Proceedings of the 34th
Conference on Decision & Control
New Orleans, LA - December 1995

Shortest 3-dimensional paths with a prescribed curvature bound

Héctor J. Sussmann*
Department of Mathematics
Rutgers University
New Brunswick, NJ 08903
sussmann@hamilton.rutgers.edu

Abstract

We present the solution of the three-dimensional case of a problem studied by A.A. Markov, L. Dubins, and J. Reeds and L. Shepp, regarding the structure of minimum-length paths with a prescribed curvature bound and prescribed initial and terminal positions and directions. In particular, we disprove a conjecture made by other authors, according to which every minimizer is a concatenation of circles and straight lines. We show that there are many minimizers —the "helicoidal arcs"— that are not of this form. These arcs are smooth and are characterized by the fact that their torsion satisfies a second-order ordinary differential equation. The solution is obtained by applying Optimal Control Theory. An essential feature of the problem is that it requires the use of Optimal Control on manifolds. The natural state space of the problem is the product of three-dimensional Euclidean space and a two-dimensional sphere. Although the problem is obviously embeddable in 6-dimensional Euclidean space, the Maximum Principle for the embedded problem yields no information, whereas a careful application of the Maximum Principle on manifolds yields a very strong result, namely, that every minimizer is either a helicoidal arc or of the form C, S, CS, SC, CSC, CCC, where C, S stand for "circle" and "segment," respectively.

1. Introduction

The purpose of this note is to announce the solution of the three-dimensional case of a problem first studied by A.A. Markov in 1889, and then by L. Dubins in 1957 and J. Reeds and L. Shepp in 1990, regarding the structure of minimum-length paths $t \to x(t)$ with a prescribed curvature bound $||x''(t)|| \leq 1$, and prescribed initial and terminal positions and tangent vectors. In particular, we present an elementary self-contained argument that disproves a widely believed conjecture, stated, e.g., by J. Reeds and L. Shepp in 1990 in [6].

Markov, in [4], and Dubins, in [2], only considered the two-dimensional case, but Dubins explicitly posed the question of what may happen in higher dimensions, so we will refer to the problem as the *Markov-Dubins problem* (MDP). More precisely, let $R > 0$, and define $\Gamma^n(R) = \cup_{\infty<a\leq b<+\infty}\Gamma^n_{a,b}(R)$, where $\Gamma^n_{a,b}(R)$ is the set of all curves $x(\cdot) : [a, b] \to \mathbb{R}^n$ that (a) are of class C^1, (b) are parametrized by arc length (i.e. $||x'(s)|| = 1$ for all s), and (c) are such that the derivative $x'(\cdot)$ is absolutely continuous and satisfies the *curvature bound* $||x''(s)|| \leq R$ for almost all $s \in [a, b]$. If $x(\cdot) : [a, b] \to \mathbb{R}^n$ is a member of $\Gamma^n(R)$, call $x(\cdot)$ an *R-minimizer* if it is of minimum length among all the paths in $\Gamma^n(R)$ that start at $x(a)$ with direction $x'(a)$ and end at $x(b)$ with direction $x'(b)$. The MDP is the problem of determining the structure of the R-minimizers.

From now on, we will always take $R = 1$ for simplicity, write $\Gamma^n_{a,b}$, Γ^n, instead of $\Gamma^n_{a,b}(1)$, $\Gamma^n(1)$, and refer to the 1-minimizers as "minimizers."

The *CSC conjecture* (cf. [6]) states, among other things, that for $n = 3$ every sufficiently short minimizer is equivalent to one of the form CSC, i.e. a concatenation of a circle, a straight-line segment, and another circle. (The precise definitions are given further below. The result actually conjectured in [6] is the stronger statement that every minimizer is CCC or CSC. Since it can be proved that a CCC arc cannot be a minimizer unless the length of the middle C is $\geq \pi$, cf. e.g. Theorem 3.1 below, the statement of [6] actually implies the CSC conjecture.) We will disprove this by studying curves with *torsion*. We introduce a quantity $N(P)$ that measures the "nonplanarity" of a pair $P = ((x_1, y_1); (x_2, y_2))$ of points in position-velocity space, and show —using arcs with very large torsion— that for all sufficiently small lengths t there is a curve $\gamma \in \Gamma^3_{0,t}$ of length t such that the nonplanarity of the pair $\partial_1\gamma = ((\gamma(0), \gamma'(0)); (\gamma(t), \gamma'(t)))$ is arbitrarily close to $\frac{4}{\pi^3}t^3$, whereas for CSC curves γ the number $N(\partial_1\gamma)$ cannot exceed $\frac{1}{8}t^3$. The fact that $\pi^3 < 32$ then disproves the CSC conjecture.

The work on the two-dimensional case by Markov in [4], Dubins in [2], and Reeds and Shepp in [6], relied heavily on very ingenious *ad hoc* methods, geometric arguments and —in [6]— computer simulations. Dubins proved in [2] that every minimizer in dimension 2 is either CCC or CSC, and then went on to state that "the nature of R-geodesics [i.e. 'R-minimizers'] for $n \geq 3$ is open."

Reeds and Shepp studied, in [6], the related problem in $\mathbb{R}^2$ where the curve $x(\cdot)$ is no longer required to be C^1, but we ask instead that there be an absolutely continuous function $y(\cdot)$ with $x'(t) = \pm y(t)$, $||y(t)|| = 1$, and $||y'(t)|| \leq 1$ for almost all t. (In other words, the "ve-

*Supported in part by NSF Grant DMS92-02554

hicle" is allowed to back up as well as move forwards.) They proved a result similar to Dubins', except that now (i) the minimizers can be concatenations of up to *five* C or S pieces, and (ii) there actually are minimizers with a more complicated structure, including whole families parametrized by an arbitrary measurable function with values in $[-1,1]$, so the precise statement of their result is no longer that *every* minimizer is "simple," but rather that every minimizer *can be replaced* by a simple one leaving unchanged the initial and terminal positions and directions as well as the length. They then stated that "These problems make sense in higher dimensions as well, but we could solve neither the forward nor reverse case mainly because we could not explicitly solve the algebraic equations involved in finding the shortest CCC from a to b in $\mathbb{R}^5$, i.e. with specified initial and terminal conditions in 3 dimensions, and so we could not use the computer to guess the answer analogous to the way we used it here. We think this answer is CCC, CSC for the forward problem [i.e. for the original Dubins problem, without 'backing up', HJS] but without much evidence or hope of doing the reduction because the equations get so complicated."

In [8] it was shown that the results of [2] and [6] for $n=2$ can be obtained without any use of the computer to guess answers, by means of a systematic approach based on applying the Maximum Principle (cf. [1], [3], [5]) of classical Optimal Control Theory, together with the more recently developed tools of Differential Geometric Control Theory (cf. [7]). Our results for $n=3$ provide a fairly complete description of the optimal paths, and rely, once again, on the tools of Optimal Control Theory, which are thus shown to be effective for a problem where other methods have failed to work. The rather long details, given in [10], will be briefly summarized in the last section. But first we proceed to present our self-contained argument disproving the CSC conjecture.

