Pi(T, Z]$ does not touch $\partial\mathcal{P}$. Furthermore, the circular arc $\Pi[X, T')$ does not touch $\partial\mathcal{P}$ because otherwise C_3 would be strongly $\mathcal{PP}-$*anchored*. Let $\vec{U}$ be a unit vector tangent to Π at T, and let $\vec{V}$ be a unit vector tangent to Π at any point on the S-segment of Π. Finally, let $ray(O_3, \vec{U})$ and $ray(O_3, \vec{V})$ denote the rays starting at O_3 in the directions $\vec{U}$ and $\vec{V}$, respectively.

First perturbation. Consider the perturbation shown as a thick solid path in Figure 3.5b, where Π has been perturbed into a path $F(\varepsilon)$ of type $CCCCS$. We define this perturbed path $F(\varepsilon)$ by specifying the position of the center $O_3^F(\varepsilon)$ of the supporting circle of its third circular arc, namely, $O_3^F(\varepsilon) = O_3 + \varepsilon\vec{U}$. The path $F(\varepsilon)$ is well defined for any ε small enough because the unit circle centered at $O_3^F(\varepsilon)$ intersects the S-segment of Π (by definition of $\vec{U}$); thus the fourth circular arc of $F(\varepsilon)$ is defined. Also, since the second C-segment of Π and the arcs $\Pi[X, T')$ and $\Pi(T, Z]$ do not touch $\partial\mathcal{P}$, $F(\varepsilon)$ is free for any sufficiently small values of $\varepsilon > 0$.

Second perturbation. To prove that the length of $F(\varepsilon)$ is a decreasing function of ε, we define a path $K_h(k)$ with two perturbation steps. The first step perturbs Π to a path $H(h)$, and the second step then perturbs $H(h)$ to a path $K_h(k)$. As we will show later, ε, h, and k can be chosen so that $F(\varepsilon) = K_h(k)$. Furthermore, we will show that $K_h(k)$ is shorter than Π.

Below, let $O_i^H(h)$, $O_i^K(k)$, $i = 1, 2, 3, 4$, denote the center of the circle supporting the i-th C-segment of paths $H(h)$ and $K_h(k)$, respectively.

First step. The path $H(h)$ (see Figure 3.6a) is of type $CCCS$ joining I to F such that $O_1^H(h)$ is identically equal to O_1, $O_3^H(h) = O_3 + h\vec{V}$, and the length of the second

circular arc is greater than π. For any h small enough, $H(h)$ is well defined.

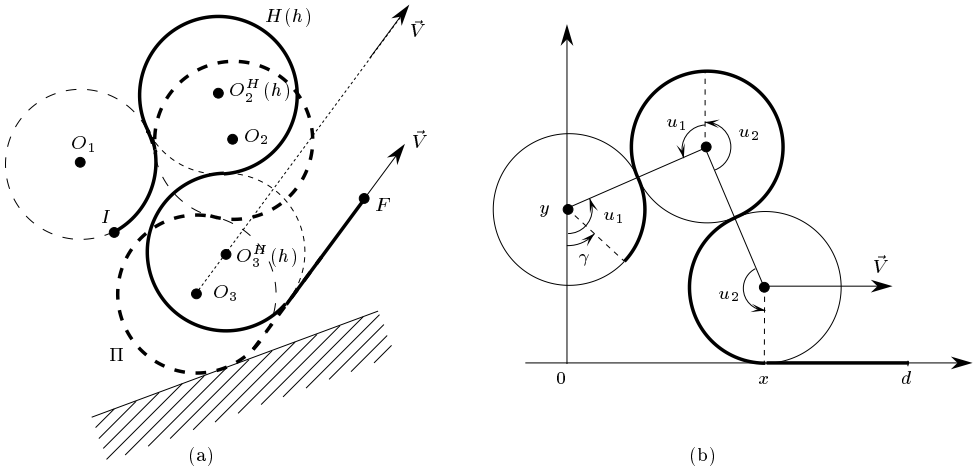


FIG. 3.6. *Length reducing perturbation: the path $H(h)$ is shorter than* Π.

Second step. The path $K_h(k)$ (see Figure 3.7) is defined as follows. For a given $H(h)$ and for a sufficiently small $k > 0$, path $K_h(k)$ is of type $CCCCS$ from I to F such that $O_1^K(k)$ and $O_2^K(k)$ are identically equal to O_1 and $O_2^H(h)$, respectively, $O_3^K(k)$ is at distance k counterclockwise from $O_3^H(h)$ along the circular arc of radius 2 centered at $O_2^K(k)$, and the fourth circular arc has length less than π.

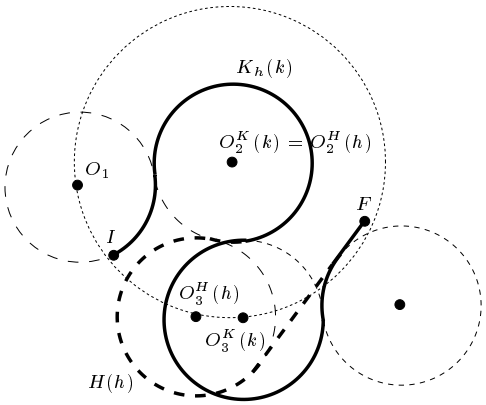


FIG. 3.7. *Dubins' length reducing perturbation: the path $K_h(k)$ is shorter than $H(h)$.*

CLAIM 3.10. *For a given sufficiently small $\varepsilon > 0$, h and k can be chosen such that $F(\varepsilon) = K_h(k)$.*

Proof. Let $\mathcal{R}$ be the open strip bounded by two lines perpendicular to the line O_2O_3 and passing through O_2 and O_3 respectively (see Figure 3.5a). Since the lengths of the circular arcs $\Pi[X,T]$ and $\Pi[X,Z]$ are greater than π by Lemmas 2.4 and 2.3(i), respectively, $ray(O_3,\vec{U})$ and $ray(O_3,\vec{V})$ are directed into $\mathcal{R}$. Thus for any sufficiently small ε, $O_3^F(\varepsilon)$ belongs to $\mathcal{R}$, and the distance between $O_3^F(\varepsilon)$ and O_2 is less than 2. Consequently, it can be shown that $O_2^F(\varepsilon)$ must lie outside $\mathcal{R}$ on the circle of radius 2 centered at O_1. It follows that, for any open neighborhood $\mathcal{N}$ of O_3, any choice of ε sufficiently small ensures that the circle of radius 2 centered at $O_2^F(\varepsilon)$

intersects $ray(O_3, \vec{U})$ and $ray(O_3, \vec{V})$ in $\mathcal{N}$ (see Figure 3.5b). Let $A(\varepsilon)$ denote the intersection of that circle with $ray(O_3, \vec{V})$ in $\mathcal{N}$; recall that the intersection, in $\mathcal{N}$, of that circle with $ray(O_3, \vec{U})$ is $O_3^F(\varepsilon)$. The polar angle of $ray(O_3, \vec{V})$ is bigger than the polar angle of $ray(O_3, \vec{U})$ by an amount smaller than π. Thus for ε small enough, the counterclockwise oriented arc, denoted $arc(A(\varepsilon), O_3^F(\varepsilon))$, of the circle of radius 2 centered at $O_2^F(\varepsilon)$ starting at $A(\varepsilon)$ and ending at $O_3^F(\varepsilon)$ is also contained in the neighborhood $\mathcal{N}$; indeed, since $A(\varepsilon)$ and $O_3^F(\varepsilon)$ converge to O_3 when ε tends to 0, $arc(A(\varepsilon), O_3^F(\varepsilon))$ also tends to O_3 when ε tends to 0. Therefore, we can choose ε such that the line segment $[O_3, A(\varepsilon)]$ and the circular arc $arc(A(\varepsilon), O_3^F(\varepsilon))$ are arbitrarily small.

Choose h equal to the length of the line segment $[O_3, A(\varepsilon)]$ and choose k equal to the length of the circular arc $arc(A(\varepsilon), O_3^F(\varepsilon))$. Then $O_3^H(h) = A(\varepsilon)$ and $O_3^K(k) = O_3(\varepsilon)$, and therefore $K(k) = F(\varepsilon)$. Moreover, we have shown that we can choose ε small enough such that h and k are arbitrarily small. □

CLAIM 3.11. *The length of $K_h(k)$ is strictly smaller than the length of Π, for any h and k sufficiently small.*

Proof. The length of $K_h(k)$ has been shown by Dubins [15] to be strictly shorter than the length of $K_h(0) = H(h)$ for any fixed h and for any small enough $k > 0$. Furthermore, the length of $K_h(0) = H(h)$ has been shown in [7] to be strictly shorter than the length of Π. For completeness, we give here the proof. Consider a path of type $CCCS$ such that the length of the second circular arc is greater than π. With the notation of Figure 3.6b, the length of the path is equal to $L = 2(u_1 + u_2) - \gamma + d - x$, where γ and d are some constants. Furthermore, we have :

$$\begin{cases} \sin(u_1) + \sin(u_2) = x/2 \\ \cos(u_1) - \cos(u_2) = (y-1)/2 \ . \end{cases}$$

By computing the derivative of each equation with respect to x and solving the system, we obtain the following solution (which is defined, since $(u_1 + u_2) \in (\pi, 2\pi)$ by hypothesis) :

$$\begin{cases} \dfrac{\partial u_1}{\partial x} = \dfrac{\sin(u_2)}{2\sin(u_1+u_2)} \\ \dfrac{\partial u_2}{\partial x} = \dfrac{\sin(u_1)}{2\sin(u_1+u_2)} \ . \end{cases}$$

Therefore,

$$\frac{\partial L}{\partial x} = \frac{\sin(u_1)+\sin(u_2)}{\sin(u_1+u_2)} - 1 = \frac{\cos\left(\frac{u_1-u_2}{2}\right) - \cos\left(\frac{u_1+u_2}{2}\right)}{\cos\left(\frac{u_1+u_2}{2}\right)}.$$

Since u_1 and u_2 are positive and $(u_1 + u_2) \in (\pi, 2\pi)$, $0 \leqslant \frac{|u_1-u_2|}{2} < \frac{|u_1+u_2|}{2} < \pi$ and thus, $\cos\left(\frac{u_1-u_2}{2}\right) > \cos\left(\frac{u_1+u_2}{2}\right)$. Furthermore, $\cos\left(\frac{u_1+u_2}{2}\right) < 0$ since $\frac{u_1+u_2}{2} \in (\frac{\pi}{2}, \pi)$. Therefore, $\partial L/\partial x$ is negative, which means that, as long as $H(h)$ is of type $CCCS$ with the second circular arc greater than π, the length of $H(h)$ is a decreasing function of h. Hence we have shown that the length of $K_h(k)$ is smaller than the length of Π for any h and k sufficiently small. □

To complete the proof of Lemma 3.4(ii), recall that $F(\varepsilon)$ is free in $\mathcal{P}$ for any ε small enough. By Claims 3.10 and 3.11, there exist arbitrarily small values of ε, h, k such that $F(\varepsilon) = K_h(k)$, where the length of $K_h(k)$ is strictly less than the length of Π. This contradicts the optimality of Π and completes the proof.

4. A Simple Algorithm. Theorem 3.1 can be used to obtain the following simple algorithm for computing an optimal path inside $\mathcal{P}$. We enumerate candidate paths of types described in Theorem 3.1. Our candidate set is guaranteed to contain an optimal path, if any exist. For each such path, we check whether it is feasible, and if so, compute its length. Finally, we either return the shortest feasible path, or report that no feasible path exists.

In order to determine whether a path is feasible, we rely on the circle-shooting data structure by Cheng *et al.* [10] that preprocesses $\mathcal{P}$ in $O(n \log n)$ time into a data structure of linear size that makes it possible to determine in $O(\log n)$ time whether a given circular arc of unit radius intersects $\partial\mathcal{P}$. This immediately implies the following lemma.

LEMMA 4.1. *$\mathcal{P}$ can be preprocessed in $O(n \log n)$ time into a data structure of linear size that enables us to determine in $O(m \log n)$ time whether a given path consisting of m C- and S-segments is feasible.*

To bound the running time of this simple algorithm, we must count the number of candidate paths to check. We note that once a path type is given, and the supporting circles for C-segments are known, there are $O(1)$ candidate paths. These are determined by the choices of the orientations for the C-segments. Hence we are interested in the number of possible supporting circles for each path type. Note that there may be $\Omega(n^2)$ $\mathcal{PP}$-anchored circles (see Lemma 5.3) and $\Omega(n)$ $\mathcal{PC}$-anchored circles.

