Visualizing Quaternion Rotation

JOHN C. HART Washington State University

GEORGE K. FRANCIS University of Illinois, Urbana and LOUIS H. KAUFFMAN University of Illinois, Chicago

Quaternions play a vital role in the representation of rotations in computer graphics, primarily for animation and user interfaces. Unfortunately, quaternion rotation is often left as an advanced topic in computer graphics education due to difficulties in portraying the fourdimensional space of the quaternions. One tool for overcoming these obstacles is the quaternion demonstrator, a physical visual aid consisting primarily of a belt. Every quaternion used to specify a rotation can be represented by fixing one end of the belt and rotating the other. Multiplication of quaternions is demonstrated by the composition of rotations, and the resulting twists in the belt depict visually how quaternions interpolate rotation.

This article introduces to computer graphics the exponential notation that mathematicians have used to represent unit quaternions. Exponential notation combines the angle and axis of the rotation into a concise quaternion expression. This notation allows the article to present more clearly a mechanical quaternion demonstrator consisting of a ribbon and a tag, and develop a computer simulation suitable for interactive educational packages. Local deformations and the belt trick are used to minimize the ribbon's twisting and simulate a natural-appearing interactive quaternion demonstrator.

Categories and Subject Descriptors: I.3.5 [Computational Geometry and Object Modeling]: Hierarchy and Geometric Transformations; I.3.6 [Methodology and Techniques]: Graphics Data Structures and Data Types

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Deformation, education, orientation interpolation, quaternions, rotation, visualization

ACM Transactions on Graphics, Vol. 13, No. 3, July 1994, Pages 256-276.

J. C. Hart is supported in part by the NSF under a Research Initiation Award CCR-9309210. Equipment was provided by the Imaging Research Laboratory, which is supported by the NSF under grant CDA-9121675.

Authors' addresses: J. C. Hart, School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164-2752; email: hart@eecs.wsu.edu; G. K. Francis, Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801; email: gfrancis@math.uiuc.edu; L. H. Kauffman, Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60680; email: u10451@uicvm. bitnet.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

^{© 1994} ACM 0730-0301/94/0700-0256 \$03.50

1. INTRODUCTION

The method of specifying rotations and orientations of coordinate systems via unit quaternions was formally introduced to the computer graphics community by the publication of Shoemake [1985]. Quaternions were used in graphics programming informally mostly by geometers because Sir William Rowan Hamilton's [Hamilton 1866] beautiful invention is not regularly taught in college. Quaternions encode rotations by four real numbers (or two complex numbers), whereas the linear representation of these transformations as 3×3 matrices requires nine. Moreover, Hamilton impressed explicit geometrical meaning into every detail of his algebraic system, which guides intuition and facilitates implementation [Francis and Kauffman 1994].

Interpolating the quaternionic representation of a sequence of rotations is more natural than doing so for the familiar Euler angles, such as yaw, pitch, and roll. The quaternions occupy a smooth, seamless, isotropic space which is a generalization of the surface of a sphere. Thus, there is no need for special care to avoid singularities, such as gimbal lock, where two rotation axes collapse into one and thus make the interpolation irreversible.

Bezier curves were used in Shoemake [1985] to spline the quaternions representing rotations, while Barr et al. [1992] used energy-minimizing curves for demonstrably smoother motions. Quaternions provide an easy mechanism for specifying an arbitrary rotation about an arbitrary axis. This has long been exploited in keyboard user interfaces, and most recently for specifying 3-dimensional rotations with a 2-dimensional mouse [Hanson 1992; Shoemake 1992].

1.1 Overview

This article builds on previous work in quaternion rotation to derive an implementation of the quaternion demonstrator. The first half of the article summarizes various recent works on the quaternions. Section 2 reviews the quaternion representation of three-dimensional rotation, based on Shoemake [1985] and Francis and Kauffman [1994], and describes the quaternion demonstrator, as devised originally in Kauffman [1987; 1991]. Section 3 describes the belt trick, summarizing Kauffman [1991] and Francis and Kauffman [1994] and explaining the mathematics behind the animation "Air on the Dirac Strings" [Sandin et al. 1993].