2. Proof that the CSC conjecture is false

Throughout this paper, the abbreviation "PAL" stands for "parametrized by arc-length." "Circle" means "arc of a PAL circle of radius one," and "segment" means "PAL straight line segment." Circles satisfy the equation $x'''=-x'$, which is equivalent to the existence of a parametric equation $x(t)=\bar{x}+\cos t\,\bar{y}+\sin t\,\bar{z}$. Conversely, any PAL curve that satisfies $x'''=-x'$ is a circle. Similarly, segments satisfy $x''=0$, which is equivalent to the existence of a parametric equation $x(t)=\bar{x}+t\,\bar{y}$, and any PAL arc that satisfies $x''=0$ is a segment. The concatenations of the form CSC are, of course, required to be in the class Γ^n. This means that they have to be C^1 curves, so that at a transition from C to S or from S to C the derivatives have to match.

Let $M^n=\mathbb{R}^n\times S^{n-1}$. Points of M^n will be called *states.* If $x(\cdot)\in\Gamma^n_{a,b}$, then we write $\Lambda(x(\cdot))=b-a$, and refer to $\Lambda(x(\cdot))$ as the *length* of $x(\cdot)$. The points $x(a)$ and $x(b)$ are, respectively, the *initial and terminal positions*, and $x'(a)$ and $x'(b)$ are the *initial and terminal directions* of $x(\cdot)$. The pairs $\partial_1^-(x(\cdot))=(x(a),x'(a))$ and $\partial_1^+(x(\cdot))=(x(b),x'(b))$ are the *initial and terminal states* of $x(\cdot)$, and the pair $\partial_1(x(\cdot))=(\partial_1^-(x(\cdot)),\partial_1^+(x(\cdot)))$ is the *boundary value* of $x(\cdot)$. A path $x(\cdot)$ is a *minimizer* $\Lambda(\bar{x}(\cdot))\le\Lambda(x(\cdot))$ for every path $\bar{x}(\cdot)\in\Gamma^n$ such that $\partial_1\bar{x}(\cdot)=\partial_1 x(\cdot)$. If $x(\cdot)$ is a minimizer and in addition the equalities $\partial_1\bar{x}(\cdot)=\partial_1 x(\cdot)$, $\Lambda(\bar{x}(\cdot))=\Lambda(x(\cdot))$ imply that $\bar{x}(\cdot)$ and $x(\cdot)$ coincide up to a translation of their time intervals, then we call $x(\cdot)$ a *strict minimizer.* Trivial geometric considerations show that given any two states $\bar{p},\hat{p}\in M^n$ there is an arc in Γ^n whose boundary value is $(\bar{p},\hat{p})$, and then an application of Ascoli's theorem shows that among all these arcs there is one of minimum length.

From now on we let $n=3$, and we just write $\Gamma_{a,b}$, Γ, M instead of $\Gamma^3_{a,b}$, Γ^3, M^3. Notice that an arc which is just CS or SC is contained in a plane.

Given a pair $P=(\bar{p},\hat{p})$ of points $\bar{p}=(\bar{x},\bar{y})$, $\hat{p}=(\hat{x},\hat{y})$ in M, we measure the *nonplanarity* of P by means of the scalar

$$N(P)\stackrel{\text{def}}{=}\Big|\langle\hat{x}-\bar{x},\hat{y}\times\bar{y}\rangle\Big|\,. \tag{1}$$

Notice that $N(P)=0$ if and only if there is a two-dimensional plane in $\mathbb{R}^3$ passing through the points $\bar{x}$, $\hat{x}$ to which the vectors $\bar{y}$, $\hat{y}$ are tangent. If $x(\cdot)\in\Gamma_{a,b}$, then the *nonplanarity* of $x(\cdot)$ is defined to be nonplanarity of the boundary value $\partial_1(x(\cdot))$, i.e. the number $|\langle x(b)-x(a),x'(b)\times x'(a)\rangle|$. Our result will then be a corollary of the following two simple facts about nonplanarity:

Proposition 2.1 Let $\varepsilon>0$. Then there exists a $\delta>0$ such that for every t in the open interval $(0,\delta)$ there is a curve in $\Gamma_{0,t}$ whose nonplanarity is $>(\frac{4}{\pi^3}-\varepsilon)t^3$.

Proposition 2.2 Let $t>0$, and let $x:[0,t]\to\mathbb{R}^3$ be any curve in $\Gamma_{0,t}$ which is of the form CSC. Then the nonplanarity of this curve is $\le\frac{t^3}{8}$.

PROOF THAT PROPOSITIONS 2.1 AND 2.2 IMPLY THAT THE CSC CONJECTURE IS FALSE. Proposition 2.1 says that for all sufficiently small t one can produce a curve in $\Gamma_{0,t}$ whose nonplanarity is equal to $\left(\frac{4}{\pi^3}+o(1)\right)t^3$. Proposition 2.2 says that for CSC curves in $\Gamma_{0,t}$, and t completely arbitrary, the nonplanarity cannot possibly exceed $\frac{1}{8}t^3$. In view of the inequality $\pi^3<32$, $\frac{4}{\pi^3}$ is strictly larger than $\frac{1}{8}t^3$. So, if t is sufficiently small, the first result implies that we can find a pair P of points $\bar{p}$, $\hat{p}$ in M that can be joined by a curve in $\Gamma_{0,t}$ and are such that $N(P)>\frac{t^3}{8}$. A minimum length path in Γ joining $\bar{p}$ and $\hat{p}$ must therefore have length $\tau\le t$. If this path was of the CSC form, then the second result would imply that the nonplanarity of this path is $\le\frac{1}{8}\tau^3$, which is $<N(P)$. But this is a contradiction because, by definition, the nonplanarity of *any* path joining $\bar{p}$ and $\hat{p}$ is $N(P)$. So the minimizing path is not CSC, and the Reeds-Shepp conjecture fails to be true.

PROOF OF PROPOSITION 2.1. We exhibit a curve explicitly. Pick $t > 0$. Let $\omega = \frac{\pi}{t}$, so $t = \frac{\pi}{\omega}$. Assume $t < \pi$. Then $\omega > 1$, so we can write $\omega^2 = 1 + \theta^2$, $\theta > 0$. Let $\mathbf{i}, \mathbf{j}, \mathbf{k}$ be the canonical basis of $\mathbb{R}^3$, i.e. $\mathbf{i} = (1,0,0)$, $\mathbf{j} = (0,1,0)$, $\mathbf{k} = (0,0,1)$. Let

$$x(s) = \frac{1-\cos\omega s}{\omega^2}\mathbf{i} + \frac{\theta^2\omega s + \sin\omega s}{\omega^3}\mathbf{j} + \frac{\theta(\omega s - \sin\omega s)}{\omega^3}\mathbf{k}\,.$$

This curve is obviously smooth and satisfies $x(0) = 0$, $x'(0) = \mathbf{j}$. It is easy to verify that x is PAL —i.e. $\|x'(s)\| = 1$— and that the curvature condition holds. (Actually, $\|x''(s)\| \equiv 1$.) Moreover, and although this fact will not be used directly, we point out that the *torsion* (cf. §3) of $x(\cdot)$ is constant and equal to θ.

