Shortest Path Solution by Epitrochoid Machine

Mark Schwartz (mdschwar@cc.owu.edu), Ohio Wesleyan University, Delaware, OH 43015, and Darryl Adams (daadams@usa.capgemini.com), Cap Gemini America, Cincinnati, OH 45242

Introduction

Let A and B be distinct points on the same side of a straight line L in the xy-plane. The shortest-path problem in calculus is to find the point U on L such that the path length $\overline{AU}+\overline{UB}$ is minimal. Let $A=(a,b)$, $B=(c,d)$. Formally, the problem is to minimize the function $F=\sqrt{(a-u)^2+(b-v)^2}+\sqrt{(c-u)^2+(d-v)^2}$, where (u,v) is any point on L. The usual minimization technique (using calculus) works, but a simple geometric solution is obtained by reflecting B through L to the point D. The situation is shown in Figure 1. The line segment $\overline{DA}$ intersects L in the

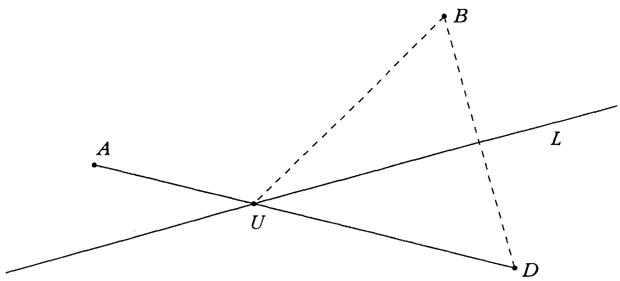

Figure 1

desired point U. The aim of this paper is to present a similar reflection-style solution, where L is replaced by the unit circle C. As explained later, equipped with cardboard and/or a symbolic and graphing package, students and instructors may find this problem appealing for an independent project or for use in a computer lab setting.

Main problem. Let C be the unit circle with parametric equations $r(t)=(u,v)$, where $u=\cos t$, $v=\sin t, 0\le t\le 2\pi$. We will generally write $U=(u,v)$. Let $A=(a,b)$ and $B=(c,d)$ be points outside of C, and assume B is "visible" from A. That is, the line segment $\overline{AB}$ does not intersect C (it is permissible if $\overline{AB}$ is tangent to C, but this situation has the trivial straight line solution). The problem is to find the shortest path from A to C to B. Throughout the discussion we assume, without loss of generality, $B=(c,0)$, with $c>1$, and A is in the upper half-plane. Figure 2 shows the situation.

We seek to minimize the function

$F(t)=\sqrt{(a-u)^2+(b-v)^2}+\sqrt{(c-u)^2+v^2}$. Differentiating with respect to t, equating to 0, and rearranging terms, we get

$$\frac{(a-u)u'+(b-v)v'}{\sqrt{(a-u)^2+(b-v)^2}}=\frac{(u-c)u'+vv'}{\sqrt{(c-u)^2+v^2}}. \tag{1}$$

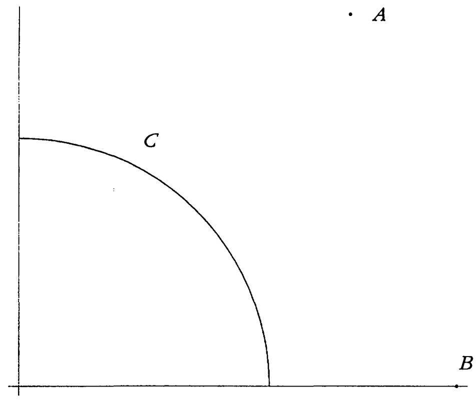

Figure 2

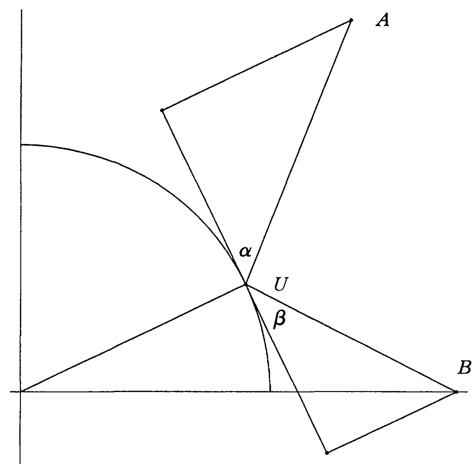

Figure 3

Consider Figure 3 where an arbitrary value of t is depicted. Let α and β be the angles of incidence and reflection, respectively. The numerator on the left side of (1) is the scalar projection of the vector $\overrightarrow{UA}$ onto the unit tangent vector $(u', v') = (-\sin t, \cos t)$ to C at U. Thus, the fraction on the left side of (1) is $\cos\alpha$. Similarly, the fraction on the right side of (1) is $\cos\beta$. It follows that $\alpha = \beta$, as some may have expected.