There are $O(1)$ Dubins path candidates. For paths of type (A.i) and (B.ii), once the $\mathcal{PP}$-anchored circle is chosen, there are $O(1)$ choices for other supporting circles, and hence $O(1)$ candidate paths. Since there are $O(n^2)$ $\mathcal{PP}$-anchored circles, there are $O(n^2)$ candidate paths for these two path types.

A path of type (A.ii) may have up to two $\mathcal{PC}$-anchored segments. Once their supporting circles are chosen, there are $O(1)$ path candidates. There are $O(n)$ potential $\mathcal{PC}$-anchored circles. If both anchored segments are present, we have $O(n^2)$ paths to check; otherwise, we have only $O(n)$. Paths of type (B.i) are also determined by a $\mathcal{PC}$-anchored circle; hence there are $O(n)$ of them.

Paths of type (B.iii), i.e., of type $C_I\bar{\bar{C}}_1S\bar{C}_2\bar{C}_3S\bar{\bar{C}}_4C_F$, present a special problem. If we know the supporting circles of the $\bar{C}\bar{C}$ subpath, the rest of the path is determined by a pair of $\mathcal{PC}$-anchored circles $\mathcal{C}_1, \mathcal{C}_4$, for which there are $O(n^2)$ possibilities. Unfortunately, there is an infinite family of supporting circles for the $\bar{C}\bar{C}$ subpath. The following result by Boissonnat and Lazard [6] allows us to consider only a finite set of $\bar{C}\bar{C}$ subpaths.

LEMMA 4.2 (Boissonnat and Lazard [6]). *Given two configurations X and Y, and two edges e, e' of $\mathcal{P}$, we can compute[4] in $O(1)$ time a finite set of paths from X to Y of type $C_1S\bar{C}_2\bar{C}_3SC_4$, where $\bar{C}_2$ and $\bar{C}_3$ are tangent to edges e and e', respectively. This set contains all optimal paths from X to Y of type $C_1S\bar{C}_2\bar{C}_3SC_4$.*

Given a pair of edges e, e' and a pair of $\mathcal{PC}$-anchored circles $\mathcal{C}_1, \mathcal{C}_4$, tangent to C_I and C_F, respectively, we choose X to be the configuration determined by the intersection of C_I and $\mathcal{C}_1$ and Y to be the configuration determined by C_F and $\mathcal{C}_4$. By the above lemma, we can compute in $O(1)$ time a constant number of candidate paths for this pair of edges and anchored circles. Doing this for all possible pairs of edges (e, e'), and pairs of supporting circles $(\mathcal{C}_1, \mathcal{C}_4)$, we determine $O(n^4)$ path candidates of type (B.iii) in $O(n^4)$ time.

[4]The computation is performed by solving four algebraic systems of three equations in three indeterminates.

In summary, the simple algorithm examines $O(n^4)$ candidate paths, and for each, spends $O(\log n)$ time checking feasibility, by Lemma 4.1 with $m \leqslant 8$. Therefore, the overall running time is $O(n^4 \log n)$.

PROPOSITION 4.3. *Given a convex polygon $\mathcal{P}$, an initial configuration I, and a final configuration F, an optimal path from I to F inside $\mathcal{P}$ can be computed in time $O(n^4 \log n)$.*

5. An Efficient Algorithm. In this section we prove additional properties of optimal paths that significantly reduce the number of candidates to examine. We have already shown that we need to consider only $O(1)$ Dubins paths and $O(n)$ candidates for paths of type (B.i). We will show that it suffices to consider only $O(1)$ candidate paths of type (A.i) and (B.ii), $O(n)$ candidate paths of type (A.ii), and $O(n^2)$ candidate paths of type (B.iii).

5.1. Computing paths of type (A.i) and (B.ii). The paths of types (A.i) and (B.ii) contain a strongly $\mathcal{PP}$-anchored C-segment $\bar{\bar{C}}$. The circle $\bar{\bar{\mathcal{C}}}$ supporting $\bar{\bar{C}}$ defines one or two pockets that contain a point of tangency of $\bar{\bar{\mathcal{C}}}$ with $\partial\mathcal{P}$ (see Figures 2.4b and 3.2). By Lemma 2.7, we know that I and F must belong to these pockets. The following lemma states that there exist at most two circles with these properties, and that they can be computed efficiently.

LEMMA 5.1. *For a fixed pair of locations I, F, there exist at most two circles that can support a strongly $\mathcal{PP}$-anchored C-segment appearing in an optimal path from I to F, and they can be computed in $O(n)$ time.*

Proof. Consider a strongly $\mathcal{PP}$-anchored segment that lies in an optimal path. Let X and Y denote its points of tangency with $\partial\mathcal{P}$, and let $\mathcal{C}$ denote its supporting circle. Assume without loss of generality that the short arc on $\mathcal{C}$ joining X and Y is $\mathcal{C}^-[X, Y]$ (see Figure 5.1). The proof of the lemma is divided into two cases, which depend on whether or not $\mathcal{C}$ is free.

Case 1: $\mathcal{C}$ is free. The center of $\mathcal{C}$ lies at a vertex of the retracted polygon of $\mathcal{P}$ (i.e., the set of points p such that the unit circle centered at p lies inside $\mathcal{P}$). By computing the retracted polygon of $\mathcal{P}$ in linear time, we get (in linear time) the set of the $O(n)$ free $\mathcal{PP}$-anchored circles, which contains $\mathcal{C}$. Each of these circles defines one pocket, and all these pockets are pairwise disjoint (see Figure 5.1). Thus only one of these pockets contains I and F; hence, by Lemma 2.7, $\mathcal{C}$ must be the circle defining this pocket. For each of the $O(n)$ free $\mathcal{PP}$-anchored circles, we can easily check, in $O(1)$ time, whether I and F belong to the corresponding pocket. Indeed (see Figure 5.1), $I, F \in \mathcal{P}$ belong to the pocket if and only if I and F are outside the circle and are in the small wedge defined by the rays emanating from the center of the circle and passing through the points of tangency of the circle with $\partial\mathcal{P}$.

Case 2: $\mathcal{C}$ is not free. $\mathcal{C}$ defines two pockets, $\Lambda_{\mathcal{C}}[X, X']$ and $\Lambda_{\mathcal{C}}[Y, Y']$ (see Figure 5.2). By Lemma 2.7, one of these pockets contains I and the other contains F. Thus $I \neq F$ and the line L_{IF} through I and F is defined. Let A and B denote the intersection points of L_{IF} with $\partial\mathcal{P}$.

With no loss of generality, suppose $I \in \Lambda_{\mathcal{C}}[X, X']$ and $F \in \Lambda_{\mathcal{C}}[Y, Y']$. The segment $[I, F]$ must pass through $\mathcal{C}$ twice since I and F are in distinct pockets. Thus the ray emanating from I in the direction $\overrightarrow{FI}$ cannot intersect $\mathcal{C}$, and therefore leaves $\Lambda_{\mathcal{C}}[X, X']$ through $\partial\mathcal{P}^-[X, X']$. Thus $A \in \partial\mathcal{P}^-[X, X']$. A similar argument shows $B \in \partial\mathcal{P}^+[Y, Y']$. Hence, $X', Y' \notin \partial\mathcal{P}^+[A, B]$ and $X, Y \in \partial\mathcal{P}^+[A, B]$.

The chain $\partial\mathcal{P}^+[A, B]$ does not properly intersect $\mathcal{C}$. Indeed (see Figure 5.2), it properly intersects neither the long arc $\mathcal{C}^+[X, Y]$, by assumption, nor the small arc

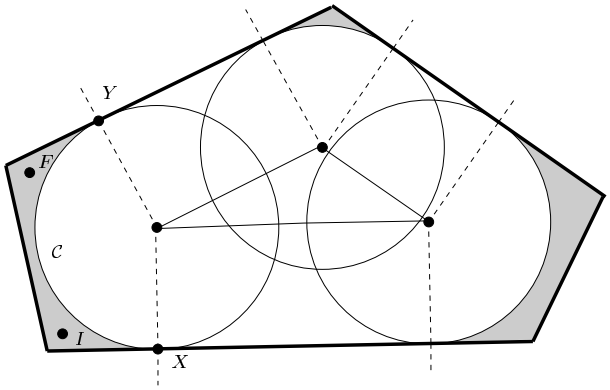


FIG. 5.1. *The free circle $\mathcal{C}$ is centered on the retracted polygon of $\mathcal{P}$.*

$\mathcal{C}^-[X,Y]$ because the first intersection between $\mathcal{C}^-[X,Y]$ (respectively $\mathcal{C}^+[Y,X]$) and $\partial\mathcal{P}$ is X' (respectively Y'), which does not belong to $\partial\mathcal{P}^+[A,B]$. It then follows from $X,Y \in \partial\mathcal{P}^+[A,B]$ that the circle $\mathcal{C}$ is a free anchored circle in the polygon $\mathcal{P}'$ obtained by extending the two edges of $\partial\mathcal{P}^+[A,B]$ that end at A,B (see Figure 5.2). Moreover, the pocket defined by $\mathcal{C}$ in $\mathcal{P}'$ contains the two pockets $\Lambda_{\mathcal{C}}[X,X']$ and $\Lambda_{\mathcal{C}}[Y,Y']$, and thus contains I and F. As before, at most one free anchored circle in $\mathcal{P}'$ defines a pocket containing I and F, and given $\mathcal{P}'$, it can be computed in $O(n)$ time.

Note finally that polygon $\mathcal{P}'$ can be determined in $O(n)$ time. This is because I and F determine the points A and B, and the turning angle of $\partial\mathcal{P}^+[A,B]$ is bigger than π. Thus, independent of any assumption about the orientation of the short arc on $\mathcal{C}$ joining X and Y, determining whether $\mathcal{P}'$ is the polygon obtained by extending the two edges (ending at A,B) of $\partial\mathcal{P}^+[A,B]$ or of $\partial\mathcal{P}^-[A,B]$ can be done simply by checking which of the turning angles of $\partial\mathcal{P}^+[A,B]$ and $\partial\mathcal{P}^-[A,B]$ is bigger than π.

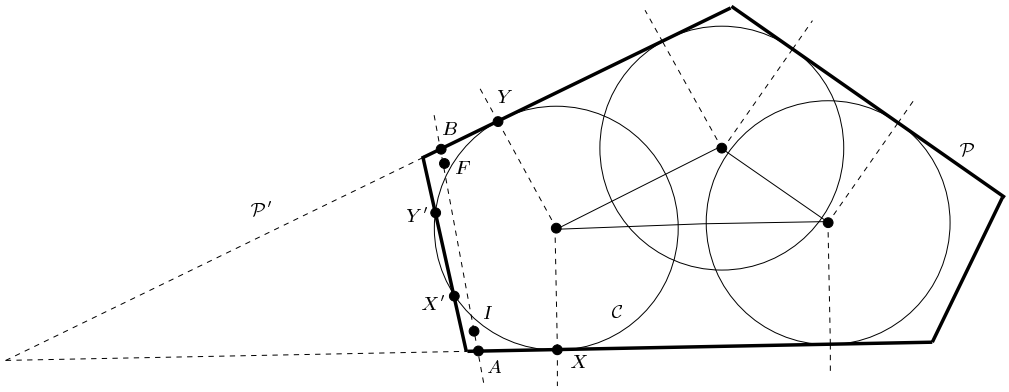


FIG. 5.2. *The non-free circle $\mathcal{C}$ is centered on the retracted chain of $\partial\mathcal{P}^+[A,B]$.*

Hence we have proved that for a fixed pair of locations I and F, there exist at most two $\mathcal{PP}$-anchored circles (one free and the other non-free) that can appear in an optimal path from I to F, and they can be computed in $O(n)$ time. □

By the above lemma, we can compute, in $O(n)$ time, a set of $O(1)$ candidate paths of types (A.i) and (B.ii). The candidate paths may be checked for feasibility in $O(\log n)$ time. Therefore:

LEMMA 5.2. *An optimal path of type (A.i) or (B.ii) can be computed in $O(n)$ time.*

The following proposition shows that Lemma 5.1 is essential for establishing the linear running time given in Lemma 5.2; indeed, an algorithm that checked all candidate strongly $\mathcal{PP}$-anchored circles to produce candidate paths would have running time $\Omega(n^2)$.

PROPOSITION 5.3. *There exist convex n-polygons that have $\Omega(n^2)$ strongly $\mathcal{PP}$-anchored circles.*

Proof. Let $n \geqslant 8$ be a multiple of 4, and let $0 < \epsilon < \tan(\pi/n)$ be a real parameter. Let $\mathcal{P}$ be an n-regular polygon centered at the origin with in-radius $(1+\epsilon)$, i.e., the distance from the origin to each side of $\mathcal{P}$ is $(1+\epsilon)$. We assume that one of the edges, say s_0, of $\mathcal{P}$ is parallel to the x-axis and lies below the x-axis; see Figure 5.3. The coordinates of the right endpoint of s_0 are $(a, 0)$, where $a = (1+\epsilon)\tan(\pi/n)$. Let the edges of $\mathcal{P}$ in the counterclockwise sense be $s_0, s_1, s_2, \ldots, s_{n-1}$.