Although (1) defines $x(s)$ for all s, we are only interested in the restriction x_t of $x(\cdot)$ to the interval $[0,t]$. To compute the nonplanarity of x_t, we first differentiate the formula for $x(s)$ and get

$$x_t'(s) = \frac{\sin\omega s}{\omega}\mathbf{i} + \frac{\theta^2 + \cos\omega s}{\omega^2}\mathbf{j} + \frac{\theta(1-\cos\omega s)}{\omega^2}\mathbf{k}\,. \quad (2)$$

If we plug in $s = t$, then $\omega t = \pi$, so

$$x_t'(t) = \frac{\theta^2 - 1}{\omega^2}\mathbf{j} + \frac{2\theta}{\omega^2}\mathbf{k}\,. \quad (3)$$

Then $x_t'(t) \times x_t'(0) = -\frac{2\theta}{\omega^2}\mathbf{i}$. On the other hand, $x_t(t) = \frac{2}{\omega^2}\mathbf{i} + \frac{\theta^2\pi t}{\omega^3}\mathbf{j} + \frac{\theta\pi}{\omega^3}\mathbf{k}$, and therefore $\left\langle x_t(t), x_t(t) \times x_t'(0)\right\rangle = -\frac{4\theta}{\omega^4} = -\frac{4\theta t^3}{\omega\pi^3}$. Clearly, ω goes to ∞ as $t \to 0$, since $\omega = \frac{\pi}{t}$. Since $\theta = \sqrt{\omega^2 - 1}$, the quotient $\frac{\theta}{\omega}$ goes to 1. Given ε, if we pick δ so small that $\frac{4\theta}{\omega\pi^3} > \frac{4}{\pi^3} - \varepsilon$ for $t < \delta$, then we see that, for $t < \delta$, the nonplanarity of x_t is $> (\frac{4}{\pi^3} - \varepsilon)t^3$, as stated. ∎

PROOF OF PROPOSITION 2.2. We assume that our curve is a concatenation of a circle, a segment, and a another circle, of lengths t_1, t_2, t_3, respectively, so that $t_1 + t_2 + t_3 = t$. Accordingly, we divide the interval $[0,t]$ into $I = [0, t_1 + t_2]$ and $J = [t_1 + t_2, t]$. Since circles and lines are given by fairly simple explicit formulas, we could compute the nonplanarity of our curve by just writing out all the formulas. We prefer, however, the following shorter and less computational argument.

Work in a coordinate system such that $x(0) = 0$. Define $f(s) = \langle x(s), x'(s) \times x'(0)\rangle$. The nonplanarity that we want to estimate is $|f(t)|$. Notice that $f(s) = 0$ for $s \in I$, since the first CS piece is contained in a plane. On the other hand, the function f is continuous and has a piecewise continuous derivative, given by $f'(s) = \langle x(s), x''(s) \times x'(0)\rangle$. The second derivative x'' has jump discontinuities at t_1 and $t_1 + t_2$. Now let us just work on the interval J, and compute one more derivative. We get

$$f''(s) = \langle x'(s), x''(s) \times x'(0)\rangle + \langle x(s), x'''(s) \times x'(0)\rangle\,,$$

and the second term is equal to $-f(s)$, because $x''' = -x'$. So f satisfies $f''(s) = g(s) - f(s)$, where $g(s) = \langle x'(s), x''(s) \times x'(0)\rangle$. Another differentiation shows —using $x''' = -x'$ again— that $g(s)$ is in fact a constant g. Since f satisfies $f'' + f = g$ on J, we have $f(s) = g + A\cos\sigma + B\sin\sigma$ for some A, B, where σ denotes $s - t_1 - t_2$. Using $f(t_1 + t_2) = 0$ we get $g + A = 0$, so $A = -g$. Also, if we let $\tau = t_1 + t_2$, we get $B = f'(\tau+) = \langle x(\tau), x''(\tau+) \times x'(0)\rangle$. So $f(s) = g(1 - \cos\sigma) + f'(\tau+)\sin\sigma$. In particular,

$$f(t_1 + t_2 + t_3) = g(1 - \cos t_3) + f'(\tau+)\sin t_3\,. \quad (4)$$

Now, $g = \langle x'(\tau), x''(\tau+) \times x'(0)\rangle$. So

$$g = \langle x'(\tau) - x'(0), x''(\tau+) \times x'(0)\rangle\,,$$

since $x'(0)$ is orthogonal to $x''(\tau+) \times x'(0)$. Since $\|x''\| \le 1$ on I, it is clear that $\|x'(\tau) - x'(0)\| \le \tau$. But $x''(\tau+)$ and $x'(0)$ are unit vectors. So we get the bound

$$|g| \le \tau\,. \quad (5)$$

Also, $x(\tau) = \int_0^\tau x'(s)ds = \tau x'(0) + \int_0^\tau\int_0^s x''(r)\,dr\,ds$, and $\|\int_0^\tau\int_0^s x''(r)\,dr\,ds\| \le \frac{\tau^2}{2}$, since $\|x''(r)\| \le 1$. So $\|x(\tau) - \tau x'(0)\| \le \frac{\tau^2}{2}$. Since

$$\begin{aligned} f'(\tau+) &= \langle x(\tau), x''(\tau+) \times x'(0)\rangle \\ &= \langle x(\tau) - \tau x'(0), x''(\tau+) \times x'(0)\rangle\,, \end{aligned}$$

and $\|x''(\tau+)\| = \|x'(0)\| = 1$, we get $|f'(\tau+)| \le \frac{\tau^2}{2}$. Using this together with (4) and (5), we get the inequality $|f(t)| \le \tau(1 - \cos t_3) + \frac{\tau^2}{2}\sin t_3$. Since $1 - \cos t_3 \le \frac{t_3^2}{2}$ and $\sin t_3 \le t_3$, it follows that

$$|f(t)| \le \frac{\tau t_3^2}{2} + \frac{\tau^2 t_3}{2}\,.$$

If we write $\tau = \mu t$, $t_3 = \nu t$, we find that $|f(t)| \le \frac{\mu\nu(\mu+\nu)}{2}t^3$. Since $\mu \ge 0$, $\nu \ge 0$, and $\mu + \nu = 1$, it follows that $\mu\nu \le \frac{1}{4}$. So $|f(t)| \le \frac{1}{8}t^3$, and our proof is complete. ∎

3. The structure of the minimizers in dimension 3

We now outline the results of [10]. Define a *helicoidal arc* in $\mathbb{R}^3$ to be a smooth PAL curve $t \to x(t)$ that has curvature 1, and is such that the torsion $t \to \tau(t)$ never vanishes, and satisfies the differential equation

$$\tau'' = \frac{3\tau'^2}{2\tau} - 2\tau^3 + 2\tau - \zeta\tau|\tau|^{1/2} \quad (6)$$

for some nonnegative constant ζ. (Recall that, if x is PAL and $\|x''(t)\| \equiv 1$, then x satisfies a differential equation $x'''(t) = -x'(t) + \tau(t)(x'(t) \times x''(t))$. The function τ is the *torsion* along x.)

