PACIFIC JOURNAL OF MATHEMATICS
Vol. 145, No. 2, 1990

OPTIMAL PATHS FOR A CAR THAT GOES BOTH FORWARDS AND BACKWARDS

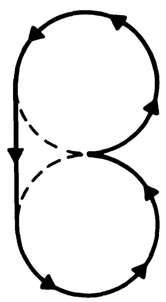
J. A. REEDS AND L. A. SHEPP

The path taken by a car with a given minimum turning radius has a lower bound on its radius of curvature at each point, but the path has cusps if the car shifts into or out of reverse gear. What is the shortest such path a car can travel between two points if its starting and ending directions are specified? One need consider only paths with at most 2 cusps or reversals. We give a set of paths which is *sufficient* in the sense that it always contains a shortest path and *small* in the sense that there are at most 68, but usually many fewer paths in the set for any pair of endpoints and directions. We give these paths by explicit formula. Calculating the length of each of these paths and selecting the (not necessarily unique) path with smallest length yields a simple algorithm for a shortest path in each case. These optimal paths or geodesics may be described as follows: If C is an arc of a circle of the minimal turning radius and S is a line segment, then it is sufficient to consider only certain paths of the form $CCSCC$ where arcs and segments fit smoothly, one or more of the arcs or segments may vanish, and where reversals, or equivalently cusps, between arcs or segments are allowed. This contrasts with the case when cusps are not allowed, where Dubins (1957) has shown that paths of the form CCC and CSC suffice.

1. Introduction. We want to find a shortest path in the plane with specified initial and final points and directions and with the further constraint that at each point the radius of curvature should be ≥ 1. This problem arose in a simple model for a robot cart which moves under computer control. The cart can shift into reverse and so the path is allowed to have cusps.

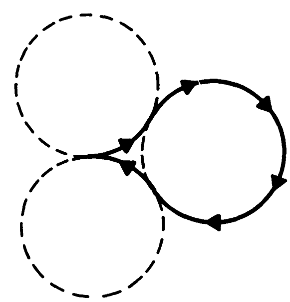
In an elegant paper, Lester Dubins (1957) solved the problem when the car cannot reverse and cusps are not allowed. Even in this case it is apparently impossible to give an explicit formula for the shortest path. Instead Dubins gives a sufficient set of paths, i.e. a set which always contains what he called a *geodesic*, or optimal path. His sufficient set is so small that there are at most 6 contenders in the set for each case of specified endpoint conditions, and it is a simple matter to find the shortest of these 6, which gives an algorithm for the solution. He showed that any geodesic can be described by one of 6 words: $\ell r \ell$,

$\ell s\ell$, ℓsr, $r\ell r$, rsr, $rs\ell$ where ℓ, r, and s stand for "go left", "go right", and "go straight", respectively. Here left and right mean anticlockwise or clockwise around a *circle of unit radius*, i.e. a tightest possible circle, and of course one always goes less than 2π around any circle. More compactly, Dubins proved that a geodesic must be a smooth curve that is piecewise circular (radius 1) or linear, with at most 3 pieces, and always takes the form CCC or CSC where C is an arc of a unit circle and S is a line segment. A *word* notation like CCC or ℓsr thus stands for the corresponding class of paths. We use subscripts on a word (as in $\ell_t s_u r_v$) to specify the length of the corresponding arcs or segments involved. Note that one or more of these lengths may vanish. For example, to choose a path which returns to the initial point but in the opposite direction, two competing paths of Dubins type suggest themselves: $\ell_{3\pi/2} s_2 \ell_{3\pi/2}$ and $\ell_{\pi/3} r_{5\pi/3} \ell_{\pi/3}$ (see Figs. A, B). It is easy to verify that both of these accomplish the job



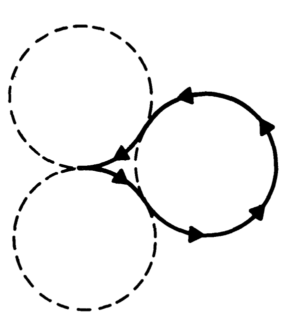
$\ell_{\frac{3\pi}{2}} s_2 \ell_{\frac{3\pi}{2}}$

FIGURE A



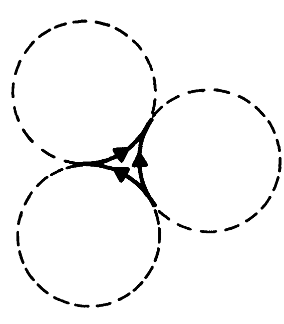
$\ell_{\frac{\pi}{3}} r_{\frac{5\pi}{3}} \ell_{\frac{\pi}{3}}$