The retracted polygon of $\mathcal{P}$ (i.e., the set of points p such that the unit circle centered at p lies inside $\mathcal{P}$) is an n-regular polygon with radius ϵ. Denote its sides by $t_0, t_1, \ldots, t_{n-1}$, where t_i is the retraction of s_i, for $0 \leqslant i < n$. For $i = 1, \ldots, n/4$, denote by $p_i = (x_i, -\epsilon)$ the intersection point of the lines supporting t_i and t_0, and by $\mathcal{C}_i$ the unit circle centered at p_i. It is easily seen that $0 < x_1 < \cdots < x_{n/4} \leqslant \epsilon$.

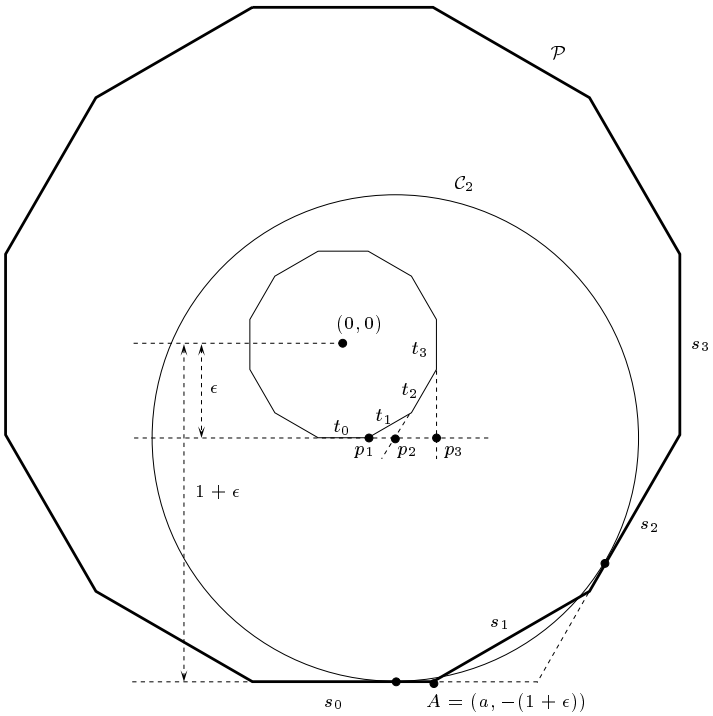


FIG. 5.3. *The circles $\mathcal{C}_i$ centered at p_i, $i = 1,\ 2,\ 3 = \frac{n}{4}$, are strongly $\mathcal{PP}$-anchored: they are tangent to s_0 and s_i, and their long arcs are free.*

Since the x-coordinate of the point at which $\mathcal{C}_i$ touches the line supporting s_0 is x_i and $x_i < \epsilon < \tan(\pi/n) < a$, $\mathcal{C}_i$ is tangent to s_0, for any $1 \leqslant i \leqslant n/4$. A symmetric argument shows that $\mathcal{C}_i$ is tangent to s_i.

Therefore, we can assign $n/4$ $\mathcal{PP}$-anchored circles to every side of $\mathcal{P}$, and the

number of $\mathcal{PP}$-anchored circles of $\mathcal{P}$ is $\Omega(n^2)$. It remains to show that these $\mathcal{PP}$-anchored circles are strongly $\mathcal{PP}$-anchored.

As can be seen from Figure 5.3, the point p_i is a vertex of the polygon defined by the two lines through t_0 and t_i, and the edges $t_{i+1}, \ldots, t_{n-1}$. Thus p_i is in the retracted polygon of the polygon formed by the two lines through s_0 and s_i, and the edges $s_{i+1}, \ldots, s_{n-1}$. Therefore, $\mathcal{C}_i$ does not properly intersect any of $s_{i+1}, \ldots, s_{n-1}$, and its long arc is free, so $\mathcal{C}_i$ is indeed a strongly $\mathcal{PP}$-anchored circle. □

5.2. A monotonicity property of CCSC paths. Subpaths of type $CCSC$ occur in both (A.ii) and (B.iii) path types. In this subsection, we ignore the polygon $\mathcal{P}$ and study paths from X to Y of type $C_1C_2SC_3$, with specified orientations on C_1 and C_3; the orientation of C_1 fixes the orientation of C_2, namely if C_1 is oriented clockwise (respectively counterclockwise), then C_2 is oriented counterclockwise (respectively clockwise).

The positions of circles $\mathcal{C}_1$ and $\mathcal{C}_3$, supporting C_1 and C_3, respectively, are considered fixed while the position of circle $\mathcal{C}_2$ is determined by M, its point of tangency with $\mathcal{C}_1$. The orientations of these three circles are fixed by the orientations of the corresponding C-segments. The S-segment is determined by the appropriate tangent line, given the orientations on C_2 and C_3. This tangent, if it exists, is unique.

For each $M \in \mathcal{C}_1$, there is at most one path, denoted $\Pi(M)$, from X to Y of type $C_1C_2SC_3$ with the specified orientations on the C_1- and C_3-segments. We are interested in how the path length $\|\Pi(M)\|$ changes as M moves along $\mathcal{C}_1$, in the same direction as the segment C_1.

For certain positions of M, one or more of the segments of $\Pi(M)$ may vanish. For example, when $M = X$, the length of the first segment C_1 changes discontinuously from 2π to 0. At such points the path length may change discontinuously, so these positions of M are called *singular points* of $\Pi(M)$.

LEMMA 5.4. *Given two configurations X and Y, the paths $\Pi(M)$ of type $C_1C_2SC_3$ from X to Y with specified orientations on the C_1- and C_3-segments admit at most six singular points, and their locations can be computed in $O(1)$ time.*

Proof. We enumerate the possibilities for a segment to vanish in the paths $\Pi(M)$. Figure 5.4 illustrates the six singular points in a path of type $C_1^+C_2^-SC_3^+$.

Segment C_1 vanishes if and only if $M = X$, so X is the only singular point such that C_1 vanishes.

Segment C_2 vanishes if and only if the path type degenerates to C_1SC_3. Then the point M on $\mathcal{C}_1$ is also on the S-segment. Since there is at most one S-segment tangent to $\mathcal{C}_1$ and $\mathcal{C}_3$ that respects their specified orientations, there is at most one singular point $M_1 \in \mathcal{C}_1$ such that C_2 vanishes.

Segment S vanishes if and only if the path type degenerates to either $C_1C_2C_3$ if C_2 and C_3 have opposite orientations, or C_1C_2 otherwise. Thus there are at most two singular points $M_2, M_3 \in \mathcal{C}_1$ (common point of C_1 and C_2) such that S vanishes.

Segment C_3 vanishes if and only if the path type degenerates to C_1C_2S. This means that L_Y, the line passing through the configuration Y, is tangent to $\mathcal{C}_2$. There are at most two circles $\mathcal{C}_2$ tangent to $\mathcal{C}_1$ and L_Y that respect the orientations of $\mathcal{C}_1$ and L_Y. Thus there are at most two singular points M_4, M_5 such that C_3 vanishes.

In total, there are no more than six singular points, and they can clearly be computed in $O(1)$ time. □

We next state the monotonicity property of the paths $\Pi(M)$.

LEMMA 5.5. *$\|\Pi(M)\|$ strictly increases as M moves along the oriented circle $\mathcal{C}_1$ between singular points, except when $\mathcal{C}_1$ and $\mathcal{C}_3$ are identical with the same orientation,*

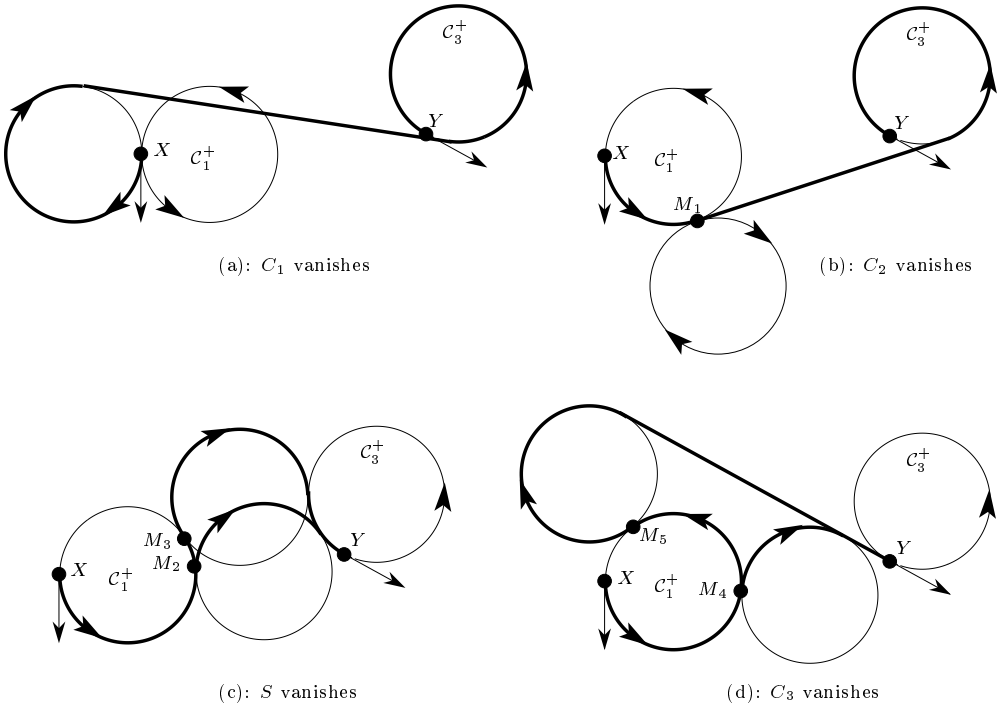


FIG. 5.4. *Paths of type $C_1^+C_2^-SC_3^+$ from X to Y and the six singular points X, M_1, M_2, M_3, M_4, and M_5 on $\mathcal{C}_X^+$.*

in which case $\|\Pi(M)\|$ is constant as M moves between singular points.

Proof. There are four possible orientation assignments for the circles: $C_1^+C_2^-SC_3^+$, $C_1^+C_2^-SC_3^-$, $C_1^-C_2^+SC_3^-$, and $C_1^-C_2^+SC_3^+$. We prove the claim for the first two cases; the other two cases can be proved using a symmetric argument.

Consider a path $\Pi(M)$. Let $\alpha_1 = \|\mathcal{C}_1^+[X, M]\|$ be the length of the first C-segment. Since there is a one-to-one mapping between α_1 and M, we can parameterize Π by α_1. Let $\alpha_2 = \alpha_2(\alpha_1)$, $\alpha_3 = \alpha_3(\alpha_1)$ be the length of the second and third C-segments of $\Pi(\alpha_1)$, and let $2s = 2s(\alpha_1)$ be the length of the S-segment of $\Pi(\alpha_1)$. Let O_i be the center of the circle $\mathcal{C}_i$ supporting C_i, $i = 1, 2, 3$. Although O_1 and O_3 are fixed, O_2 depends on α_1. By definition

$$L(\alpha_1) = \|\Pi(\alpha_1)\| = \alpha_1 + \alpha_2 + 2s + \alpha_3. \tag{5.1}$$

As M moves continuously on $\mathcal{C}_1$, the length of every segment in path $\Pi(M)$ changes continuously, except at singular points and at points for which $\Pi(M)$ is not defined (i.e., when $\mathcal{C}_2$ and $\mathcal{C}_3$ have opposite orientation and properly intersect). It follows that the segment lengths are piecewise differentiable functions of α_1, and that L is a piecewise differentiable function of α_1 on the intervals of $[0, 2\pi)$ where the path $\Pi(\alpha_1)$ is defined. For a function $f(\alpha_1)$, we will use $f'(\alpha_1)$ to denote $\partial f/\partial\alpha_1$. Then

$$L'(\alpha_1) = 1 + \alpha_2' + 2s' + \alpha_3'. \tag{5.2}$$

We call a value of α_1 singular if the corresponding point M on $\mathcal{C}_1$ is singular. The lemma can now be restated as follows: In the open intervals between singular

points, $L' > 0$ almost everywhere (i.e., at all but a finite number of points) except when $O_1 = O_3$ and $\Pi(M)$ is of type $C_1^+C_2^-SC_3^+$ in which case $L' = 0$ everywhere. The proof is divided into two cases depending on whether O_1 and O_3 are equal.