Theorem 1 *For the Markov-Dubins problem in dimension three, every minimizer is either (a) a helicoidal arc or (b) a concatenation of three pieces each of which is a circle or straight line. For a minimizer of the form CCC, the middle circle has length $\ge \pi$ and $< 2\pi$.* ∎

The results of the previous section show that concatenations of circles and segments are not enough to obtain all minimizers, so *some* helicoidal arcs must be minimizers. Actually, a stronger result is true, namely,

Theorem 2 *Every helicoidal arc corresponding to a value of ζ such that $\zeta > 0$ is a local strict minimizer.* ∎

(A "local strict minimizer" is an arc $x : [a,b] \to \mathbb{R}^3$ such that there is a $\delta > 0$ with the property that the restriction of x to every subinterval of $[a,b]$ of length $\leq \delta$ is a strict minimizer.)

Remark 3.1 A constant function $\tau = c$ is a solution of (6) for some $\zeta \geq 0$ if and only if $|c| \leq 1$. This means that curves of constant torsion are not locally optimal if the torsion is large. In particular, the curves x_t considered in §2 are not locally optimal if $t < \frac{\pi}{2}$. □

We now outline the main steps of the proof of Theorem 1, and make some remarks about that of Theorem 2. To begin with, we introduce the velocity y as a new variable, taking values in the unit sphere S^2 in $\mathbb{R}^3$, and work with the class $\mathcal{A}$ of arcs of the form $t \to (x(t), x'(t))$, where $x(\cdot) \in \Gamma$. So $\mathcal{A} = \cup_{-\infty<a\leq b<+\infty}\mathcal{A}_{a,b}$, where $\mathcal{A}_{a,b}$ is the set of all M-valued absolutely continuous arcs $t \to \gamma(t) = (x(t), y(t))$, $t \in [a,b]$, such that $x' \equiv y$ and $||y'|| \leq 1$ almost everywhere. (Recall that $M = \mathbb{R}^3 \times S^2$.) If $\gamma = (x,y) \in \mathcal{A}_{a,b}$ is such an arc, we write $\partial\gamma = \partial_1 x = (\gamma(a), \gamma(b))$, $\Lambda(\gamma) = b - a$. We let $\mathcal{M}$ be the class of all arcs $\gamma = (x,y) \in \mathcal{A}$ that minimize time among all arcs $\tilde{\gamma} \in \mathcal{A}$ such that $\partial\gamma = \partial\tilde{\gamma}$.

We can realize $\mathcal{A}$ as a set of trajectories of a control system Σ by writing the dynamical equations as

$$\Sigma : \qquad x' = y \ , \ y' = y \times w \ , \tag{7}$$

where the control w is restricted to taking values in $\mathbb{B}^3$, the closed unit ball in $\mathbb{R}^3$. The controls are measurable functions $t \to w(t) \in \mathbb{B}^3$. We let $\mathcal{T}_c(\Sigma)$ denote the set of all pairs $\xi = (\gamma, w)$, where $w : I \to \mathbb{B}^3$ is a measurable map defined on some interval I, and $\gamma = (x,y)$ is a solution of (7) defined on I. We let $\mathcal{T}(\Sigma)$ denote the set of γ such that $(\gamma, w) \in \mathcal{T}_c(\Sigma)$ for some w, and use $\mathcal{A}(\Sigma)$ (resp. $\mathcal{A}_c(\Sigma)$) for the set of $\gamma \in \mathcal{T}(\Sigma)$ (resp. $(\gamma, w) \in \mathcal{T}_c(\Sigma)$) whose domain of definition is a compact interval. The elements of $\mathcal{T}(\Sigma)$, $\mathcal{T}_c(\Sigma)$, $\mathcal{A}(\Sigma)$, $\mathcal{A}_c(\Sigma)$ will be called, respectively, *trajectories*, *controlled trajectories*, *arcs* and *controlled arcs* of Σ. So $\mathcal{A}$ is exactly $\mathcal{A}(\Sigma)$, the set of arcs of Σ.

To characterize the minimum-time arcs, we use the Maximum Principle (abbr. MP). We consider a new control system Σ^* —the *Hamiltonian lift* of Σ—with state space $N = M \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$ and dynamical equations —for the variable $(x, y, \lambda, \mu, \nu) \in N$— given by

$$\Sigma^* : \quad \frac{d}{dt}(x, y, \lambda, \mu, \nu) = (\nabla_{\lambda,\mu} H \, , \, -\nabla_{x,y} H, 0) \, , \tag{8}$$

where H is the Hamiltonian, defined by

$$H = H(x, y, \lambda, \mu, \nu, w) = \langle \lambda, y \rangle + \langle \mu, y \times w \rangle + \nu \ . \tag{9}$$

Then the evolution of x and y is given by (7) (so in particular (x,y) evolves in M, i.e. $||y(t)|| = 1$ for all t if $||y(t)|| = 1$ for some t), while that of λ, μ, ν is governed by the "adjoint equations"

$$\lambda' = 0 \, , \ \ \mu' = -\lambda - w \times \mu \, , \ \ \nu' = 0 \, . \tag{10}$$

(To compute $\nabla_y H$, use the the cross-product identity $\langle A, B \times C \rangle = \langle B, C \times A \rangle$ to conclude that $H = \langle \lambda, y \rangle + \langle y, w \times \mu \rangle + \nu$.) We define the sets $\mathcal{T}(\Sigma^*)$, $\mathcal{T}_c(\Sigma^*)$, $\mathcal{A}(\Sigma^*)$, $\mathcal{A}_c(\Sigma^*)$ of *lifted trajectories*, *lifted controlled trajectories*, *lifted arcs*, and *lifted controlled arcs* in an obvious way. If we write $\Pi(x, y, \lambda, \mu, \nu, w) = (x, y, w)$, then it is clear that $\Xi = (x, y, \lambda, \mu, \nu, w)$ is a lifted controlled trajectory if and only if $\xi = \Pi \circ \Xi$ is a controlled trajectory and (10) holds. (In that case, Ξ will be called a *lift* of ξ.) We call a lifted controlled trajectory $\Xi = (x, y, \lambda, \mu, \nu, w)$ *H-minimizing* if $H(\Xi(t)) = \min\{H(x(t), y(t), \lambda(t), \mu(t), \nu(t), v) : v \in \mathbb{B}^3\}$ for a.e. t.