Orthotomic. Solving (1) for t is impossible in general. Instead, we develop a geometric solution obtained by reflecting the point B through the circle C. To motivate this, consider the inverse problem. Suppose U is a fixed point on C. For B fixed, we determine the locus of points A for which U is the minimizing point in the shortest path problem. Given U and B, the angle of reflection β is determined. Hence, the desired locus is the ray with initial point at U forming angle of incidence α equal to β. Referring to Figure 4, a convenient way to construct this is to let D be the reflection of B in the tangent line to C at U. Then the ray $\overrightarrow{DU}$ contains the required locus. That the correct locus has been constructed follows from the

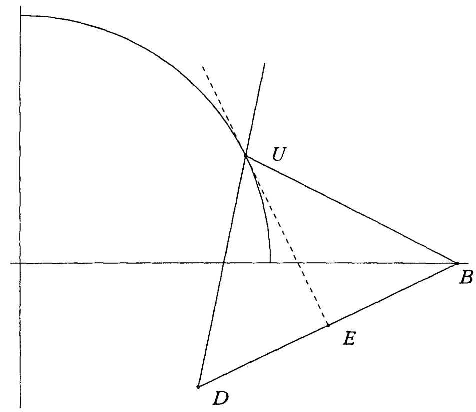

Figure 4

congruence of $\triangle DEU$ and $\triangle BEU$. If U is allowed to vary, we obtain a locus of reflection points D, the equation of which is determined as follows. First, the point E in the figure is obtained from the vector projection of the vector $\overrightarrow{BU}$ onto the unit normal vector $\vec{N} = (-u, -v)$ to C at U. In fact, $\overrightarrow{BE} = (\overrightarrow{BU} \cdot \vec{N})\vec{N}$. Thus, $\overrightarrow{BD}$, having the same direction as $\overrightarrow{BE}$ but twice the magnitude, is given by

$$\overrightarrow{BD} = 2\left(\overrightarrow{BU} \cdot \vec{N}\right)\vec{N}. \tag{2}$$

The locus of reflection points D is a curve, called the *orthotomic* of C (relative to fixed B). The orthotomic is well known in the study of curves. The reader may obtain additional background by consulting **[1]**, **[3]**, or **[4]**. For our purposes, the orthotomic is the generalized version of the reflection used in the linear shortest path problem. Recalling $U = (u, v) = (\cos t, \sin t)$ and $B = (c, 0)$, (2) gives

$$D - (c, 0) = 2[(u - c, v) \cdot (-u, -v)](-u, -v) = 2(1 - cu)(u, v). \tag{3}$$

The result is shown in Figure 5, where U has been allowed to vary over the entire circle. In fact, the orthotomic of C (relative to a point outside of C) is a familiar curve. As t varies, the right side of (3) becomes the polar coordinate graph of

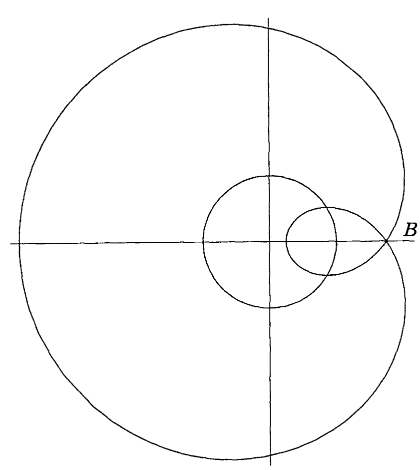

Figure 5

$f(\theta) = 2(\ 1 - c\cos\theta)$, which we recognize as a limaçon (rotated, for general B) with a loop.

Solution. Returning to the main problem, we are given points A and B, and seek the point U on C to minimize the path length $\overrightarrow{AU} + \overrightarrow{UB}$. First, reflect B through C to get the orthotomic. The question is how to connect the orthotomic to A. In the linear case where the orthotomic degenerates to a single point, the situation is clear. In the present case, let D be the reflection of B through C at any point U. As described above (inverse problem) the minimizing U is that point for which the ray $\overrightarrow{DU}$ passes through A. To find the minimizing U, we develop a simple mechanical device which does the job. First, each point on the orthotomic may be constructed geometrically as follows. Refer to Figure 6 and note that c is the distance from B to