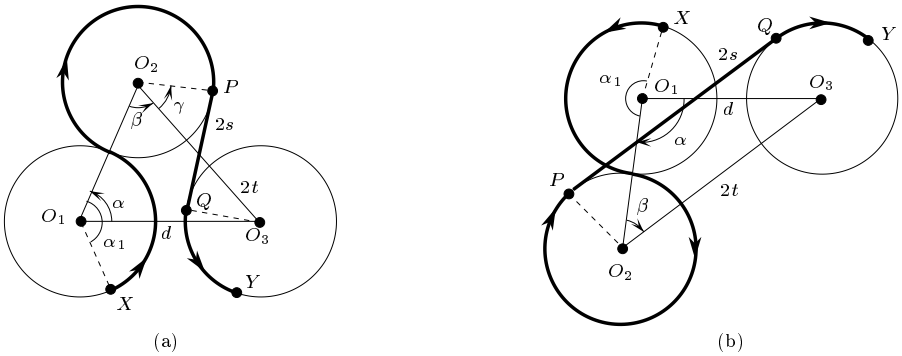


FIG. 5.5. *Path $\Pi(M)$ of type (a) $C_1^+C_2^-SC_3^+$, and (b) $C_1^+C_2^-SC_3^-$.*

Case 1: O_1 is distinct from O_3. Consider the triangle $\triangle O_1O_2O_3$. See Figure 5.5. We have $\|O_1O_2\| = 2$; let $d = \|O_1O_3\|$ and let $2t = \|O_2O_3\|$. We also define two (counterclockwise) oriented angles $\alpha = \angle O_3O_1O_2$ and $\beta = \angle O_1O_2O_3$. Both angles depend on α_1. Since C_1 is oriented counterclockwise, $\alpha_1 - \alpha$ is a constant, and therefore $\alpha' = 1$.

In view of the above discussion, it is sufficient to prove that $L'(\alpha_1) > 0$ for any nonsingular value of α_1 such that the path $\Pi(\alpha_1)$ is defined and $\alpha \not\equiv 0 \pmod{\pi}$; indeed there are only a finite number of values α_1 such that $\alpha \equiv 0 \pmod{\pi}$. Since α_1 is not singular, we can assume in the following that $t \neq 0$.

By applying the cosine law to $\triangle O_1O_2O_3$, we obtain

$$4t^2 = 4 + d^2 - 4d\cos\alpha.$$

By differentiating the above equality and noting that $\alpha' = 1$, we get $2tt' = d\sin\alpha$. Applying the sine law to $\triangle O_1O_2O_3$ gives $2t\sin\beta = d\sin\alpha$, because α and β have the same sign, by definition (see Figure 5.5). It follows that

$$t' = \sin\beta. \tag{5.3}$$

We first consider the case when $\Pi(M)$ is of type $C_1^+C_2^-SC_3^-$. Recall that $\psi(X)$ is the polar angle of the tangent vector for configuration X. This angle is constant for all configurations along an S-segment. On the other hand, the angle increases by δ after traversing a C^+-segment of length δ, and decreases by the same amount upon traversing a C^--segment of the same length. We therefore have the following

$$\psi(Y) \equiv \psi(X) + \alpha_1 - \alpha_2 - \alpha_3 \pmod{2\pi}. \tag{5.4}$$

Since X and Y are fixed, we have $1 - \alpha_2' - \alpha_3' = 0$ or $\alpha_2' + \alpha_3' = 1$. Substituting this into (5.2) gives

$$L' = 2 + 2s'.$$

The S-segment is a translate of the segment O_2O_3 (see Figure 5.5b). Thus $s = t$, and hence $s' = t' = \sin\beta$, by (5.3). Thus $L' = 2 + 2\sin\beta$. Since $\Pi(M)$ is of type $C^+C^-SC^-$,

$$\beta + \pi/2 + \alpha_2 \equiv 0 \pmod{2\pi}$$

(see Figure 5.5b); indeed, $\beta = \angle O_1O_2O_3$, $\pi/2 = \angle O_3O_2P$, and $\alpha_2 = \angle PO_2O_1$. Thus $\beta \equiv 3\pi/2 \pmod{2\pi}$ if and only if $\alpha_2 \equiv 0 \pmod{2\pi}$, which only occurs at a singular point (by definition of singular points). Therefore $L'(\alpha_1) > 0$ for any nonsingular value of α_1 for which the path $\Pi(\alpha_1)$ is defined.

We now turn to the case in which $\Pi(M)$ is of type $C_1^+C_2^-SC_3^+$. Then (5.4) is replaced by

$$\psi(Y) \equiv \psi(X) + \alpha_1 - \alpha_2 + \alpha_3 \pmod{2\pi},$$

so $\alpha_3' = \alpha_2' - 1$. Substituting this into (5.2) gives

$$L' = 2(s' + \alpha_2').$$

In order to find an expression for α_2', it is convenient to define the oriented angle $\gamma = \angle O_3O_2P$, where P is the common point between the segments C_2 and S (see Figure 5.5a). Recall that $\alpha_2 = \angle PO_2O_1$ and $\beta = \angle O_1O_2O_3$. Thus $\gamma + \alpha_2 + \beta \equiv 0 \pmod{2\pi}$, which implies $\gamma' + \alpha_2' + \beta' = 0$ and

$$L' = 2(s' - \gamma' - \beta'). \tag{5.5}$$

We now find expressions for $s' - \gamma'$ and β'.

With P and Q denoting, respectively, the first and last points of the S-segment, it is easy to see that the segments O_2O_3, PQ, O_2P, and O_3Q form two congruent right triangles (see Figure 5.5a). Thus we have $s^2 + 1 = t^2$, whence $ss' = tt' = t\sin\beta$, using (5.3). Further, $\tan\gamma = s$, so $\gamma' = s'\cos^2\gamma$. Combining the two results,

$$s' - \gamma' = s'\sin^2\gamma = s'\left(\frac{s}{t}\right)^2 = \frac{s}{t}\sin\beta. \tag{5.6}$$

The final derivative needed is β', which again follows from the cosine law applied to triangle $\triangle O_1O_2O_3$ (see Figure 5.5a):

$$d^2 = 4 + 4t^2 - 8t\cos\beta.$$

After a differentiation and rearrangement, this yields $t\beta'\sin\beta = t'(\cos\beta - t)$. Substituting for t' using (5.3) and noting that $\alpha \not\equiv 0 \pmod{\pi}$ implies $\sin\beta \neq 0$, we obtain

$$\beta' = \frac{1}{t}(\cos\beta - t). \tag{5.7}$$

Combining (5.5), (5.6) and (5.7) yields

$$L' = \frac{2}{t}(s\sin\beta + t - \cos\beta).$$

$\Pi(\alpha_1)$ is defined only when $\mathcal{C}_2$ and $\mathcal{C}_3$ do not properly intersect. Thus $t \geqslant 1$ wherever L is defined, thereby implying that $(t - \cos\beta) \geqslant 0$ and that

$$\begin{aligned}
L' \leqslant 0 &\Leftrightarrow t - \cos\beta \leqslant -s\sin\beta \\
&\Rightarrow t^2 - 2t\cos\beta + \cos^2\beta \leqslant s^2\sin^2\beta = (t^2-1)\sin^2\beta \\
&\Rightarrow t^2\cos^2\beta - 2t\cos\beta + 1 \leqslant 0 \\
&\Rightarrow (t\cos\beta - 1)^2 \leqslant 0 \\
&\Rightarrow \cos\beta = \frac{2}{2t}.
\end{aligned}$$

Hence $L' \leqslant 0$ only if $\angle O_3O_1O_2 = \alpha \equiv \pi/2 \pmod{\pi}$ (see Figure 5.5a). Therefore, $L' > 0$ almost everywhere it is defined.

Case 2: O_1 and O_3 are equal. When $\Pi(M)$ is of type $C_1^+C_2^-SC_3^-$ (see Figure 5.6a), $\alpha_2 = 3\pi/2$ and $s = 2$; thus $L' = 1 + \alpha_3'$. Equation (5.4) still holds, and thus $\alpha_3' = 1$. Therefore, $L' = 2 > 0$ everywhere it is defined.

When $\Pi(M)$ is of type $C_1^+C_2^-SC_3^+$ (see Figure 5.6b), the circles $\mathcal{C}_1$ and $\mathcal{C}_3$ coincide and have the same orientation. Thus both segments C_2 and S vanish, $\Pi(M)$ degenerates to a C-segment, and consequently $L' = 0$ everywhere except when $M = X$ or Y, where L is not differentiable. □

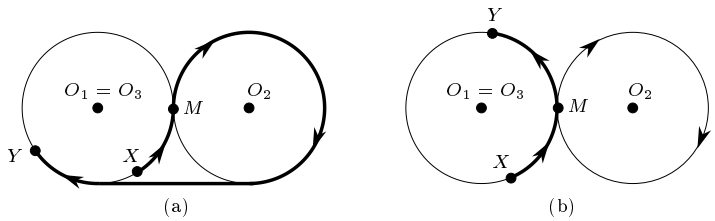


FIG. 5.6. *When $O_1 = O_3$: (a) $\Pi(M)$ is of type $C_1^+C_2^-SC_3^-$; (b) $\Pi(M)$ of type $C_1^+C_2^-SC_3^+$ degenerates to a single C-segment, for any $M \in \mathcal{C}_1$.*

5.3. Computing type (A.ii) paths. As mentioned in Section 4, we can compute in $O(n \log n)$ time the feasible candidates of type (A.ii) paths with at most one $\mathcal{PC}$-anchored segment. If the path is of type $C_I\bar{\bar{C}}S\bar{\bar{C}}C_F$, a naive analysis gives $O(n^2)$ candidates to check. Using Lemma 5.5, we reduce the number of candidates to $O(n)$ and compute them in $O(n \log n)$ time, as follows.

Fix the orientations of the terminal C-segments, and let $\mathcal{C}_I$ and $\mathcal{C}_F$ denote the circles supporting C_I and C_F, respectively. Let us assume that $\mathcal{C}_I$ is oriented counterclockwise. Let $\mathcal{K}_I$ be the sequence of $\mathcal{PC}$-anchored circles that touch $\mathcal{C}_I$ and that are free, sorted by their points of tangency with $\mathcal{C}_I$. The set $\mathcal{K}_F$ is defined analogously for $\mathcal{PC}$-anchored circles tangent to $\mathcal{C}_F$. The sets $\mathcal{K}_I$ and $\mathcal{K}_F$ can be computed in $O(n \log n)$ time, and they have $O(n)$ elements.

By Lemma 3.6, the circles $\bar{\bar{\mathcal{C}}}_1$ and $\bar{\bar{\mathcal{C}}}_2$ supporting the $\bar{\bar{C}}$-segments in an optimal path Π of type $C_I\bar{\bar{C}}_1S\bar{\bar{C}}_2C_F$ are free (otherwise, $\bar{\bar{C}}_1$ or $\bar{\bar{C}}_2$ would be a terminal C-segment). Therefore, the $\bar{\bar{C}}_1$-segment of Π lies on a circle of $\mathcal{K}_I$, and the $\bar{\bar{C}}_2$-segment lies on a circle of $\mathcal{K}_F$. Suppose $\mathcal{C}_2 \in \mathcal{K}_F$ supports the $\bar{\bar{C}}_2$-segment of Π. This fixes the terminal configuration of the subpath $C_I\bar{\bar{C}}_1S\bar{\bar{C}}_2$. By Lemma 5.4, there are at most six singular points, say, $\Sigma_0 = I, \ldots, \Sigma_5$, on $\mathcal{C}_I$ with respect to $\mathcal{C}_2$, sorted in the counterclockwise sense.

For $0 \leqslant i \leqslant 5$, let $\mathcal{K}_I(i) \subseteq \mathcal{K}_I$ be the contiguous subsequence of circles that touch $\mathcal{C}_I$ at a point in the arc $\mathcal{C}_I[\Sigma_i, \Sigma_{i+1}]$. By Lemma 5.5, only the first circle of $\mathcal{K}_I(i)$ is a candidate for $\mathcal{C}_1$. Hence at most six circles in $\mathcal{K}_I$ are candidates for $\mathcal{C}_1$. For each $0 \leqslant i \leqslant 5$, by performing a binary search on $\mathcal{K}_I$, we can find, in $O(\log n)$ time, the first circle of $\mathcal{K}_I$ whose point of tangency with $\mathcal{C}_I$ lies after Σ_i in the counterclockwise sense. Obviously, this is the first circle of $\mathcal{K}_I(i)$. We can therefore compute in $O(\log n)$ time at most six candidate paths for a fixed $\mathcal{C}_2 \in \mathcal{K}_F$. By examining each $\mathcal{C}_2 \in \mathcal{K}_F$ in turn, we compute $O(n)$ candidate paths in $O(n \log n)$ time. We repeat a similar procedure for all possible orientations of C_I and C_F. We can therefore conclude the following.