The MP says that, if $\xi = (\gamma, w) \in \mathcal{A}_c$ and $\gamma = (x,y)$ is a minimizer, then there exists a lift $\Xi = (x, y, \lambda, \mu, \nu, w)$ which is H-minimizing and such that (i) $\nu \geq 0$, (ii) $H(\Xi(t)) = 0$ for a.e. t, and (iii) Ξ is nontrivial, in the sense that

$$||\lambda|| + ||\mu(t) \times y(t)|| > 0 \ \text{for all} \ t \, . \tag{11}$$

(The reason for this formulation of the nontriviality property is as follows: we really ought ot be working on the cotangent bundle T^*M of M, rather than on $M \times \mathbb{R}^3 \times \mathbb{R}^3$. That is, (λ, μ) should be regarded as a covector on M at (x,y) or, equivalently, μ should be regarded as a linear functional $z \to \langle \mu, z \rangle$ on the tangent space $T_y S^2$. This functional vanishes iff $\mu \times y = 0$, so (11) says precisely that $(\lambda, \mu) \neq 0$ *as a covector on M*. The distinction between (11) and the "naïve" nontriviality condition $(\lambda, \mu) \neq (0,0)$ is crucial, as explained in Remark 3.2 below.)

From now on, an H-minimizing lift Ξ of a $\xi \in \mathcal{A}_c(\Sigma)$ for which (i), (ii) and (iii) above hold will be said to be an *MP controlled arc*, and an *MP lift* of ξ.

To translate the minimization condition into more familiar terms, we first use $\langle A, B \times C \rangle = \langle B, C \times A \rangle$ again, to rewrite H as $H = \langle \lambda, y \rangle + \langle w, \mu \times y \rangle + \nu$. Write

$$\varphi \stackrel{\text{def}}{=} \langle \lambda, y \rangle \, , \quad W \stackrel{\text{def}}{=} y \times \mu \, . \tag{12}$$

It is then clear that H is minimized by taking

$$w = -\frac{\mu \times y}{||\mu \times y||} \ \text{if} \ \mu \times y \neq 0 \, . \tag{13}$$

If $\mu \times y = 0$, then any value of w in $\mathbb{B}^3$ is minimizing. We can capture both cases in one formula by writing the minimization condition as $||\mu \times y|| w = -\mu \times y$, i.e. as

$$||W|| w = W \, . \tag{14}$$

Similarly, (11) and the condition that $H = 0$ say that

$$||\lambda|| + ||W|| > 0 \ \text{and} \ ||W|| = \varphi + \nu \, . \tag{15}$$

We now make a deeper analysis of the lifted trajectories. Notice that, as long as $W \neq 0$, (7) and (10) constitute a closed, smooth system of ordinary differential equations for x, y, λ, μ, ν, if we use the formula $w = \frac{W}{||W||}$ for w. Let us call a $\Xi \in \mathcal{T}_c(\Sigma^*)$ "nice" if W never vanishes along Ξ. We will be particularly interested in understanding how nice lifted trajectories connect up with trajectories for which $W = 0$.

To carry out our analysis, it is useful to to reduce the dimension of the system by finding integrals of motion, i.e. functions that are constant along trajectories. To find such integrals, we use the control theory version of Noether's Theorem (cf. Sussmann [9]). Our system is clearly invariant under the 6-dimensional group of rigid motions of $\mathbb{R}^3$. Consider first the action of the translations. For $v \in \mathbb{R}^3$, let τ_v be the translation $x \to x+v$. Then τ_v acts on the state and control variables of our system via $\tau_v(x,y,w) = (x+v,y,w)$. The infinitesimal generators of the action of the translations on M are the vector fields X_v given by $X_v(x,y) = (v,0)$. The Hamiltonian function corresponding to X_v is $h_{X_v} = \langle\lambda, v\rangle$. By Noether's Theorem, this function is constant along MP controlled arcs. Since this is true for every vector v, we have rediscovered the fact that λ itself has to be constant, as we already knew from (10).

A much more interesting conservation law is derived by using rotational invariance. A rotation matrix $R \in SO(3)$ acts on $M \times \mathbb{B}^3$ via $R(x,y,w) = (Rx, Ry, Rw)$. The infinitesimal generators are the skew-symmetric matrices $A \in so(3)$. To each such matrix there corresponds a vector field Y_A on M, given by $Y_A(x,y) = (Ax, Ay)$. The Hamiltonian h_{Y_A} corresponding to Y_A is given by $h_{Y_A} = \langle\lambda, Ax\rangle + \langle\mu, Ay\rangle$. So this quantity has to be conserved for every A. Now recall that the skew-symmetric transformations or $\mathbb{R}^3$ are exactly the maps of the form $u \to v \times u$, where $v \in \mathbb{R}^3$. So what we have shown is that the expression $\langle\lambda, v\times x\rangle + \langle\mu, v\times y\rangle$ has to be conserved for every vector v. Using $\langle A, B\times C\rangle = \langle B, C\times A\rangle$ again, we can rewrite this expression as $\langle v, x\times\lambda\rangle + \langle v, y\times\mu\rangle$. Since this has to be constant for every v, we conclude that the vector

$$V \stackrel{\text{def}}{=} x\times\lambda + y\times\mu = x\times\lambda + W \qquad (16)$$

is conserved. (This can be verified directly using (10) and (14). Indeed, $W' = y'\times\mu + y\times\mu'$, so (10) implies that $W' = (y\times w)\times\mu - y\times\lambda - y\times(w\times\mu)$. The identity $A\times(B\times C) = (A\times B)\times C + B\times(A\times C)$ then gives $y\times(w\times\mu) = (y\times w)\times\mu + w\times(y\times\mu)$. Then (14) gives $w\times(y\times\mu) = w\times W = ||W||w\times w = 0$, so $y\times(w\times\mu) = (y\times w)\times\mu$. Then

$$W' = -y\times\lambda = \lambda\times y. \qquad (17)$$

This implies that $V' = y\times\lambda + W' = 0$.)

Moreover, we can also conclude that λ and V cannot both vanish. (Indeed, if $\lambda = 0$ and $V = 0$, then $W = 0$, so $||\lambda|| + ||W|| = 0$, contradicting (15).)

We now consider a controlled arc $\xi = (\gamma, w)$ such that $\gamma = (x,y)$ is optimal, and let $\Xi = (x,y,\lambda,\mu,\nu,w)$ be an MP lift of ξ. It is useful to introduce the scalar conserved quantity

$$C \stackrel{\text{def}}{=} \langle\lambda, V\rangle\,. \qquad (18)$$

Clearly, $C = \langle\lambda, W\rangle$, since $\lambda \perp x\times\lambda$. Then (14) implies

$$C = \psi||W||\,,\ \text{where}\ \psi \stackrel{\text{def}}{=} \langle\lambda, w\rangle\,. \qquad (19)$$

Now assume that Ξ is nice, so $W \neq 0$ along Ξ. We can then write the equations for γ in the form

$$x' = y\ ,\ y' = y\times\left(\frac{V - x\times\lambda}{||V - x\times\lambda||}\right), \qquad (20)$$

with full assurance that the denominator never vanishes.