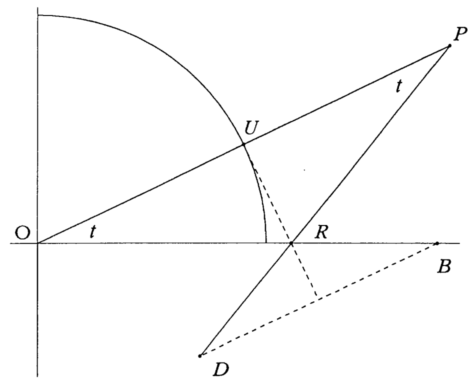

Figure 6

the origin O. Fix a value of t and let $U = (\cos t, \sin t)$. We seek the corresponding point D on the orthotomic. Let $P = 2(\cos t, \sin t)$ which is the point at distance 2 from O along the ray $\overrightarrow{OU}$. Construct angle t counterclockwise from the ray $\overrightarrow{PU}$, giving a new ray $\overrightarrow{PR}$. The point D is at distance c from P along $\overrightarrow{PR}$. This construction is correct, since $\triangle ORP$ and $\triangle DBR$ are similar isosceles triangles. They have a common perpendicular bisector to their bases which passes through their common vertex and is tangent to C at U. Hence, D is the reflection of B through the tangent line to C at U. This construction has an easy physical realization, which serves as the model for our mechanical solution. Using the terminology of **[2]** and **[4]**, the orthotomic under study is an *epitrochoid*, and is obtained by attaching a tracing point at distance c from the center of another unit circle and rolling the structure along the circumference of C. In Figure 6, one can easily imagine the rolling circle centered at P. Figure 7 illustrates a mechanical device which can be

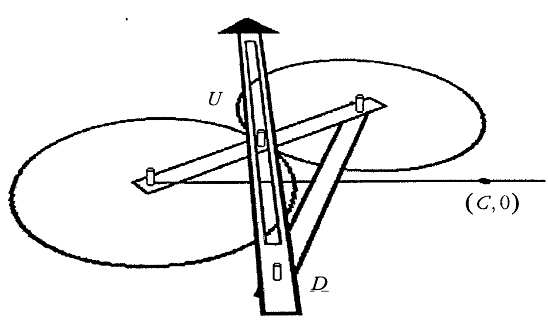

Figure 7

used to solve our shortest-path problem. The device consists of two unit circles, one fixed and one rolling, connected by a coupling of length 2 through their centers. A pin U is placed at the midpoint of the coupling. A tracing point D is at the end of an arm of length c attached to the center of the rolling circle. Finally, a pointer is hinged to this arm at D. The pointer is grooved and straddles the pin U. At all times, the pin U is located at the contact point of the two circles, and the pointer lies on the ray $\overrightarrow{DU}$. Figure 8 shows the solution of the shortest path problem when

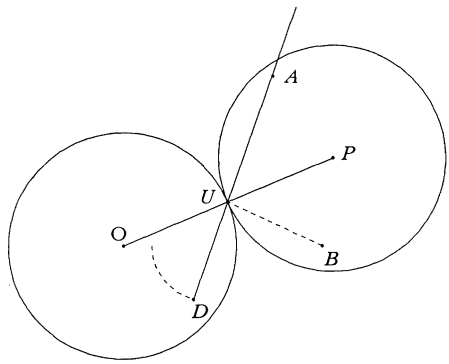

Figure 8

$B = (c, 0)$. If the rolling circle is initially placed in contact with C at the point $(1, 0)$, then the tracing point D is at $(2 - c, 0)$ and the pointer coincides with the x-axis pointing to the right. In Figure 8, the rolling circle has moved through an angle t until the pointer is over A. The location of the pin $U = (\cos t, \sin t)$ is the minimizing point we seek.

The authors constructed the machine in Figure 7 out of cardboard, and programmed a computer simulation. Perhaps ironically in this day of teaching and learning calculus with its concomitant use of technology, we found the cardboard machine particularly rewarding to construct and experiment with. Neither project was difficult, and it is anticipated that one or the other (or both) would be appreciated by calculus students working on this problem.

References

1. J. W. Bruce and P. J. Giblin, *Curves and Singularities*, 2nd ed., Cambridge University Press, 1992.
2. Leon M. Hall, Trochoids, roses, and thorns—beyond the Spirograph, This JOURNAL 23 (1992) 20–35.
3. I. R. Porteous, *Geometric Differentiation for the intelligence of curves and surfaces*, Cambridge University Press, 1994.
4. Robert C. Yates, *Curves and Their Properties*, reprint of 1952 edition, National Council of Teachers of Mathematics, 1974.