LEMMA 5.6. *An optimal path of type (A.ii) can be computed in $O(n \log n)$ time.*

5.4. Computing type (B.iii) paths. Let Π be an optimal path of the form $\Pi_I \bar{C}_2 \bar{C}_3 \Pi_F$, i.e., of type (B.iii). Suppose we know the edges e, e' that are tangent to $\bar{C}_2$ and $\bar{C}_3$, respectively.

If Π does not contain any $\bar{\bar{C}}$-segment in Π_I or Π_F, then Π is of type $C_I S \bar{C}_2 \bar{C}_3 S C_F$. We can compute Π in $O(\log n)$ time using Lemmas 4.1 and 4.2.

Next consider the case in which Π_I and Π_F each contains a $\bar{\bar{C}}$-segment, i.e., Π is of type $C_I \bar{\bar{C}} S \bar{C}_2 \bar{C}_3 S \bar{\bar{C}} C_F$. We show that, given e and e', we can compute, in $O(\log n)$ time, a set of $O(1)$ candidate circles that contains the $\bar{\bar{C}}$-segments of Π. Given this set, we can compute in $O(\log n)$ time the shortest feasible path of the above type, using Lemmas 4.1 and 4.2. Thus by considering all $\Theta(n^2)$ pairs of edges of $\mathcal{P}$, we can compute in $O(n^2 \log n)$ time a set of $O(n^2)$ candidate paths for this case. However, we will see later (in Lemma 5.15) that it suffices to consider fewer pairs of edges of $\mathcal{P}$ in some cases.

5.4.1. Properties of paths. We first establish some simple properties of an optimal path Π of type $C_I \bar{\bar{C}}_1 S \bar{C}_2 \bar{C}_3 S \bar{\bar{C}}_4 C_F$. Assume without loss of generality that $\bar{C}_2$ and $\bar{C}_3$ are oriented clockwise and counterclockwise, respectively. By Lemma 3.3, the $\bar{\bar{C}}_1$-segment is oriented clockwise, and the $\bar{\bar{C}}_4$-segment is oriented counterclockwise, i.e., Π is of type $C_I^+ \bar{\bar{C}}_1^- S \bar{C}_2^- \bar{C}_3^+ S \bar{\bar{C}}_4^+ C_F^-$. Let $\bar{\bar{\mathcal{C}}}_1$, $\bar{\mathcal{C}}_2$, $\bar{\mathcal{C}}_3$, and $\bar{\bar{\mathcal{C}}}_4$ denote the circles supporting the C-segments $\bar{\bar{C}}_1$, $\bar{C}_2$, $\bar{C}_3$, and $\bar{\bar{C}}_4$, respectively.

LEMMA 5.7. *If an optimal path is of type $C_I \bar{\bar{C}}_1 S \bar{C}_2 \bar{C}_3 S \bar{\bar{C}}_4 C_F$, then the circles $\bar{\bar{\mathcal{C}}}_1$, $\bar{\mathcal{C}}_2$, $\bar{\mathcal{C}}_3$, and $\bar{\bar{\mathcal{C}}}_4$ are free.*

Proof. Lemma 3.6 directly yields that $\bar{\bar{\mathcal{C}}}_1$ and $\bar{\bar{\mathcal{C}}}_4$ are free. Suppose now for a contradiction that $\bar{\mathcal{C}}_3$ is not free. As before, we assume that the orientations are such that $\Pi = C_I^+ \bar{\bar{C}}_1^- S \bar{C}_2^- \bar{C}_3^+ S \bar{\bar{C}}_4^+ C_F^-$. Let T be the point of tangency between $\bar{C}_2$ and $\bar{C}_3$. Moving along $\bar{C}_3^+$, let X be the last point of tangency between $\bar{C}_3$ and $\partial\mathcal{P}$ (see Figure 5.7). Starting at X and moving along $\bar{\mathcal{C}}_3^+$, let Y be the first proper intersection point between $\bar{\mathcal{C}}_3$ and $\partial\mathcal{P}$.

By Lemma 2.4, the length of $\bar{C}_3$ between T and X is greater than π, i.e., $\|\bar{C}_3^+[T, X]\| > \pi$. It follows that $\bar{\mathcal{C}}_3$, X and Y define a pocket $\Lambda_{\bar{\mathcal{C}}_3}[X, Y]$ (see Figure 5.7). By Lemma 2.7, this pocket contains $\Pi[X, F]$ and therefore contains the C-segment $\bar{\bar{C}}_4$. But the pocket does not contain the circle $\bar{\bar{\mathcal{C}}}_4$, by Lemma 2.6. The path $\bar{C}_3 S \bar{\bar{C}}_4$ enters the pocket at X, and since $\bar{\bar{\mathcal{C}}}_4$ is free, it is possible to escape the pocket by extending segment $\bar{\bar{C}}_4$. This contradicts Lemma 2.7, establishing that $\bar{\mathcal{C}}_3$ is free. A symmetric argument shows that circle $\bar{\mathcal{C}}_2$ is free. □

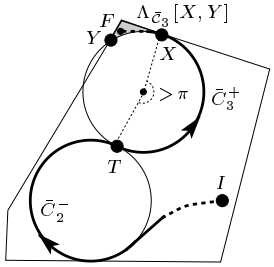


FIG. 5.7. *Illustration of the proof of Lemma 5.7.*

We now introduce the following simple definitions. Given a circle $\mathcal{C}$, a point $M \in \mathcal{C}$ is called *free on* $\mathcal{C}$ if and only if the circle tangent to $\mathcal{C}$ at M is free. Given a circle $\mathcal{C}$ and a point $X \in \mathcal{C}$, a point $M \in \mathcal{C}$ is called the *first free point after* X *along* $\mathcal{C}^+$ if and only if M is free on $\mathcal{C}$ and no point on the arc $\mathcal{C}^+[X, M)$ is free. In

Figure 5.8, M^* is the first free point after M_L along $\mathcal{C}_I^+$. Note that M and X might coincide. The circle tangent to $\mathcal{C}$ at the first free point after X is called the *first free circle after X along $\mathcal{C}^+$*.

We now show that, given I, F, e, and e', we can compute in $O(\log n)$ time a set of $O(1)$ candidate circles that contain the $\bar{\bar{C}}$-segments of any optimal path from I to F of type $C_I^+ \bar{\bar{C}}_1^- S \bar{C}_2^- \bar{C}_3^+ S \bar{\bar{C}}_4^+ C_F^-$ such that $\bar{C}_2^-$ and $\bar{C}_3^+$ are tangent to e and e', respectively. We show how to compute candidate circles for $\bar{\bar{C}}_1$; computing candidate circles for $\bar{\bar{C}}_4$ is similar.

5.4.2. Computing $\bar{\bar{C}}_1$. We identify two circles $\mathcal{C}'$ and $\mathcal{C}''$ that are the candidate circles for $\bar{\bar{C}}_1$. See Figure 5.8. $\mathcal{C}'$ is the first free circle after I along $\mathcal{C}_I^+$. Such a circle $\mathcal{C}'$ exists, by Lemma 5.7, if the optimal path from I to F is of type $C_I^+ \bar{\bar{C}}_1^- S \bar{C}_2^- \bar{C}_3^+ S \bar{\bar{C}}_4^+ C_F^-$. For simplicity, we define $\mathcal{C}''$ in the local frame where the line L through e is horizontal and below $\mathcal{P}$. If the distance between L and $\mathcal{C}_I^+$ is greater than 2, then $\mathcal{C}''$ is not defined. Otherwise, there exist two circles that are above L and tangent to both $\mathcal{C}_I^+$ and L. Let $\mathcal{C}_L$ be the leftmost of these two circles, and let M_L be its point of tangency with $\mathcal{C}_I^+$. Let $\mathcal{C}''$ be the first free circle after M_L along $\mathcal{C}_I^+$. Note that $\mathcal{C}'$ depends only on I and $\mathcal{P}$ and that $\mathcal{C}''$ depends only on I, e, and $\mathcal{P}$.

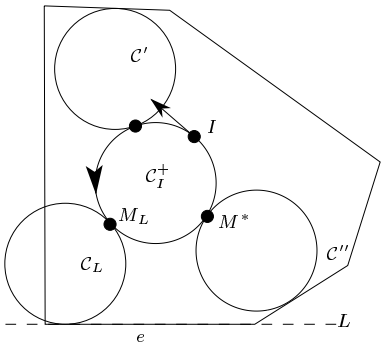


FIG. 5.8. *Definition of $\mathcal{C}'$ and $\mathcal{C}''$.*

LEMMA 5.8. *After preprocessing $\mathcal{P}$ in $O(n \log n)$ time, for a given edge e, $\mathcal{C}'$ and $\mathcal{C}''$ can be computed in $O(\log n)$ time.*

Proof. Let Γ be the circle of radius 2 concentric with $\mathcal{C}_I^+$ (see Figure 5.9). Let $\mathcal{R}$ be the *retracted polygon* of $\mathcal{P}$ with respect to a unit circle, i.e., $\mathcal{R}$ is the set of points p such that the unit circle centered at p lies inside $\mathcal{P}$; $\mathcal{R}$ is a convex polygonal region with at most n edges, and it can be computed in linear time. We preprocess $\mathcal{R}$ in $O(n \log n)$ time using the algorithm by Cheng *et al.* [10], so that given a unit-radius circle $\mathcal{C}$ and a point $M \in \mathcal{C}$, we can compute in $O(\log n)$ time the first intersection point of $\mathcal{C}$ and $\mathcal{R}$ as we walk along $\mathcal{C}$ in the clockwise (or counterclockwise) sense.

Let I_Γ (respectively M_Γ) be the intersection point between Γ and the ray emanating from the center of $\mathcal{C}_I^+$ and going through I (respectively M_L), and let O' be the first intersection point between Γ and $\mathcal{R}$ starting at I_Γ and moving along Γ^+. Note that $O' = I_\Gamma$ if I_Γ lies inside $\mathcal{R}$. The center of $\mathcal{C}'$ is O'. Indeed, by definition of $\mathcal{R}$, the circle centered at O' is free, and any circle (of unit radius) centered at a point on $\Gamma^+[I_\Gamma, O')$ is not free. Since $\mathcal{C}'$ does not depend on the edge e, we can compute it in $O(n)$ time in the preprocessing stage once and for all. The center of $\mathcal{C}''$ is the first intersection point between Γ and $\mathcal{R}$ starting at M_Γ and moving along Γ^+, and it can be computed in $O(\log n)$ time. □

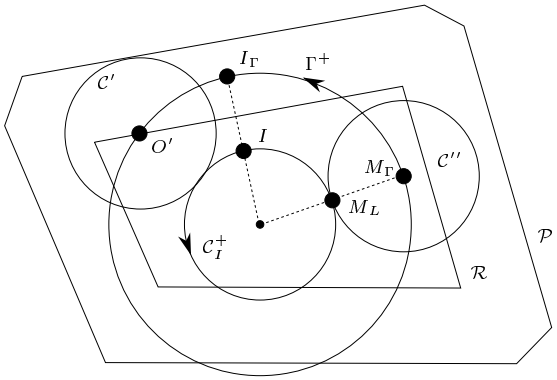


FIG. 5.9. *Illustration of the proof of Lemma 5.8.*

We now state a key lemma, which we prove in the following section.

LEMMA 5.9. *If* Π *is an optimal path of type* $C_I^+ \bar{\bar{C}}_1^- S \bar{C}_2^- \bar{C}_3^+ S \bar{\bar{C}}_4^+ C_F^-$, *then* $\bar{\bar{C}}_1$ *is supported by* $\mathcal{C}'$ *or* $\mathcal{C}''$.

5.4.3. Proof of Lemma 5.9. Let T be the configuration on Π at the common point between $\bar{C}_2$ and $\bar{C}_3$. See Figure 5.10. As before (in Section 5.2), any choice of a point $M \in \mathcal{C}_I^+$ defines one path $\Pi(M)$ of the form $C_I^+ C_1^- S C_2^-$, which begins at I and ends at T, and where $\mathcal{C}_I^+$ and $\mathcal{C}_1^-$ are tangent at M. Since the C-segments C_1^- and C_2^- have the same orientation, the path $\Pi(M)$ always exists, though it might not be free. Let $M^* \in \mathcal{C}_I^+$ be the point such that $\Pi(M^*)$ is a subpath of the optimal path Π. It follows that $\Pi(M^*)$ is an optimal path from I to T. We will show that M^* is the first free point after I or M_L.