FIGURE B



$r_{\frac{\pi}{3}} \ell_{\frac{5\pi}{3}} r_{\frac{\pi}{3}}$

FIGURE C



$\ell^{+}_{\frac{\pi}{3}} r^{-}_{-\frac{\pi}{3}} \ell^{+}_{\frac{\pi}{3}}$

FIGURE D

of reversing in place but the second is shorter and turns out to be optimal by Dubins's theorem, as is the symmetrically reflected path, $r_{\pi/3}\ell_{5\pi/3}r_{\pi/3}$ (see Fig. C). Note that in a crude dimensional sense, the number of free parameters t, u, v exactly matches the number (3) of conditions of endpoint and direction. For each of the words $\ell s \ell$, ℓsr, rsr and $rs\ell$ it is not hard to show that there is at most one path obeying the end conditions.

Although there may actually be two distinct paths of form $\ell_t r_u \ell_v$ or of form $r_t \ell_u r_v$, Dubins shows [**4**, p. 513, Sublemma] that only one of them, with $u > \pi$, can be a geodesic.

With this fact it is not hard to show that there is at most one geodesic for each word, although there may be two different words which are both optimal, as in the reverse-in-place example.

Dubins has given an effective algorithm for the forward problem, but what if the car can reverse?

We remark that A. A. Markov in 1887 (Markov (1887), see also [Krein-Nudelman, p. 17]) considered and solved various related versions of these problems in his work on laying railroad track connecting already existing sections of track. Other papers of interest are Melzak (1961) and Dubins (1961).

If cusps are allowed in the path then we must consider words built from ℓ^+, ℓ^-, r^+, r^-, s^+, and s^-, where ℓ^+ means turning to the left while going forwards, ℓ^- means turning to the left while going backwards, etc. Note that a path of the form $\ell^+ r^-$, for example, has a cusp whereas $\ell^+ r^+$ or $\ell^- r^-$ has no cusp. (In car-driving terms, the letters ℓ, r, s refer to the steering wheel and the signs $+$ and $-$ refer to the gear shift.)

It is easy to see that cusps can sometimes shorten Dubins's paths; for example, in the reverse-in-place problem, the path $\ell^+_{\pi/3} r^-_{\pi/3} \ell^+_{\pi/3}$ (see Fig. D) reverses in place and is shorter (in fact, it is optimal). Here the superscript indicates the direction taken along the corresponding arc or segment. We give a set of words in $\ell^\pm$, $r^\pm$, $s^\pm$ which give a solution to the reverse problem analogous to that of Dubins for the forward problem. These are more compactly given in C, S notation:

$$(1.1) \quad \begin{gathered} C^+C^-C^+, \quad C^+C^-C^-, \quad C^+S^+C^+, \quad C^+C^+_uC^-_uC^-, \\ C^+C^-_uC^-_uC^+, \quad C^-C^+_{\pi/2}S^+C^+, \quad C^-C^+_{\pi/2}S^+C^+_{\pi/2}C^-, \\ C^+C^+C^-, \quad C^-S^+C^+_{\pi/2}C^+ \end{gathered}$$

together with the words obtained by reversing all the signs. Here C

stands for either ℓ or r so CC means either ℓr or $r\ell$. A $C^+_{\pi/2}$ or $C^-_{-\pi/2}$ means the corresponding ℓ or r must be of length $\pi/2$, and the combination $C_u C_u$ in (1.1) means the two corresponding circular segments have equal lengths. For more precision we will use the somewhat redundant convention that in $C^\pm_t$, $\ell^\pm_t$, $r^\pm_t$, or $S^\pm_t$ the sign of t should match the direction, i.e. $t > 0$ when the car is going forward, $t < 0$ when it is reversing a distance $|t|$ along a left-circle, right-circle, or straight line. Note that in a family of paths such as $C^+_t C^-_u C^-_v C^+_w$ the number (4) of free parameters t, u, v, w is one more than the number (3) of end conditions, and so there typically is a manifold of solutions, t, u, v, w for given end conditions. Optimizing among the paths of the family gives an extra equation such as $v = u$, or $v = \pi/2$, so that in each case in (1.1) only 3 free parameters remain. Of course any length not $\pi/2$ can vanish, so subwords are included as well. We will show (in the remark below the proof of Lemma 3) that there are paths—for example, in the class $C^+C^-C^+C^-C^+C^-$—which are not of any of the above types and which are geodesics but in all such cases there are equally short geodesics inside our class (1.1). This phenomenon of geodesics that are not in the sufficient set does not occur in the simpler forward problem. In a still more compact notation that avoids $\pm$, we may write the list of sufficient special paths as