These sections form the basis for this article's original contribution, found in the second half of the article. Beginning with Section 4, techniques from differential geometry model the quaternion demonstrator, regulating the twists and motions of the belt. Section 5 describes the resulting implementation and outlines directions for further research.

1.2 Background

An object \mathscr{O} is assumed to be defined with respect to some canonical coordinate frame. The *orientation* of \mathscr{O} is represented by a rotation that takes the object from its canonical coordinate frame to its current state.

This article relies heavily on the deformation techniques developed in Barr [1984]. We use globally and locally specified deformations. A globally specified deformation alters explicitly the positions of points in an object whereas a *locally specified deformation* affects the tangent space of an object, and new positions result only after an integration over the deformed tangent space.

2. THE QUATERNIONS

The four-dimensional space, \mathbb{H} , of quaternions is spanned by the real axis, and three further orthogonal axes, spanned by vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$, called the *principal imaginaries*, which obey Hamilton's rules

$$i^2 = j^2 = k^2 = ijk = -1.$$
 (1)

These imaginaries signify the three-dimensional vectors

$$i = (1, 0, 0),$$

 $j = (0, 1, 0),$
 $k = (0, 0, 1).$

Multiplication of these imaginaries resembles a cross product

$$ij = k, \quad jk = i, \quad ki = j,$$

$$ji = -k, \quad kj = -i, \quad ik = -j$$
(2)

and is clearly noncommutative. Quaternion multiplication causes rotation: multiplication on the right by **j** causes a 90 degree rotation in four-dimensional space, rotating the **i** axis into the **k** axis, and rotating the **k** axis into the $-\mathbf{i}$ axis. Quaternion multiplication differs from the cross product in that $\mathbf{i} = \mathbf{j}\mathbf{j} = \mathbf{k}\mathbf{k} = -1$ whereas $\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0$.

A quaternion $q = r + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ consists of a real part r and a pure part $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ [Hamilton 1866]. We will call quaternions with zero real part (r = 0) pure quaternions. Pure quaternions will also be simultaneously represented as a column vector

$$\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$
 (3)

Under this notation, the same symbol can simultaneously represent both a vector and a pure quaternion, depending on its context. For example,

$$\mathbf{v}^2 = -\mathbf{v} \cdot \mathbf{v} \tag{4}$$

because the LHS of (4) treats \mathbf{v} as a pure quaternion whereas the RHS of (4) treats \mathbf{v} as a vector.

Let $q_1 = a_1 + \mathbf{v}_1$ and $q_2 = a_2 + \mathbf{v}_2$ be two quaternions. Their sum is

$$q_1 + q_2 = (a_1 + a_2) + (\mathbf{v}_1 + \mathbf{v}_2),$$

and their product is

$$q_1q_2 = a_1a_2 - \mathbf{v}_1 \cdot \mathbf{v}_2 + a_1\mathbf{v}_2 + a_2\mathbf{v}_1 + \mathbf{v}_1 \times \mathbf{v}_2.$$

The quaternion $q = a + \mathbf{v}$ also decomposes into $a + b\mathbf{u}$ which resembles a complex number, where the imaginary \mathbf{u} is a unit three-vector

$$\mathbf{u} = \begin{pmatrix} x/b \\ y/b \\ z/b \end{pmatrix} = \frac{x}{\|\mathbf{v}\|}\mathbf{i} + \frac{y}{\|\mathbf{v}\|}\mathbf{j} + \frac{z}{\|\mathbf{v}\|}\mathbf{k},$$

such that $||\mathbf{u}|| = 1$, and x, y, z are the same coordinates used in (3). The pure unit-magnitude quaternion **u** resembles the imaginary **i** from the complex plane in that $\mathbf{u}^2 = -1$.