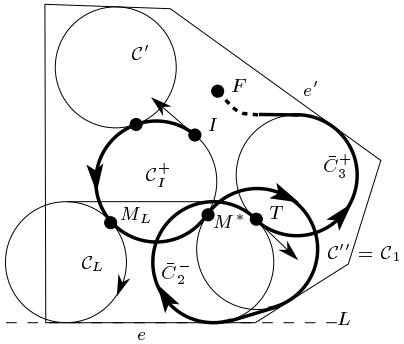


FIG. 5.10. *Circles* $\mathcal{C}'$ *and* $\mathcal{C}''$ *with a (supposedly) optimal path* Π.

We consider different cases based on which of the singular points exist on $\mathcal{C}_I^+$. See Figure 5.11. We first introduce some notation in order to distinguish different singular points. Let $M_1 \in \mathcal{C}_I$ be the point such that the C-segment C_1^- in $\Pi(M)$ vanishes (i.e., $\Pi(M)$ is of type $C_I^+ S C_2^-$); M_1 is only defined when $\mathcal{C}_I$ and $\mathcal{C}_2$ do not properly intersect. Assume for simplicity that T is the lowest point of $\mathcal{C}_2^-$, and let L_0 be the horizontal half-line lying to the right of T. Let $\tilde{\mathcal{C}}_1$ and $\tilde{\mathcal{C}}_1'$ be the two circles (if they exist) tangent to $\mathcal{C}_I$ with center on the horizontal line through the center of $\mathcal{C}_2^-$, and let M_2 and M_2' be their respective common points with $\mathcal{C}_I$; assume without loss of generality that M_2 is left of M_2'. The point M_2 (respectively M_2') is a singular point of $\Pi(M)$ (at which C_2^- vanishes) if and only if $\tilde{\mathcal{C}}_1$ (respectively $\tilde{\mathcal{C}}_1'$)

touches L_0. Otherwise, the C_2-segment of $\Pi(M_2)$ (respectively $\Pi(M_2')$) has length π (see Figure 5.11b).

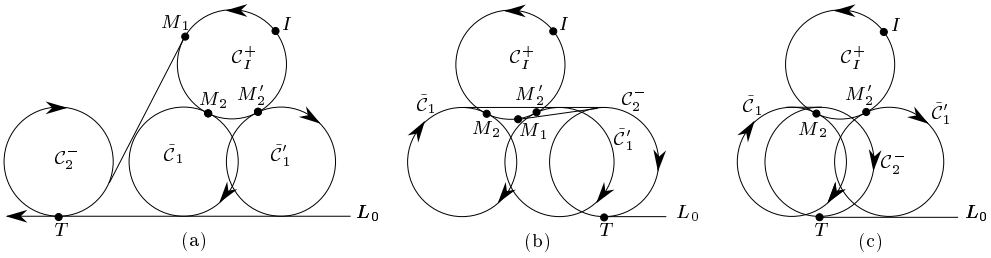


FIG. 5.11. *Singular points of* $\Pi(M)$. *In (a):* $\{I, M_1, M_2, M_2'\}$. *In (b):* $\{I, M_1\}$. *In (c):* $\{I, M_2'\}$.

Since C_1- and C_2-segments of $\Pi(M)$ have the same orientation, the S-segment vanishes if and only if the path type $C_I^+ C_1^- S C_2^-$ degenerates to $C_I^+ C^-$. Thus if the S-segment vanishes, the C_1- or C_2-segment also vanishes. Therefore, in view of the discussion in Section 5.2, only the following points can be singular points:

- I (C_I^+ vanishes),
- M_1, if $\mathcal{C}_I$ and $\mathcal{C}_2$ do not properly intersect (C_1^- vanishes),
- M_2, if $\tilde{\mathcal{C}}_1$ exists and touches L_0 (C_2^- vanishes), and
- M_2', if $\tilde{\mathcal{C}}_1'$ exists and touches L_0 (C_2^- vanishes).

There are three cases to consider, depending on the relative positions of $\mathcal{C}_1$ and $\mathcal{C}_2$.

(i) The distance between $\mathcal{C}_I$ and the line supporting L_0 is at most 2, and $\mathcal{C}_2$ lies to the left of $\tilde{\mathcal{C}}_1$, i.e., both $\tilde{\mathcal{C}}_1$ and $\tilde{\mathcal{C}}_1'$ touch L_0; see Figure 5.11a. In this case $\mathcal{C}_2$ does not intersect $\mathcal{C}_I$, and therefore M_1 also exists. The singular points are thus $\{I, M_1, M_2, M_2'\}$.

(ii) Either the distance between $\mathcal{C}_I$ and the line supporting L_0 is greater than 2, or neither $\tilde{\mathcal{C}}_1$ nor $\tilde{\mathcal{C}}_1'$ touches L_0; see Figure 5.11b. In both cases, $\mathcal{C}_I$ and $\mathcal{C}_2$ do not intersect, so M_1 exists. The singular points are therefore $\{I, M_1\}$.

(iii) The distance between $\mathcal{C}_I$ and the line supporting L_0 is at most 2 and $\mathcal{C}_2$ lies between $\tilde{\mathcal{C}}_1$ and $\tilde{\mathcal{C}}_1'$. In this case, $\tilde{\mathcal{C}}_1$ does not touch L_0, and $\mathcal{C}_I$ intersects $\mathcal{C}_2$, so the singular points are $\{I, M_2'\}$; see Figure 5.11c.

Before proving for each case that M^* is the first free point along $\mathcal{C}_I^+$ after I or M_L, we state a few claims, which we will need for the proof.

CLAIM 5.10. *M^* is not a singular point of* $\Pi(M)$.

Proof. If M^* is a singular point, the type of $\Pi(M^*)$ degenerates, contradicting that Π is of type $C_I^+ \bar{\bar{C}}_1^- S \bar{C}_2^- \bar{C}_3^+ S \bar{\bar{C}}_4^+ C_F^-$. □

CLAIM 5.11. *M^* is the first free point along $\mathcal{C}_I^+$ after a singular point of* $\Pi(M)$.

Proof. By Lemma 5.7, M^* is free on $\mathcal{C}_I^+$. If there exists $M' \neq M^*$ free on $\mathcal{C}_I^+$ such that $\mathcal{C}_I^+(M', M^*]$ does not contain any singular point, then the path $\Pi(M')$ exists because C_1^- and C_2^- have the same orientation. $\Pi(M')$ is free because the first C_I^+-segment of $\Pi(M')$ is part of the feasible path $\Pi(M^*)$, the circle $\mathcal{C}_1^-$ is free by definition of M', and the circle $\mathcal{C}_2^-$ is free (by Lemma 5.7). Finally, $\|\Pi(M')\| < \|\Pi(M^*)\|$ by the monotonicity property (Lemma 5.5), which contradicts the optimality of $\Pi(M^*)$. □

CLAIM 5.12. *If M_2 and M_2' are defined (but not necessarily singular points), then (i) the length function $\|\Pi(M)\|$ increases at $M = M_2$, and (ii) $M^* \in \mathcal{C}_I^+(M_2, M_2')$.*

Proof. If M_2 is not a singular point, then by Lemma 5.5, $\|\Pi(M)\|$ increases at M_2. If M_2 is a singular point, then $\|\Pi(M)\|$ jumps by 2π at M_2 (see Figure 5.11a).

As for (ii), the length of the last C-segment C_2 of $\Pi(M)$ is greater that π if and only if the center of $\mathcal{C}_1$ lies below the center of $\mathcal{C}_2$ (see Figure 5.11), that is $M \in \mathcal{C}_I^+(M_2, M_2')$. Since C_2 is a nonterminal C-segment of the optimal path Π, $\|C_2\| > \pi$, and therefore $M^* \in \mathcal{C}_I^+(M_2, M_2')$. □

CLAIM 5.13. *If M_1 exists, then the circular arc $\mathcal{C}_I^+[M_1, M^*]$ contains I or M_2'. If $I \notin \mathcal{C}_I^+[M_1, M^*]$, then M_2' is a singular point.*

Proof. If $I \notin \mathcal{C}_I^+[M_1, M^*]$, then $\Pi(M_1)$ is free because the first C-segment $\mathcal{C}_I^+[I, M_1]$ of $\Pi(M_1)$ is part of $\Pi(M^*)$ and $\mathcal{C}_2^-$ is free by Lemma 5.7. Thus $\Pi(M_1, M^*]$ contains a singular point, because otherwise $\|\Pi(M_1)\| < \|\Pi(M^*)\|$ by the monotonicity property (Lemma 5.5) and Claim 5.10, and thus $\Pi(M^*)$ is not optimal, a contradiction. If M_2' also does not lie in $\mathcal{C}_I^+[M_1, M^*]$, then M_2 is a singular point and lies on this arc. By Claim 5.12, $\|\Pi(M_1)\| < \|\Pi(M_2)\| < \|\Pi(M^*)\|$, a contradiction. Hence either I or M_2' lies on $\mathcal{C}_I^+[M_1, M^*]$, and M_2' is a singular point if I does not lie in this arc. □

We now prove for each of the three cases stated above that M^* is the first free point after I or M_L.

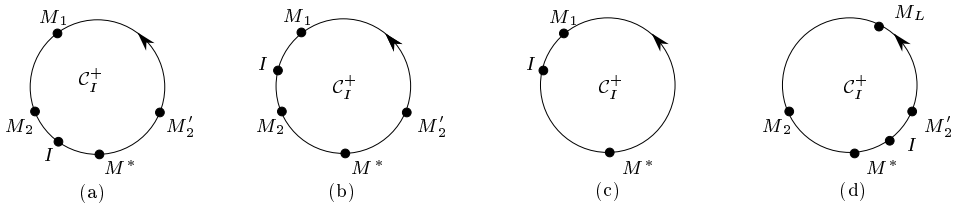


FIG. 5.12. *Some positions of the singular points $\{I, M_1, M_2, M_2'\}$ in (a) and (b), $\{I, M_1\}$ in (c), and $\{I, M_2'\}$ in (d).*

Case (i) The singular points are $\{I, M_1, M_2, M_2'\}$. Since $\mathcal{C}_2$ lies to the left of $\tilde{\mathcal{C}}_1$, one can easily show that $M_1 \in \mathcal{C}_I^+[M_2', M_2]$ (see Figure 5.11a). Refer now to Figures 5.12a and b. By Claim 5.12, $M^* \in \mathcal{C}_I^+(M_2, M_2')$. It follows that $\mathcal{C}_I^+[M_1, M^*]$ does not contain M_2' and thus contains I (by Claim 5.13). Thus $\mathcal{C}_I^+(I, M^*]$ does not contain any singular point except possibly M_2. If $M_2 \notin \mathcal{C}_I^+(I, M^*]$ (Figure 5.12a), then M^* is the first free point after I because, by Claim 5.11, M^* is the first free point after a singular point. Even if $M_2 \in \mathcal{C}_I^+(I, M^*]$ (Figure 5.12b), M^* is the first free point after I. Indeed, if the first free point after I along $\mathcal{C}_I^+$ is $M' \in \mathcal{C}_I^+[I, M_2)$, then $\Pi(M')$ is free because the first arc $\mathcal{C}_I^+[I, M']$ of $\Pi(M')$ is part of Π and the second and third C-segments of $\Pi(M')$ are free by definition of M' and by Lemma 5.7, respectively. Moreover, by Lemma 5.5 and Claim 5.12, $\|\Pi(M')\| < \|\Pi(M_2)\| < \|\Pi(M^*)\|$, a contradiction.

Case (ii) The singular points are $\{I, M_1\}$. See Figure 5.12c. Since M_2' is not a singular point, by Claim 5.13, $I \in \mathcal{C}_I^+[M_1, M^*]$. Consequently, $\mathcal{C}_I^+(I, M^*]$ does not contain any singular point. Therefore, by Claim 5.11, M^* is the first free point after I.

Case (iii) The singular points are $\{I, M_2'\}$. As before, if $\mathcal{C}_I^+(I, M^*]$ does not contain any singular point, M^* is the first free point after I. We thus consider the case in which $M_2' \in \mathcal{C}_I^+(I, M^*]$ (see Figure 5.12d). Since M_2 and M_2' exist, by Claim 5.12, $M^* \in \mathcal{C}_I^+(M_2, M_2')$. It thus follows that M^* is the first free point after M_2' (by Claims 5.11 and 5.12). Thus in order to show that M^* is the first free point after M_L, it is sufficient to prove that $M_L \in \mathcal{C}_I^+[M_2', M_2]$,

which is equivalent to proving that the length of the last C-segment C_2 of $\Pi(M_L)$ is at most π. This can be shown as follows.
We assume for simplicity that the edge e (tangent to the C-segment C_2^-) is horizontal and below $\mathcal{P}$ (see Figure 5.10); to be consistent, we no longer assume that T is the lowest point of $\mathcal{C}_2^-$. By Lemma 2.4, the arc length of $\bar{C}_2$ in Π from its point of tangency with the edge e to T must be at least π. In other words, T must be in the right half of $\bar{\mathcal{C}}_2$. On the other hand, by definition of M_L, $\mathcal{C}_L$ is the leftmost circle of all the unit circles tangent to L from above that intersect $\mathcal{C}_I$. Since $\mathcal{C}_2$ is tangent to L from above and properly intersects $\mathcal{C}_I$ (because M_1 is not defined in this case), the top point of $\mathcal{C}_L$ is left of the top point of $\mathcal{C}_2$. Thus, since T is on the right half of $\bar{\mathcal{C}}_2$, the arc length of C_2 in $\Pi(M_L)$ is less than π (see Figure 5.10).