In addition, V cannot be arbitrary, because (16) implies that $V - x\times\lambda \perp y$, since $W \perp y$. Moreover, the number $\nu = ||W|| - \langle\lambda, y\rangle$ has to be a nonnegative constant.

Now, for a $V, \lambda \in \mathbb{R}^3$, $\nu \in \mathbb{R}$, we write

$$\Omega_{V,\lambda} = \{(x,y) \in \mathbb{R}^3\times\mathbb{R}^3 : V - x\times\lambda \neq 0\}\,, \qquad (21)$$

and let $\Omega^o_{V,\lambda,\nu}$ be the set of those $(x,y) \in \Omega_{V,\lambda}$ for which $||y||=1$, $V - x\times\lambda \perp y$ and $\nu = ||V - x\times y|| - \langle\lambda, y\rangle$. Let us call a solution of (20) "good" if it is contained in $\Omega^o_{V,\lambda,\nu}$ for some $\nu \geq 0$. We have shown that *every nice MP controlled arc Ξ gives rise to a good solution γ of (20).*

Conversely, we show that *every solution $\gamma = (x,y)$ of (20) that goes through a point of a set $\Omega^o_{V,\lambda,\nu}$ —for some real ν— is entirely contained in $\Omega^o_{V,\lambda,\nu}$* (that is, *the set $\Omega^o_{V,\lambda,\nu}$ is invariant under the flow defined on $\Omega_{V,\lambda}$ by (20)*), *and if $\nu \geq 0$ then γ is a good solution of (20) and arises from a nice MP controlled trajectory.* To prove this, we first let $\tilde{W} = V - x\times\lambda$, and define w via $w = \frac{\tilde{W}}{||\tilde{W}||}$. Then $y' = y\times w$, so $y' \perp y$. Since there is a t_0 such that $||y(t_0)|| = 1$, we conclude that $||y(t)|| = 1$ for all t, so γ is M-valued. Since $||w(t)|| = 1$, w is a control, and γ is a trajectory for w. We pick a time t_0 in the domain of γ such that $y(t_0) \perp \tilde{W}(t_0)$, and then choose a μ_0 such that $y(t_0)\times\mu_0 = \tilde{W}(t_0)$. We then define $\mu(t)$ by solving the adjoint equation $\mu' = -\lambda - w\times\mu$ with $\mu(t_0) = \mu_0$. We now have to show that the minimization condition holds, and this will follow if we prove that $\tilde{W} = W$, where $W = y\times\mu$. The desired equality holds for $t = t_0$. The adjoint equation gives $W' = -(y\times\lambda) - w\times W$, and it is clear that $\tilde{W}' = -y\times\lambda$. So $(\tilde{W} - W)' = w\times W = ||\tilde{W}||\tilde{W}\times W = ||\tilde{W}||\tilde{W}\times(W - \tilde{W})$. Therefore $\tilde{W} - W$ is a solution of a linear homogeneous O.D.E. Since $\tilde{W}(t_0) - W(t_0) = 0$, we conclude that $W(t) = \tilde{W}(t)$ for all t. So $w = \frac{W}{||W||}$, and the minimization condition holds. Next, we define $\nu(t) = ||W(t)|| - \langle\lambda, y(t)\rangle$. Since $W' = \lambda\times y$, we have $\frac{d}{dt}(||W||) = \langle w, W'\rangle = \langle w, \lambda\times y\rangle = \langle\lambda, y\times w\rangle = \langle\lambda, y'\rangle$. So $\nu' \equiv 0$. So ν is a constant, and then γ is entirely contained in $\Omega^o_{V,\lambda,\nu}$. If $\nu \geq 0$, then γ is good, and $\Xi = (x,y,\lambda,\mu,\nu,w)$ is an MP lift of $\xi = (\gamma, w)$, completing the proof of our statement.

Next, we have to study the maximal good solutions of (20). Suppose $\gamma : I \to \Omega^o_{V,\lambda,\nu}$ is such a solution, and let $\mu(\cdot)$ be as above. (Naturally, μ need not be unique.) We show that if $C \neq 0$ then $I = \mathbb{R}$. (The maximal solutions corresponding to $C = 0$ will be studied later.) To see this, let $\hat{t} = \sup I$, $\tilde{t} = \inf I$. Suppose $\hat{t} < \infty$. Then the limits $\hat{x}$, $\hat{y}$ of $x(t)$, $y(t)$ as $t \to \hat{t}-$ exist, because $||x'(t)|| = ||y'(t)|| = 1$. To prove that the solution can be extended to the right of $\hat{t}$ —which would be a contradiction— it suffices to show that $(\hat{x},\hat{y}) \in \Omega_{V,\lambda}$. To see this, notice that $C = \langle\lambda, V\rangle = \langle\lambda, W(t)\rangle$, since $\langle\lambda, x(t)\times\lambda\rangle = 0$. Since C is constant and nonzero, the vector $\hat{W} = V - \hat{x}\times\lambda = \lim_{t\to\hat{t}-}(V - x(t)\times\lambda) = \lim_{t\to\hat{t}-} W(t)$ must satisfy $\langle\lambda, \hat{W}\rangle = C \neq 0$. Therefore $\hat{W} \neq 0$, so $(\hat{x},\hat{y}) \in \Omega_{V,\lambda}$. This shows that $\hat{t} = +\infty$, and a similar argument proves that $\tilde{t} = -\infty$.

The solutions of (20) are obviously smooth. We now show that the good solutions corresponding to $C \neq 0$ are precisely the helicoidal arcs defined earlier, and analyze the good solutions for $C = 0$. So we assume until further notice that $\gamma = (x, y)$ is a good solution —for a given λ, V, ν— and let μ, W, w be as above, so $\Xi = (x, y, \lambda, \mu, \nu, w)$ is a nice MP lift of $\xi = (\gamma, w)$.

The vectors w, y —we will from now on use this notation rather than $w(t)$, $y(t)$— are orthogonal and of unit length, and $y' = y \times w$, so the triples (y, w, y') and $(y, y', -w)$ are positively oriented orthonormal bases of $\mathbb{R}^3$. Therefore $y = w \times y'$ and $w = -y \times y'$. Moreover, every vector $v \in \mathbb{R}^3$ has a unique expression

$$v = \langle v, y\rangle y + \langle v, y'\rangle y' + \langle v, w\rangle w \,. \qquad (22)$$