$$\begin{aligned}(1.2)\quad & C|C|C, \quad CC|C, \quad CSC, \quad CC_u|C_uC, \quad C|C_uC_u|C, \\ & C|C_{\pi/2}SC, \quad C|C_{\pi/2}SC_{\pi/2}|C, \quad C|CC, \quad CSC_{\pi/2}|C\end{aligned}$$

where $|$ means reverse direction. There are 48 different words in $\ell^\pm$, $s^\pm$, $r^\pm$ when C in (1.1) or (1.2) is replaced by $C = \ell$ or $C = r$. Some of these 48 words have 2 formulas for an actual path of its word type. There are at most 68 formulas in any given case. Table 1 summarizes the 48 words and 68 formulas.

Dubin's proof is different from ours. He shows that there are geodesics for any endpoint conditions, i.e. the infimum is achieved, and then proves the lemma that any geodesic of length less than $\pi/8$ must be a CSC. It then follows easily that every geodesic must be a finite word in C and S, and then using a series of special arguments reduces all finite words to CCC or CSC.

We use advanced calculus to deduce our result from Dubin's theorem and then in §7, we use the same general method to outline a separate proof of Dubin's theorem itself. In fact, we do not see how

TABLE 1

This table lists the 48 words in our sufficient set, together with their shorthand names as used in (1.1) and (1.2). The last column gives the segment length formulas for the given word.