5.4.4. Computing the overall path. By Lemmas 5.8 and 5.9, we can compute, in $O(\log n)$ time, two candidates for the circle supporting segment $\bar{\bar{C}}_1$. We can similarly compute two candidates for the circle supporting segment $\bar{\bar{C}}_4$. By Lemma 4.2, this gives us $O(1)$ candidate paths, for which we may check the feasibility in $O(\log n)$ time. Hence, given two edges e and e' of $\mathcal{P}$, we can compute in $O(\log n)$ time an optimal path of type $C_I\bar{\bar{C}}_1S\bar{C}_2\bar{C}_3S\bar{\bar{C}}_4C_F$, where $\bar{C}_2$ and $\bar{C}_3$ are tangent to e and e', respectively.

If the optimal path is of type (B.iii) with only one $\bar{\bar{C}}$-segment in Π_I or Π_F, we get similar results. For example, if an optimal path is of type $C_I\bar{\bar{C}}_1S\bar{C}_2\bar{C}_3SC_F$, then $\bar{\bar{\mathcal{C}}}_1$ and $\bar{\mathcal{C}}_2$ are free, and $\bar{\bar{C}}_1$ is supported by $\mathcal{C}'$ or $\mathcal{C}''$ as defined above. Thus we obtain the following lemma.

LEMMA 5.14. *Let e, e' be edges of $\mathcal{P}$. In $O(\log n)$ time, we can compute an optimal path of type $\Pi_I\bar{C}_2\bar{C}_3\Pi_F$, where $\Pi_I \in \{C_I\bar{\bar{C}}S, C_IS, C_I, S\}$, $\Pi_F \in \{S\bar{\bar{C}}C_F, SC_F, C_F, S\}$ and where $\bar{C}_2$ and $\bar{C}_3$ are tangent to e and e', respectively.*

5.4.5. Finding candidate pairs of edges. Now we describe how to find a suitable set of pairs of edges $\mathcal{E}$ such that if an optimal path from I to F is of type (B.iii) (i.e., $\Pi_I\bar{C}_2\bar{C}_3\Pi_F$), then the pair of edges (e, e') tangent to $\bar{C}_2$ and $\bar{C}_3$ is in the set $\mathcal{E}$.

Agarwal *et al.* [1] showed that if an optimal path from I to F is of type $\Pi_I\bar{C}_2\bar{C}_3\Pi_F$ such that $\bar{C}_2$ and $\bar{C}_3$ are nonterminal, then $\mathcal{C}_I$ intersects $\bar{\mathcal{C}}_3$ (the circle supporting $\bar{C}_3$), and $\mathcal{C}_F$ intersects $\bar{\mathcal{C}}_2$ (the circle supporting $\bar{C}_2$). Thus the center of $\bar{\mathcal{C}}_3$, which is at most distance 1 from the boundary of the polygon, is at most at distance 3 from I. Since the centers of $\bar{\mathcal{C}}_2$ and $\bar{\mathcal{C}}_3$ are distance 2 apart, they are each at distance less than 5 from I. Thus edges e and e' are at distance less than 6 from I. By symmetry, they are also at distance less than 6 from F. Therefore, we can consider $\mathcal{E}$ to be the set of pairs of edges of $\mathcal{P}$ that are at distance less than 6 from both I and F. Let k denote the number of edges of $\mathcal{P}$ at distance less than 6 from both I and F. Then $|\mathcal{E}| = k^2$, and $\mathcal{E}$ can be computed in $O(k^2)$ time. Lemma 5.14 then gives:

LEMMA 5.15. *An optimal path of type (B.iii) can be computed in $O(k^2 \log n)$ time.*

Putting everything together, we obtain the following.

THEOREM 5.16. *Given a convex polygon $\mathcal{P}$, an initial configuration I, and a final configuration F, an optimal path from I to F inside $\mathcal{P}$ can be computed in time $O((n + k^2)\log n)$, where k is the number of edges of $\mathcal{P}$ at distance less than 6 from both I and F.*

Proof. We have shown in Section 4 and in Lemmas 5.2, 5.6 and 5.15 that the Dubins paths and the optimal paths of type (A.i), (A.ii), (B.i), and (B.ii) can be com-

puted in $O(n \log n)$ time, while paths of type (B.iii) can be computed in $O(k^2 \log n)$ time. Choosing the shortest among all those paths yields an optimal path. □

6. Conclusion. For an obstacle-free environment, Dubins' classification of optimal path types yields a constant time algorithm for computing optimal paths [15]. On the other hand, in the presence of general polygonal obstacles, the optimal path planning problem is NP-hard [29]. In this paper, we have studied a very simple environment, the inside of a convex polygon. We have given a classification of optimal path types and an $O(n^2 \log n)$ algorithm for optimal path planning. We have found that, surprisingly, the number of straight or circular segments composing optimal paths is bounded by a constant, independent of the number n of sides of the convex polygon.

The running time of the algorithm and the constant bound on the number of segments in an optimal path lead us to speculate that other simple environments may also be amenable to polynomial time algorithmic solution. However, we caution that although the environment we have considered is simple, our algorithm results from considerable technical analysis.

Our techniques and results may provide useful tools for further study of these problems. In particular, we call attention to two properties of moderate paths that we believe are interesting and possibly useful in their own right:

(i) A feasible path entering the interior of a pocket can never escape the pocket (Lemma 2.7).

(ii) The length of a path of type $CCSC$ from X to Y is a piecewise strictly increasing function of the length of the first C-segment (Lemma 5.5).

Note that Property (ii) holds regardless of the environment.

The theory of curvature-constrained path planning is relatively unexplored, so many problems remain. We mention some specific open problems arising from our work, and then conclude with a more general one.

First, we ask whether our classification of optimal path types inside a convex polygon is tight; that is, does each possible type of optimal path given in our classification actually arise for some $\mathcal{P}$, I, and F? We believe that the answer is yes although we have no formal proof of this (see Remark 3.2 for details). Also, we ask whether there is a polygon $\mathcal{P}$ in which all types arise.

Next, in our optimal path planning algorithm, the most time consuming part lies in the computation of the optimal paths of type (B.iii) (see Theorem 3.1). Indeed, if we eliminate type (B.iii) from consideration, the complexity of our algorithm drops to $O(n \log n)$. Furthermore, paths of type (B.iii) are rather complex and thus may be difficult to track by a mobile robot. This situation suggests two lines of investigation. First, the paths of type (B.iii) should be studied thoroughly in order to understand when they can be optimal inside a convex polygon (or amid moderate obstacles [1, 6]). We believe that optimal paths can be of type (B.iii) only if the polygon is "small". In other words, it is possible that optimal paths of type (B.iii) only arise as artifacts of tightly constricted environments (see Figure 3.1c). For example, we know (see Theorem 5.16) that if the terminal configurations are distance greater than 6 from the boundary of the polygon, then an optimal path cannot be of type (B.iii). A second approach to handling paths of type (B.iii) is to simply drop them from consideration. In that case, we ask whether we can preprocess the scene such that, for any query of terminal configurations, we can compute a shortest path among the remaining types in sub-linear time.

To conclude with a general problem for future research, we ask for the specification

of a realistic notion of feasibility that rejects hard-to-follow paths, such as paths of type (B.iii), while admitting fast computation of optimal feasible paths.

Acknowledgments. This paper arose from research begun at the International Workshop on Bounded Radius of Curvature, organized by Sue Whitesides and held at the Bellairs Research Institute of McGill University, March 9-16, 1997. We would like to thank Hazel Everett, Micha Godau, Sylvain Petitjean, Kaleem Siddiqi, and Steve Wismath for many useful discussions.

REFERENCES

[1] P. K. Agarwal, P. Raghavan, and H. Tamaki. Motion planning for a steering-constrained robot through moderate obstacles. In *Proc. of the 27th Annu. ACM Sympos. Theory Comput.*, 1995, pp. 343–352.

[2] H.-K. Ahn, O. Cheong, J. Matoušek, and A. Vigneron. Reachability by paths of bounded curvature in convex polygons. In *Proc. of the 32th Annu. ACM Sympos. Theory Comput.*, 2000, pp. 251–259.

[3] J. Barraquand and J.-C. Latombe. Nonholonomic multi-body mobile robots: Controllability and motion planning in the presence of obstacles. *Algorithmica*, 10 (1993), 121–155.

[4] J-D. Boissonnat and X-N. Bui. Accessibility region for a car that only moves forwards along optimal paths. Rapport de recherche 2181, INRIA, 1994.

[5] J.-D. Boissonnat, A. Cérézo, and J. Leblond. Shortest paths of bounded curvature in the plane. *Internat. J. Intell. Syst.*, 10 (1994), 1–16.

[6] J.-D. Boissonnat and S. Lazard. A polynomial-time algorithm for computing a shortest path of bounded curvature amidst moderate obstacles. In *Proc. of the 12th Annu. ACM Sympos. Comput. Geom.*, 1996, pp. 242–251.

[7] J.-D. Boissonnat and S. Lazard. A polynomial-time algorithm for computing a shortest path of bounded curvature amidst moderate obstacles. Rapport de recherche 2887, INRIA, 1996. Accepted in *Internat. J. Comput. Geom. Appl.*.

[8] X.-N. Bui, P. Souères, J.-D. Boissonnat , and J.-P. Laumond. The shortest path synthesis for non-holonomic robots moving forwards. In *Proc. of the IEEE Internat. Conf. Robot. Autom.*, 1994, pp. 2–7.

[9] L. Calabi and J.A. Riley. The skeletons of stable plane sets. Scientific Report SR3-5711, Parke Math. Labs., 1967.

[10] S.-W. Cheng, O. Cheong, H. Everett, and R. van Oostrum. Hierarchical vertical decomposition, ray shooting, and circular arc queries in simple polygons. In *Proc. of the 15th Annu. ACM Sympos. Comput. Geom.*, 1999, pp. 227–236.

[11] J. Canny, B. R. Donald, J. Reif, and P. Xavier. Kinodynamic motion planning. *J. ACM*, 40 (1993), 1048–1066.

[12] J. Canny, A. Rege, and J. Reif. An exact algorithm for kinodynamic planning in the plane. *Discrete Comput. Geom.*, 6 (1991), 461–484.

[13] G. Desaulniers. On shortest paths for a car-like robot maneuvering around obstacles. *Robotics and Autonomous Systems*, 17 (1996), 139–148.

[14] B. R. Donald and P. Xavier. A provably good approximation algorithm for optimal-time trajectory planning. In *Proc. of the IEEE Internat. Conf. Robot. Autom.*, 1989, pp. 958–963.

[15] L. E. Dubins. On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal positions and tangents. *Amer. J. Math.*, 79 (1957), 497–516.

[16] S. Fortune and G. Wilfong. Planning constrained motion. *Ann. Math. Artif. Intell.*, 3 (1991), 21–82.

[17] D. Halperin, L. Kavraki, and J.-C. Latombe. Robotics. In *CRC Handbook of Discrete and Computational Geometry* (J. Goodmand and J. O'Rourke, eds.), CRC Press, Boca Raton, NY, 1997, pp. 755–778.

[18] P. Jacobs and J. Canny. Planning smooth paths for mobile robots. In *Nonholonomic Motion Planning* (Z. Li and J. Canny, eds.), Kluwer Academic Publishers, Norwell, MA, 1992, pp. 271–342.

[19] P. Jacobs, J.-P. Laumond, and M. Taix. Efficient motion planners for nonholonomic mobile robots. *In Proc. of the IEEE/RSJ Internat. Workshop on Intell. Robots and Systems*, 1991, pp. 1229–1235.