Using (22) with $v = y'' + y$, and observing that $0 = \frac{d}{dt}\langle y', y\rangle = \langle y'', y\rangle + \langle y', y'\rangle = \langle y'', y\rangle + \langle y, y\rangle = \langle y'' + y, y\rangle$, and $\langle y'', y'\rangle = \langle y, y'\rangle = 0$, we find that $y'' + y$ is a multiple of w, so

$$y'' + y = \tau(y \times y') = -\tau w \qquad (23)$$

for some $\tau = \tau(t)$. By definition, τ is the torsion. To compute τ, we first determine y'' using (7), and get the identity $y'' = y \times w' + y' \times w$. Since $y = w \times y'$, we get $y'' + y = y \times w'$. To compute w' we use $w = \frac{W}{||W||}$, so that $w' = \frac{W' - \langle W', w\rangle w}{||W||}$. We compute W' from (17), and get $W' = \lambda \times y$. Using (22) with $v = \lambda$, we write

$$\lambda = \varphi y + \varphi' y' + \psi w \,. \qquad (24)$$

Then $\lambda \times y = \varphi' y' \times y + \psi w \times y$, so $W' = \lambda \times y = \varphi' w - \psi y'$, and $W' - \langle W', w\rangle w = -\psi y'$. So $w' = -\frac{\psi}{||W||} y'$. Then $y'' + y = y \times w' = -\frac{\psi}{||W||} y \times y'$, showing that $\tau = -\frac{\psi}{||W||}$. Then (15) and (19) imply that

$$\tau = -\frac{\psi^2}{C} = -\frac{C}{||W||^2} = -\frac{C}{(\varphi + \nu)^2} \,. \qquad (25)$$

If we take the inner product of (23) with λ and use (12) and (19), we get $\varphi'' + \varphi = -\tau\psi$. Then (15), (19) and (25) imply

$$\varphi'' + \varphi = \frac{C^2}{(\varphi + \nu)^3} \,. \qquad (26)$$

For a given constant C, the global existence theory for the solutions of the second order O.D.E. (26) on the half-plane $P_\nu = \{(\varphi, \varphi') : \varphi > -\nu\}$ is easily studied by remarking that

$$\kappa \stackrel{\text{def}}{=} \varphi^2 + \varphi'^2 + \frac{C^2}{(\varphi + \nu)^2} \qquad (27)$$

is constant along solutions. (This is easily verified directly, by differentiating (27) and using (26). We study the solutions of (26) on P_ν only, because at this point we are looking at nice Ξ, for which $||W|| > 0$, so $\varphi > -\nu$.) If $C > 0$, the constancy of κ implies that the solutions of (26) on P_ν cannot have explosions in finite time.

When $C = 0$ the situation is different, since in that case nothing prevents φ from approaching the value $-\nu$ in finite time. The general solution of (26) when $C = 0$ is $\varphi = A\cos(t - t_0)$, with A and t_0 arbitrary, $A \geq 0$, and then those maximal solutions for which $0 \leq A < \nu$ are globally defined, whereas those for which $0 < A \leq \nu$ are defined on intervals of the form $(t_0 - \alpha, t_0 + \alpha)$, where α is characterized by $A\cos\alpha = -\nu$, $0 < \alpha \leq \pi$. Since $\nu \geq 0$, we see that 2α is always $\geq \pi$. (The possibility that $A = \nu = 0$ is excluded, because it implies $\varphi \equiv 0$, so $||W|| = 0$ by (15), contradicting the fact that Ξ is nice.)

If $C \neq 0$, then we can use the transformation $\tau = -C(\varphi + \nu)^{-2}$ —cf. (25)— to turn (26) into an O.D.E. for τ by means of a routine calculation. The result is precisely Equation (6), with $\zeta = \frac{2\nu}{\sqrt{|C|}}$. The global existence theory for (26) then gives a similar theory for (6), and the global solutions τ will be positive if $C < 0$, negative if $C > 0$.

The above remarks already prove that every optimal γ that has an MP lift for which $C \neq 0$ is a helicoidal curve. To prove Theorem 1, we must show that the optimal γ's for which $C = 0$ are as in (b) of Theorem 1. Let $\xi = (\gamma, w) : [a, b] \to M \times \mathbb{B}$ be a controlled arc such that $\gamma = (x, y)$ is optimal and ξ has an MP lift $\Xi = (x, y, \lambda, \mu, \nu, w)$ for which $C = 0$. Let $L = \{t \in I : W(t) \neq 0\}$. Then L is relatively open in $[a, b]$. So L is the union of a finite or countable set $\mathcal{J}$ of pairwise disjoint relatively open subintervals of $[a, b]$.

We first treat the case when $L = \emptyset$. In this case, $||W|| \equiv 0$. It follows from (15) and (17) that $\lambda \neq 0$ and $\lambda \times y = 0$. Since $||y|| = 1$ and y is continuous, it follows that y =constant. So *when $L = \emptyset$ γ is a straight line.*

Remark 3.2 The previous seemingly trivial step is in fact crucial, and depends in a fundamental way on the fact that we have formulated the MP on a manifold. The key point is the implication $||W|| = 0 \Rightarrow \lambda \neq 0$, which follows from (11). If we had used $\mathbb{R}^6$ rather than $\mathbb{R}^3 \times S^2$ as our state space, then the nontriviality condition would just have said that $(\lambda, \mu) \neq (0, 0)$. It is easy to see that *for every trajectory of Σ one can find λ, μ, ν that satisfy all the conditions of the MP other than (11), as well as the weaker nontriviality property $(\lambda, \mu) \neq (0, 0)$.* Therefore, *if the problem is formulated in $\mathbb{R}^6$ rather than in $\mathbb{R}^3 \times S^2$, then every trajectory is an extremal and the Maximum Principle gives no information whatsoever.* □

From now on we assume that $L \neq \emptyset$, so $\mathcal{J} \neq \emptyset$. Let $\tilde{\mathcal{J}}$ be the set of $J \in \mathcal{J}$ such that J contains one of the endpoints a, b. Then $\tilde{\mathcal{J}}$ has at most two elements, and the members of $\hat{\mathcal{J}} = \mathcal{J} \backslash \tilde{\mathcal{J}}$ are open intervals. It is clear that the restriction of γ to each $J \in \mathcal{J}$ is a good solution of (20), and is in fact maximal if $J \in \hat{\mathcal{J}}$. On each $J \in \mathcal{J}$, (23) holds with $\tau = 0$. So $y'' + y = 0$, i.e. $x''' = -x'$. Since x is PAL, we conclude that x restricted to J is a circle. As explained above, $\varphi(t) = A_J\cos(t - t_J)$ on each $J \in \mathcal{J}$. The number κ defined in (27) —which in principle might have depended on J— is equal to $||\lambda||^2$, because (24) implies that $||\lambda||^2 = \varphi^2 + \varphi'^2 + \psi^2$, and $\psi \equiv 0$ on J, since $C = 0$ and $||W|| > 0$. On the other hand, κ obviously equals A_J^2. So $\varphi(t) = A\cos(t - t_J)$ on each $J \in \mathcal{J}$, where $A = \sqrt{\kappa} = ||\lambda||$.

If $||\lambda|| < \nu$, i.e. $A < \nu$, then the maximal good solutions of (20) are globally defined, so $\mathcal{J} = \{[a, b]\}$. Therefore γ itself is a circle. Naturally, the length has to be $< 2\pi$ for otherwise γ would contain a loop and would not be optimal.