explicit	(1.1) form	(1.2) form	Section 8 formula
$l^+r^-l^+$	$C^+C^-C^+$	$C \mid C \mid C$	(8.3), two roots
$l^-r^+l^-$	$C^-C^+C^-$	$C \mid C \mid C$	(8.3), two roots
$r^+l^-r^+$	$C^+C^-C^+$	$C \mid C \mid C$	(8.3), two roots
$r^-l^+r^-$	$C^-C^+C^-$	$C \mid C \mid C$	(8.3), two roots
$l^+r^-l^-$	$C^+C^-C^-$	$C \mid CC$	(8.4), two roots
$l^-r^+l^+$	$C^-C^+C^+$	$C \mid CC$	(8.4), two roots
$r^+l^-r^-$	$C^+C^-C^-$	$C \mid CC$	(8.4), two roots
$r^-l^+r^+$	$C^-C^+C^+$	$C \mid CC$	(8.4), two roots
$l^-r^-l^+$	$C^-C^-C^+$	$CC \mid C$	(8.4), two roots
$l^+r^+l^-$	$C^+C^+C^-$	$CC \mid C$	(8.4), two roots
$r^-l^-r^+$	$C^-C^-C^+$	$CC \mid C$	(8.4), two roots
$r^+l^+r^-$	$C^+C^+C^-$	$CC \mid C$	(8.4), two roots
$l^+r^+_ul^-_ur^-$	$C^+C^+_uC^-_uC^-$	$CC_u \mid C_uC$	(8.7), two roots
$l^-r^-_ul^+_ur^+$	$C^-C^-_uC^+_uC^+$	$CC_u \mid C_uC$	(8.7), two roots
$r^+l^+_ur^-_ul^-$	$C^+C^+_uC^-_uC^-$	$CC_u \mid C_uC$	(8.7), two roots
$r^-l^-_ur^+_ul^+$	$C^-C^-_uC^+_uC^+$	$CC_u \mid C_uC$	(8.7), two roots
$l^+r^-_ul^-_ur^+$	$C^+C^-_uC^-_uC^+$	$C \mid C_uC_u \mid C$	(8.8)
$l^-r^+_ul^+_ur^-$	$C^-C^+_uC^+_uC^-$	$C \mid C_uC_u \mid C$	(8.8)
$r^+l^-_ur^-_ul^+$	$C^+C^-_uC^-_uC^+$	$C \mid C_uC_u \mid C$	(8.8)
$r^-l^+_ur^+_ul^-$	$C^-C^+_uC^+_uC^-$	$C \mid C_uC_u \mid C$	(8.8)
$l^+r^-_{\pi/2}s^-l^-$	$C^+C^-_{\pi/2}S^-C^-$	$C \mid C_{\pi/2}SC$	(8.9)
$l^-r^+_{\pi/2}s^+l^+$	$C^-C^+_{\pi/2}S^+C^+$	$C \mid C_{\pi/2}SC$	(8.9)
$r^+l^-_{\pi/2}s^-r^-$	$C^+C^-_{\pi/2}S^-C^-$	$C \mid C_{\pi/2}SC$	(8.9)
$r^-l^+_{\pi/2}s^+r^+$	$C^-C^+_{\pi/2}S^+C^+$	$C \mid C_{\pi/2}SC$	(8.9)
$l^-s^-r^-_{\pi/2}l^+$	$C^-S^-C^-_{\pi/2}C^+$	$CSC_{\pi/2} \mid C$	(8.9)
$l^+s^+r^+_{\pi/2}l^-$	$C^+S^+C^+_{\pi/2}C^-$	$CSC_{\pi/2} \mid C$	(8.9)
$r^-s^-l^-_{\pi/2}r^+$	$C^-S^-C^-_{\pi/2}C^+$	$CSC_{\pi/2} \mid C$	(8.9)
$r^+s^+l^+_{\pi/2}r^-$	$C^+S^+C^+_{\pi/2}C^-$	$CSC_{\pi/2} \mid C$	(8.9)
$l^+r^-_{\pi/2}s^-r^-$	$C^+C^-_{\pi/2}S^-C^-$	$C \mid C_{\pi/2}SC$	(8.10)
$l^-r^+_{\pi/2}s^+r^+$	$C^-C^+_{\pi/2}S^+C^+$	$C \mid C_{\pi/2}SC$	(8.10)
$r^+l^-_{\pi/2}s^-l^-$	$C^+C^-_{\pi/2}S^-C^-$	$C \mid C_{\pi/2}SC$	(8.10)
$r^-l^+_{\pi/2}s^+l^+$	$C^-C^+_{\pi/2}S^+C^+$	$C \mid C_{\pi/2}SC$	(8.10)
$r^-s^-r^-_{\pi/2}l^+$	$C^-S^-C^-_{\pi/2}C^+$	$CSC_{\pi/2} \mid C$	(8.10)
$r^+s^+r^+_{\pi/2}l^-$	$C^+S^+C^+_{\pi/2}C^-$	$CSC_{\pi/2} \mid C$	(8.10)
$l^-s^-l^-_{\pi/2}r^+$	$C^-S^-C^-_{\pi/2}C^+$	$CSC_{\pi/2} \mid C$	(8.10)
$l^+s^+l^+_{\pi/2}r^-$	$C^+S^+C^+_{\pi/2}C^-$	$CSC_{\pi/2} \mid C$	(8.10)
$l^+s^+r^+$	$C^+S^+C^+$	CSC	(8.2)
$l^-s^-r^-$	$C^-S^-C^-$	CSC	(8.2)
$r^+s^+l^+$	$C^+S^+C^+$	CSC	(8.2)
$r^-s^-l^-$	$C^-S^-C^-$	CSC	(8.2)
$l^+s^+l^+$	$C^+S^+C^+$	CSC	(8.1)
$l^-s^-l^-$	$C^-S^-C^-$	CSC	(8.1)
$r^+s^+r^+$	$C^+S^+C^+$	CSC	(8.1)
$r^-s^-r^-$	$C^-S^-C^-$	CSC	(8.1)
$l^+r^-_{\pi/2}s^-l^-_{\pi/2}r^+$	$C^+C^-_{\pi/2}S^-C^-_{\pi/2}C^+$	$C \mid C_{\pi/2}SC_{\pi/2} \mid C$	(8.11), two roots
$l^-r^+_{\pi/2}s^+l^+_{\pi/2}r^-$	$C^-C^+_{\pi/2}S^+C^+_{\pi/2}C^-$	$C \mid C_{\pi/2}SC_{\pi/2} \mid C$	(8.11), two roots
$r^+l^-_{\pi/2}s^-r^-_{\pi/2}l^+$	$C^+C^-_{\pi/2}S^-C^-_{\pi/2}C^+$	$C \mid C_{\pi/2}SC_{\pi/2} \mid C$	(8.11), two roots
$r^-l^+_{\pi/2}s^+r^+_{\pi/2}l^-$	$C^-C^+_{\pi/2}S^+C^+_{\pi/2}C^-$	$C \mid C_{\pi/2}SC_{\pi/2} \mid C$	(8.11), two roots

to formulate a result analogous to his short geodesic lemma for the reverse problem, so his methods do not seem to work in a straightforward way. We show instead that any curve can be approximated by a word in $C^{\pm}$ and $S^{\pm}$. Then we show that any word in $C^{\pm}$, $S^{\pm}$ can be reduced to one in at most 5 letters without increasing its length. Finally we reduce to a word on the list by using ideas similar to the last part of Dubin's proof.

Although we give a rigorous proof of our assertions, we used a computer to empirically determine a sufficient list of words as follows: Given a set W of words, we tested W for insufficiency by generating the endpoint conditions randomly and first finding the best path in W. If a shorter path can be found by concatenating two paths in W, then W is insufficient. Using this method and pruning, we eventually arrived at and convinced ourselves that we had a minimal sufficient set.

