1 Introduction

Motivated by questions of unfolding polyhedra, Joseph O’Rourke asked at the 2007 Canadian Conference in Computational Geometry (CCCG), “For a ball rolling on a plane but not twisting, what is the length of the shortest path that will return the ball to its starting point, but oriented upside down?” The “no twist” condition, which would be satisfied naturally by a racquetball on a rubberized surface, is expressed mathematically in the next two sections.

Hammersley [1] posed a variant of this problem in optimal control: for a unit sphere lying on the plane at \((x_0, y_0)\) and having initial orientation \(C_0\), determine the shortest path for it to roll without twisting that brings it to point \((x_1, y_1)\) with orientation \(C_1\). Although the solution to Hammersley’s problem does not in general have a closed form, Arthurs and Walsh [2] have given an expression as a boundary-value problem with 10 coupled partial differential equations, and this expression reveals some curvature properties of the optimal path.

In this note we focus on O’Rourke’s variant, for both polygonal and smooth paths. In Section 2, we use numerical optimization in MATLAB to construct examples of optimal polygonal paths with \(k\) line segments in the plane, which form \(k\) arcs of great circles on the sphere without twisting. In Section 3, we sketch how the PDEs of Arthurs and Walsh give curvature properties of the optimal continuous curve, depicted in Figure 1.

2 Polygonal curve results

We first consider rolling polygonal paths in the plane, which corresponds to following arcs of great circles on the sphere. In this case, the no twist condition is simply that the angles at vertices are the same on the plane and on the sphere. Notice that rolling on an equilateral triangle with side length \(\pi\) upends a sphere in \(3\pi\).

We define a configuration of the plane and sphere system by \(\rho = (p, q)\), where \(p\) is the point on the plane that contacts the sphere and \(q\) is the orientation of the sphere, which we can represent by a quaternion.\(^1\)

If a unit sphere rolls along the polar vector \(v = (r, \theta)\) in the plane, then it rotates counterclockwise about \(v^\perp = (-\sin(\theta), \cos(\theta), 0)\) by an angle of \(r\), giving quaternion \(q_{r, \theta} = \cos(r/2) + \sin(r/2) (\sin(\theta) i + \cos(\theta) j)\). Thus, from configuration \((p, q)\) the sphere would move to \((p + r (\cos \theta, \sin \theta), q_{r, \theta} q)\).

Now, we wish to find a shortest \(n\)-segment path that returns to the origin in the plane, and goes from the south to north pole on the sphere. Thus, we want \(n\) vectors that sum to zero (so the first \(n-1\) determine the last vector) and that minimize total length plus a penalty \(\lambda \gg 0\) times the distance that the final contact ends from the north pole. Using MATLAB’s \texttt{fminsearch} command, which implements a Nelder-Mead simplex method for minimization [3], we find the examples of Table 1. The paths on the plane appear to converge to the teardrop shape of Figure 1 with length \(< 2.44\pi\).

\(^1\)Any vector \(v = (x, y, z)\) has a representation as a pure imaginary quaternion \(v = (x i + y j + z k)\), where \(i, j,\) and \(k\) are the imaginaries, which multiply like unit coordinate vectors: e.g., \(i^2 = 1\) and \(ij = k = -ji\). The rotation clockwise by angle \(\theta\) around \(v\) can be represented by the unit quaternion \(q = \cos(\theta/2) + \sin(\theta/2) v/\|v\|\); the actual rotation of another vector \(u\) is \(u q u^* q^*\), where the conjugate \(q^*\) is obtained by negating the imaginary terms of \(q\). Rotation by \(q_1\) then \(q_2\) is the product \(q_2 q_1^*\), since \(q_2(q_1 u q_1^*) q_2^* = (q_2 q_1) u (q_2 q_1)^*\).
3 Connection with Optimal Control

The rolling sphere problem can also be viewed as a problem in optimal control, where we need to choose a control function \( h(t) \) over time that will steer the sphere/plane system from its initial to its final configuration at minimum cost. We use the configuration \( \rho = (p, q) \) defined in the previous section. Note that the configuration \( \rho \) has 6 scalar variables, but five degrees of freedom since \( \|q\| = 1 \).

We take the unit vector along the sphere’s axis of rotation as the control. We will rotate the sphere at unit velocity, so the parameter for time, \( t \), can be replaced by path length \( \ell \). Since the sphere rolls without sliding or twisting, we obtain derivative conditions:

\[
\dot{p} = hk \\
\dot{q} = \frac{1}{2}hq
\]

The trajectories in the configuration space which satisfy these conditions are valid trajectories of the system. The control functions that give valid trajectories from start to end configurations are feasible controls. The feasible controls with minimum cost are known as optimal controls.

As in the previous section, we can use Lagrange multipliers, \( \lambda \) to combine the conditions into the minimization. Then, the Pontryagin’s Maximum principle [4] gives a necessary condition for optimal controls. The basic idea is that if \( c^*(\ell) \) is an optimal path between \( \rho_{\text{start}} \) and \( \rho_{\text{end}} \), then for any point \( c^*(\ell_0) \) on this path, the path from \( \rho_{\text{start}} \) to \( c^*(\ell_0) \) and from \( c^*(\ell_0) \) to \( \rho_{\text{end}} \) are also optimal. Let \( X \) be the configuration space of the system. Thus every path, \( c \), from \( \rho_{\text{start}} \) to \( \rho_{\text{end}} \) is a path in \( X \). Associated with every point \( c(\ell) \) on \( c \) is the length of the path traced on the plane, \( \ell \). We adjoin this cost to the configuration of the system to get the new state \( (\ell, \rho) \in \mathbb{R} \times X \), and use tangent conditions to link the values of \( \dot{\lambda} \)s and \( \dot{\rho} \)s.

This leads to defining the Hamiltonian \( H = \lambda^T f(\rho, h) \), where \( f(\rho, h) = (\ell, \rho) \) is the derivative of the lifted configuration. The tangent conditions are:

\[
\dot{\rho} = \frac{\partial H}{\partial \lambda} \\
\dot{\lambda} = -\frac{\partial H}{\partial \rho}
\]  

Pontryagin’s principle implies that at optimal control, \( H \) is maximum with respect to the control and equal to zero.

The result is a system of coupled differential equations where the boundary values of \( \rho \) are determined by the start and end configurations. Arthurs and Walsh [2] related the boundary values of \( \lambda \)s to curvature of the path. Their work demonstrates that the curvature is linearly related to the \( x \) coordinate for the curve of Figure 1, something that we first conjectured by observing the discretized curves in the previous section. Approximate solution of these differential equations give the same teardrop shape as observed in the previous section. The full paper presents the details.

Acknowledgment This work was partially supported by a grant from the National Science Foundation.

References