Next, we assume that $||\lambda|| \geq \nu$. Then each $J \in \hat{\mathcal{J}}$ has length $2\alpha_J \geq \pi$, where α_J is determined by A and ν, so α_J is in fact determined by $||\lambda||$ and ν and is equal to an α which is independent of J. Therefore all the intervals $J \in \hat{\mathcal{J}}$ have the same length $2\alpha \geq \pi$, and the intervals $J \in \tilde{\mathcal{J}}$ have length $\leq 2\alpha$. So $\hat{\mathcal{J}}$ is finite, and then the closed set $Q = [a, b] \backslash L$ $(= \{t : W(t) = 0\})$ is also a finite union of intervals. If $Q = \emptyset$, then $\mathcal{J} = \{[a, b]\}$ and γ is once again a circle of length $< 2\pi$.

Now assume that $Q \neq \emptyset$, and let $\mathcal{K}$ be the set of connected components of Q. Then $\mathcal{K}$ is finite and nonvoid, and each $K \in \mathcal{K}$ is a compact interval. We will consider separately the cases $||\lambda|| > \nu$ and $||\lambda|| = \nu$.

Assume first that $||\lambda|| > \nu$. Pick a $K \in \mathcal{K}$, $K = [c, d]$. Then either $a < c$ or $d < b$, since we are assuming that $L \neq \emptyset$. Suppose $d < b$. Then $W'(d) = \lambda \times y(d)$, so $||\lambda||^2 = ||W'(d)||^2 + \varphi(d)^2 = ||W'(d)||^2 + (\nu - ||W(d)||)^2 = ||W'(d)||^2 + \nu^2$, using $||W(d)|| = 0$ together with (15) and (17). Therefore $W'(d) \neq 0$, so $W(t) \neq 0$ for t close enough to d, as long as $t \neq d$. So $c = d$. Then Q is a finite set, and γ is a finite concatenation of circles all of which have the same length $2\alpha \geq \pi$ except possibly for the first and last ones, that are of length $\leq 2\alpha$. Moreover, for each $J \in \mathcal{J}$, the corresponding circle is contained in a two-dimensional plane P_J, orthogonal to the vector w. (It follows from $y'' + y = 0$ that $w = y' \times y$ is constant on J.) If $\bar{t} \in Q$, $a < \bar{t} < b$, then $W(t) = (t - \bar{t})W'(\bar{t}) + o(|t - \bar{t}|)$, for t near $\bar{t}$. So $w(t) = \mathrm{sgn}(t - \bar{t})W'(\bar{t}) + o(1)$, for t near $\bar{t}$, $t \neq \bar{t}$. Therefore the left and right limits $w(\bar{t}-)$, $w(\bar{t}+)$ satisfy $w(\bar{t}-) + w(\bar{t}+) = 0$. Since w is constant to the left and to the right of $\bar{t}$, we see that w changes sign at $\bar{t}$. But then the planes P_J corresponding to the two intervals J such that $\bar{t} \in \mathrm{Clos}(J)$ must coincide. This shows that the entire arc $x(\cdot)$ is contained in a plane P in $\mathbb{R}^3$. It follows in particular that $x(\cdot)$ minimizes length among all the arcs in $\mathcal{A}$ that have the same initial and terminal conditions as $x(\cdot)$ and are contained in P. So $x(\cdot)$ is a solution of the two-dimensional version of our problem. The solution of the problem in dimension 2 was obtained by Dubins in [2]. He showed that every optimal trajectory is a concatenation of at most three pieces, at most one of which is a straight line segment, while the others are circles. Using this result, we conclude that our arc γ is CC or CCC. (The fact that the middle C has length $\geq \pi$, which has been established above, also follows from Dubins' result.)

Finally, we consider the case when $||\lambda|| = \nu$. Now the zeros of W no longer need to be isolated, so there may be nontrivial intervals belonging to $\mathcal{K}$. As shown before in our analysis of the case $L = \emptyset$, the restriction of γ to every such interval is a straight line segment. On the other hand, the set $\hat{\mathcal{J}}$ must be empty, because if $J \in \hat{\mathcal{J}}$ then the corresponding circle would have length 2α, where $\cos\alpha = -\frac{\nu}{||\lambda||} = -1$. Then $\alpha = \pi$. This means that $x(\cdot)$ contains a full loop, contradicting optimality. So γ is C or S or C or CS or SC or CSC, as stated. This concludes the proof or Theorem 1.

We conclude with some brief comments about the proof of Theorem 2 (cf. [10] for full details). To prove the local optimality of the helicoidal arcs, one has to invert the transformation used to go from (26) to (6) and show that to every solution τ of (6), for a $\zeta \geq 0$, there corresponds a solution φ of (26), if we pick a nonzero C in an arbitrary fashion, and define $\nu = \frac{1}{2}\zeta\sqrt{|C|}$. Using this, one shows that if a smooth PAL curve x satisfies $||x''|| \equiv 1$, and its torsion τ is a solution of (6), then the corresponding controlled arc $\xi = (x, x', w)$ has an MP lift Ξ. (Here $w = y' \times y$.) The proof of optimality is then done by a Hamilton-Jacobi technique, embedding Ξ in a suitable field of MP controlled arcs.

References

[1] Berkovitz, L.D., *Optimal Control Theory*, Springer-Verlag, New York, 1974.

[2] Dubins, L.E., "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), pp. 497-516.

[3] Lee, E.B. and L. Markus, *Foundations of Optimal Control Theory*, J. Wiley, New York (1967).

[4] Markov, A.A., "Some examples of the solution of a special kind of problem in greatest and least quantities," (in Russian) *Soobshch. Karkovsk. Mat. Obshch.* 1 (1887), pp. 250-276.

[5] Pontryagin, L.S., V.G. Boltyansky, R.V. Gamkrelidze and E.F. Mischenko, *The Mathematical Theory of Optimal Processes*, Wiley, New York, 1962.

[6] Reeds, J.A. and L.A. Shepp, "Optimal paths for a car that goes both forwards and backwards," *Pacific J. Math.* 145 (1990), pp. 367-393.

[7] Sussmann, H.J., "Lie brackets, real analyticity and geometric control theory," in *Differential Geometric Control Theory*, R.W. Brockett, R.S. Millman and H.J. Sussmann Eds., Birkhäuser Boston Inc. (1983), pp. 1-115.

[8] Sussmann, H.J., and G. Tang, "Shortest paths for the Reeds-Shepp car: a worked out example of the use of geometric techniques in nonlinear optimal control," Rutgers Center for Systems and Control (Sycon) Report 91-10, September 1991. To appear in *SIAM. J. Control.*

[9] Sussmann, H.J., "Symmetries and Integrals of Motion in Optimal Control," to appear in *Geometry in Nonlinear Control and Differential Inclusions,* A. Fryszkowski, B. Jakubczyk, W. Respondek, T. Rzezuchowski Eds., Banach Center Publications, Institute of Mathematics of the Polish Academy of Sciences.

[10] Sussmann, H.J., "Shortest paths with a Prescribed Bound on the Curvature: the Three-dimensional Case," In preparation.