Once we had guessed at W, we used the computer again to help do the extensive algebra in the large number of cases involved to verify that a rigorous proof could be given by the method outlined above. Finally, we found that the proof could be simplified (§2), so that it can easily be followed by an ordinary human without a computer to check the details. But we think that we could never have found the right set of words without using a computer.

In §7 we outline a proof of Dubin's theorem by our method.

In §8 we give a list of formulas to compute the lengths of each of the 68 actual path-solutions for each of the 48 word types suitable for algorithmic implementation.

2. Admissible paths. For us, the state of a car at a given instant t, where t is arclength, is completely specified by its position $(x(t), y(t))$ in the plane. An *admissible path* or *curve* is a function $\gamma(t) = (x(t), y(t), \phi(t))$ for which we can find measurable functions ε and η for which

$$
\begin{aligned}
(2.1)\qquad x(t) &= x(0) + \int_0^t \varepsilon(\tau)\cos\phi(\tau)\,d\tau, \\
y(t) &= y(0) + \int_0^t \varepsilon(\tau)\sin\phi(\tau)\,d\tau, \quad \text{where} \\
\phi(t) &= \phi(0) + \int_0^t \eta(\tau)\,d\tau
\end{aligned}
$$

and where $\varepsilon(\tau) = \pm 1$ and $|\eta(\tau)| \leq 1$ for each τ. In words, a car can move only forwards or backwards in its own direction $\phi(\tau)$ with speed $\sqrt{\dot{x}^2 + \dot{y}^2}(\tau) = |\varepsilon(\tau)| = 1$ and cannot change its direction $\phi(\tau)$ faster than one radian per time unit so that its turning radius is at least one or the curvature of its path is at most one (since the curvature is the reciprocal of the turning radius).

Note that we must allow infinite acceleration at a cusp where $\varepsilon(t)$ changes sign instantly. The problem where the acceleration must satisfy $\ddot{x} + \ddot{y} \leq a < \infty$ is more difficult and is not treated here. However, for slowly moving vehicles, such as carts, this seems like a reasonable compromise to achieve tractability. How does one characterize the class of paths (x_t, y_t) which satisfy (2.1)? Any path with piecewise constant circles of radius ≥ 1 and/or line segments suffices, but one can take $\{t : \varepsilon(t) = 1\}$ to be an arbitrary measurable set. A somewhat complicated condition on (x_t, y_t), $0 \leq t \leq T$, is: It is assumed that t is length along the path, so that $\dot{x}_t^2 + \dot{y}_t^2 \equiv 1$. Then it is necessary and sufficient that for $\phi(t) = \tan^{-1}(\dot{y}_t/\dot{x}_t)$ we have

$$|\phi(t+s) - \phi(t)| \leq s \quad \text{for } 0 \leq t \leq t+s \leq T. \tag{2.2}$$

This is an immediate consequence of the fact that a Lip 1 function ϕ as in (2.2) is the integral of its derivative η as in (2.1). Finally, $\varepsilon(\tau)$ is automatically uniquely defined by (2.1) since $\dot{x}_t^2 + \dot{y}_t^2 = 1$.

The *track* of an admissible curve $\gamma(t) = (x(t), y(t), \phi(t))$ is $\overline{\gamma}(t) = (x(t), y(t))$. By differentiating (2.1) we get

$$\begin{aligned} \dot{x}(t) &= \varepsilon(t)\cos\phi(t), \\ \dot{y}(t) &= \varepsilon(t)\sin\phi(t), \\ \dot{\phi}(t) &= \eta(t), \end{aligned} \tag{2.3}$$

so both $\gamma(t)$ and $\overline{\gamma}(t)$ are rectifiable. If g is admissible with $g(t_0) = a$ and $g(t_1) = b$ with $t_0 < t_1$ we call its restriction $\gamma = g|_{[t_0, t_1]}$ to the domain $t_0 \leq t \leq t_1$ an *admissible path leading from* a *to* b *and define its length* $L(\gamma) = t_1 - t_0$. Our problem in short: given an arbitrary a and b in $\mathbb{R}^3$ find an admissible γ minimizing $L(\gamma)$.

3. Summary of results. A *word* is a finite string in the letters C, S, $|$ and with some abuse of notation is also thought of as a path or as a set of paths. Each path in $C|CS$, for example, starts somewhere, goes along a circle of radius 1 for some distance ≥ 0, then has a cusp, then goes along the other circle tangent to the first circle at the