[20] J.-C. Latombe. A fast path-planner for a car-like indoor mobile robot. *In Proc. of the 9th National Conf. on Artif. Intell.*, 1991, pp. 659–665.
[21] J.-C. Latombe. *Robot Motion Planning.* Kluwer Academic Publishers, Norwell, MA, 1991.
[22] J.-P. Laumond. Finding collision-free smooth trajectories for a non-holonomic mobile robot. *In Proc. of the Internat. Joint Conf. on Artif. Intell.*, 1987, pp. 1120–1123.
[23] J.-P. Laumond, P. Jacobs, M. Taix, and R. Murray. A motion planner for nonholonomic mobile robots. *IEEE Trans. on Robot. and Autom.*, 10 (1994), 577–593.
[24] J.-P. Laumond, M. Taix, and P. Jacobs. A motion planner for car-like robots based on global/local approach. In *Proc. of the IEEE/RSJ Internat. Workshop on Intell. Robots and Systems*, 1990, pp. 765–773.
[25] B. Mirtich and J. Canny. Using skeletons for nonholonomic path planning among obstacles. In *Proc. of the 9th IEEE Internat. Conf. Robot. Autom.*, 1992, pp. 2533–2540.
[26] C. Ó'Dúnlaing. Motion-planning with inertial constraints. *Algorithmica*, 2 (1987), 431–475.
[27] G. Pestov and V. Ionin. On the largest possible circle imbedded in a given closed curve. *Dok. Akad. Nauk SSSR*, 127 (1959), 1170–1172.
[28] J. A. Reeds and L. A. Shepp. Optimal paths for a car that goes both forwards and backwards. *Pacific J. Math.*, 145 (1990), 367–393.
[29] J. Reif and H. Wang. The complexity of the two dimensional curvature-constrained shortest-path problem. In *Proc. of the 3rd Workshop on the Algo. Found. of Robotics* (P. K. Agarwal, L. E. Kavraki, and M. Mason, eds.). A. K. Peters, Natick, MA, 1998, pp. 49–58.
[30] W. Rudin. *Real and complex analysis.* McGraw-Hill, 1966.
[31] J. T. Schwartz and M. Sharir. Algorithmic motion planning in robotics. In *Algorithms and Complexity*, volume A of *Handbook of Theoretical Computer Science* (J. van Leeuwen, ed.), Elsevier, Amsterdam, 1990, pp. 391–430.
[32] S. Sekhavat, P. Švestka, J.-P. Laumond, and M. Overmars. Multi-level path planning for non-holonomic robots using semi-holonomic subsystems. In *Workshop on the Algo. Found. of Robotics* (J.-P. Laumond and M. Overmars, eds.), A. K. Peters, Wellesley, MA, 1996, pp. 79–96.
[33] J. Sellen. Planning paths of minimal curvature. *In Proc. of the IEEE Internat. Conf. Robot. and Autom.*, 1995.
[34] J. Sellen. Approximation and decision algorithms for curvature constrained path planning: A state-space approach. In *Proc. of the 3rd Workshop on the Algorithmic Foundations of Robotics* (P. K. Agarwal, L. E. Kavraki, and M. Mason, eds.). A. K. Peters, Natick, MA, 1998, pp. 59–68.
[35] J. Serra. *Image analysis and mathematical morphology, 2: theoretical advances.* Academic press, New York, NY, 1988.
[36] P. Souères and J.-P. Laumond. Shortest path synthesis for a car-like robot. *IEEE Trans. on Automatic Control*, 41(5) (1996), 672–688.
In Proc. 34th Conf. on Decision and Control, pp. 3306– 3312, 1995.
[37] H. J. Sussmann. Shortest 3-dimensional paths with a prescribed curvature bound. In *Proc. of the 34th Conf. on Decision & Control*, 1995, pp. 3306–3311.
[38] H. J. Sussmann and G. Tang. Shortest paths for the Reeds-Shepp car: a worked out example of the use of geometric techniques in nonlinear optimal control. Report SYCON-91-10, Rutgers University, New Brunswick, NJ, 1991.
[39] P. Svestka and M. H. Overmars. Motion planning for car-like robots using a probabilistic learning approach. *J. Robot. Res.*, 16(2) (1997), 119–143.
[40] M. Vendittelli, J.-P. Laumond, and C. Nissoux. Obstacle distance for car-like robots. *IEEE Trans. on Robot. and Autom.*, 15(4) (1999), 678–691.
[41] H. Wang and P. K. Agarwal. Approximation algorithms for curvature constrained shortest paths. In *Proc. of the 7th ACM-SIAM Sympos. Discrete Algorithms*, 1996, pp. 409–418.
[42] G. Wilfong. Motion planning for an autonomous vehicle. In *Proc. of the IEEE Internat. Conf. Robot. Autom.*, 1988, pp. 529–533.
[43] G. Wilfong. Shortest paths for autonomous vehicles. In *Proc. of the IEEE Internat. Conf. Robot. and Autom.*, 1989, pp. 15–20.

Appendix A. Closed Moderate Paths.

The purpose of this section is to prove Proposition 2.5, which states:

> *A simple moderate path* Π *such that the initial and final locations coincide (the initial and final configurations may differ) bounds a region that contains at least one disk of unit radius. Moreover, if the initial and final configurations coincide and if* Π *is not a circle of unit radius, then the region bounded by* Π *contains at least two distinct disks of unit radius.*

Proof. We prove these results using some properties of the skeletons. Note that this is an original approach in the field of nonholonomic motion planning. We first recall a definition of skeletons for which several variants are considered in the literature; different terms in use include *medial-axis, central set*, and *cut-locus*. We use here the formulation using maximal disks [35].

Skeletons. Let $\mathcal{R}$ be an open set of $\Re^2$ bounded by a simple closed curve. A *maximal disk* is an open disk (of positive radius) included in $\mathcal{R}$, but not included in any other disk contained in $\mathcal{R}$. The *skeleton* of $\mathcal{R}$, denoted $\mathcal{S}(\mathcal{R})$, is the locus of the centers of all the maximal disks. For any $x \in \mathcal{R}$, let $\rho(x)$ denote the intersection between the closure of the maximal disk centered at x and the boundary of $\mathcal{R}$:

$$\rho(x) = \{y \in \partial\mathcal{R} \mid \|xy\| = \min_{z \in \partial\mathcal{R}} \|xz\|\}.$$

Now let $I = F$ denote the initial and final location on the path Π, and let $\mathcal{R}$ be the open region bounded by Π. Assume that $\mathcal{R}$ is not a disk of radius at least 1, as otherwise, the result is obvious.

The underlying idea of the proof is the following. The closure of skeleton $\mathcal{S}(\mathcal{R})$ has at least two distinct endpoints (i.e., nodes of degree 1 of the graph) x and x'; indeed, $\mathcal{S}(\mathcal{R})$ does not contain any cycle since Π is simple. One of these endpoints, say x, is necessarily distinct from the terminal location I of Π. The point x cannot lie on Π because Π is C^1 continuous everywhere except (possibly) at $I \neq x$, therefore $x \in \mathcal{R}$. Since x is an endpoint of the skeleton, $\rho(x)$ is connected (i.e., $\rho(x)$ is a point or a circular arc). It follows that the maximal disk at x is osculating[5] Π at a point $P \neq I$, and its radius is at least 1. If the initial and final orientations of Π are also equal, x and x' are both centers of maximal disks osculating Π and thus their radii are greater or equal to 1.

Formally, we directly show, using a result by Calabi and Riley [9], the following result.

Claim A.1. *There exists a maximal disk in* $\mathcal{R}$ *such that the contact points between its boundary and* Π *are connected.*

Proof. Let D be any maximal disk in $\mathcal{R}$ and x be its center, and suppose for a contradiction that $\rho(x)$ is not connected. See Figure A.1. Then there exist four points $y_1 \neq y_2, y_1' \neq y_2'$ in $\partial D \cap \Pi$ such that the circular arcs $\partial D^+[y_1, y_2]$ and $\partial D^+[y_1', y_2']$ do not strictly contain any point of $\rho(x)$.

[5] Two curves are osculating at a point P if and only if they are tangent at P and have the same signed curvature at P. Here Π is C^1 continuous, and piecewise C^2 continuous, everywhere except possibly at I. If Π is C^1 but not C^2 at P, we say that a disk is osculating Π at P if and only if they are tangent at P, the signed curvature of the disk is equal to the signed curvature (at P) of one of the two portions (C^2) of Π ending at P, and the other portion does not properly intersect the disk in a neighborhood of P.

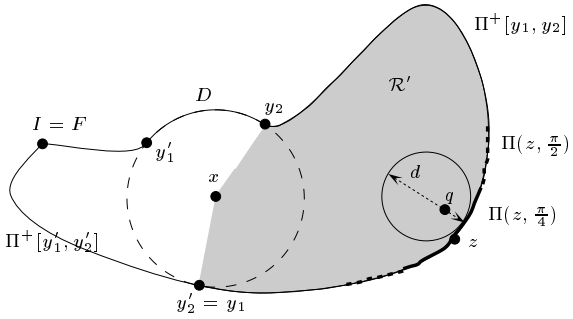


FIG. A.1. *For the proof of Proposition 2.5.*

One of the two arcs $\Pi^+[y_1, y_2]$ and $\Pi^+[y_1', y_2']$ is C^1 continuous because these two arcs do not overlap and Π is C^1 continuous everywhere except possibly at $I = F$. Say, without loss of generality, that $\Pi^+[y_1, y_2]$ is C^1 continuous, and denote by $\mathcal{R}'$ the open region bounded by $\Pi^+[y_1, y_2] \cup [x, y_1] \cup [x, y_2]$ (e.g., the shaded region in Figure A.1). By a result of Calabi and Riley [9, Proposition 10], either $\mathcal{R}'$ contains a skeleton point $x_0 \in \mathcal{SR}$ for which $\rho(x_0)$ is connected, or $\Pi^+(y_1, y_2)$ (without its endpoints) contains a point which is the limit of skeleton points contained in $\mathcal{R}'$.

Suppose on the contrary that $\Pi^+(y_1, y_2)$ contains a point z that is the limit of skeleton points $(z_i)_{i\in\mathbb{N}}$ contained in $\mathcal{R}$. Since $\Pi^+(y_1, y_2)$ is C^1, Π is C^1 at z. We get a contradiction because, as we prove below, a point z of $\partial\mathcal{R}$ is a limit point of points of the skeleton only if $\partial\mathcal{R}$ is not differentiable at z. Indeed, let $\Pi(z, \frac{\pi}{2})$ be the arc of Π of length $\frac{\pi}{2}$ centered at z (if the length of Π is smaller than $\frac{\pi}{2}$, we choose a smaller value for the length of $\Pi(z, \frac{\pi}{2})$), and let $\Pi(z, \frac{\pi}{4})$ be the arc of Π, centered at z, of length half the length of $\Pi(z, \frac{\pi}{2})$, and finally, let d be the distance between $\Pi(z, \frac{\pi}{4})$ and $\Pi \setminus \Pi(z, \frac{\pi}{2})$ (note that $d \leqslant \frac{\pi}{4} < 1$). By a result of Dubins [15, Proposition 6], $\Pi(z, \frac{\pi}{2})$ does not intersect any open disk of unit radius tangent to $\Pi(z, \frac{\pi}{4})$. It follows that Π does not intersect any open disk of radius $d/2$ tangent to $\Pi(z, \frac{\pi}{4})$. Now any point $q \in \mathcal{R}$ close enough to z belongs to a normal to $\Pi(z, \frac{\pi}{4})$ at distance $\mu < d/2$ from $\Pi(z, \frac{\pi}{4})$. Thus the disk of radius μ centered at q is not maximal because it is included in a disk of radius $d/2$ tangent to $\Pi(z, \frac{\pi}{4})$, which is included in $\mathcal{R}$. Therefore, any point $q \in \mathcal{R}$ in a small enough neighborhood of z does not belong to the skeleton. □

By Claim A.1, $\mathcal{R}$ contains a point x_0 of its skeleton such that $\rho(x_0)$ is connected. Thus $\rho(x_0)$ is a circular arc or is reduced to a point. If $\rho(x_0)$ is a circular arc, then the constraint on the average curvature of Π implies that this circular arc of Π has a radius greater than or equal to 1, and thus the radius of the maximal disk at x_0 is greater than or equal to 1. Otherwise, if $\rho(x_0)$ is reduced to a point, say y_0, the maximal disk at x_0 osculates Π at y_0, and thus its radius is greater than or equal to 1 (indeed, recall that the constraint on the average curvature implies that Π is piecewise C^2 continuous and its curvature is less than or equal to 1 almost everywhere). This concludes the proof when the initial and final orientations of Π differ.

If the initial and final configurations are equal, the path Π is C^1 continuous everywhere and the previous arguments hold for both regions $\mathcal{R}'$ and $\mathcal{R}''$ bounded, respectively, by $\Pi^+[y_1, y_2] \cup [x, y_1] \cup [x, y_2]$ and $\Pi^+[y_1', y_2'] \cup [x, y_1'] \cup [x, y_2']$ (note however that when $\rho(x)$ is connected, $\mathcal{R}''$ is empty but then the maximal disk centered at x is

of radius at least 1). □