Let $q = a + b\mathbf{u}$ be a quaternion. Its conjugate is $\overline{q} = a - b\mathbf{u}$, and its magnitude is

$$\|q\| = q\bar{q} = \bar{q}q = \sqrt{a^2 + b^2}.$$
 (5)

2.1 Quaternion Rotation

Rotations in computer graphics are typically represented by quaternions of unit magnitude [Shoemake 1985], which we will call *unit quaternions*. The unit quaternions $\{q : ||q|| = 1\}$ form a hypersphere $\mathbb{S}^3 \subset \mathbb{H}$. In particular, for any unit quaternion $q \in \mathbb{S}^3$, (5) implies

$$q^{-1}=\bar{q}.$$
 (6)

In other words, to invert a unit quaternion, we simply negate its pure part. In Shoemake [1985], a rotation of θ about the axis **u** was represented as the unit quaternion

$$q = \cos\frac{1}{2}\theta + \sin\frac{1}{2}\theta \mathbf{u}$$

which matches the complex-like form of a quaternion $q = a + b\mathbf{u}$, where a is the real component and b the imaginary component along the new imaginary axis specified by the unit vector (pure quaternion) \mathbf{u} . The abbreviation $e^{\mathbf{i}\theta} = \cos \theta + \mathbf{i} \sin \theta$, borrowed from complex analysis, has a long history of use in the engineering sciences. We can similarly represent the aforementioned unit quaternion q more concisely using exponential notation as

$$q = e^{\frac{1}{2}\theta \mathbf{u}}.$$

In the same manner that engineers read the expression $e^{i\theta}$, the reader should likewise understand the notation $e^{(1/2)\theta \mathbf{u}}$ not as e to some imaginary power but simply as the quaternion that represents a rotation of θ about the axis \mathbf{u} .

Exponential notation was chosen to represent quaternion rotations herein to promote consistency between the converging fields of computer graphics and mathematics. Such quaternion exponential notation has a lengthy history in mathematics (e.g., as the so-called exponential map in different geometry [Spivak 1965]). However, since quaternion multiplication is noncommutative, likewise quaternion exponentiation does not in general follow the rules of real or complex exponentiation (e.g., $e^{(1/2)u_1u_1}e^{(1/2)u_2u_2} \neq$

 $e^{(1/2)(\theta_1 \mathbf{u}_1 + \theta_2 \mathbf{u}_2)}$). Quaternion exponentiation is formally defined in Francis and Kauffman [1994], along with a discussion of its properties illustrated by several examples.

Given a unit quaternion q that represents a rotation, the question remains of how to apply this rotation to an arbitrary vector (pure quaternion) $\mathbf{v} \in \mathbb{R}^3$. From Shoemake [1985], we find two results: a function R(q) that returns the 3×3 (nonhomogeneous) transformation matrix corresponding to the rotation represented by q, and a two-to-one correspondence between unit quaternions \mathbb{S}^3 and the space of all rotations $\mathbb{SO}(3)$ (the group of special-orthogonal 3×3 matrices). As a consequence of these two results, we have

$$R(q)\mathbf{v} \equiv q\,\mathbf{v}q^{-1}.\tag{7}$$

The LHS of (7) treats **v** as a column vector and yields a new column vector of the same length by left-multiplying the special-orthogonal matrix returned by the function R. The RHS of (7) treats **v** as a pure quaternion and yields a new pure quaternion of the same magnitude. In fact, both apply the rotation represented by the unit quaternion q to **v**. We denote rotations with the notation on the LHS of (7), but implement rotations more efficiently using the formula on the RHS of (7). (Since $q \in S^3$, q^{-1} simplifies to \bar{q} .)

Hence, the otherwise complicated procedure of rotating a vector about an arbitrary axis simplifies in our notation to

 $R(e^{rac{1}{2} heta \mathbf{u}})\mathbf{v}$

which rotates $\mathbf{v} \in \mathbb{R}^3$ about the $\mathbf{u} \in \mathbb{S}^2$ axis by an angle of θ [Francis and Kauffman 1994].

2.2 The Quaternion Demonstrator

The quaternion demonstrator [Kauffman 1987; 1991] is a mechanical unit quaternion multiplier. It consists primarily of a ribbon, called the *belt*, with one end fixed and the other end free. Fastened to the free end of the belt is a rectangle, called the *tag*. (An alternative demonstrator was discovered by Kauffman and E. Oshins [Kauffman 1991] that uses only one human arm.)

The orientations of the tag, along with the twists in the belt, represent the unit quaternions. The tag is inscribed with labels indicating the quaternion it currently represents. The top side of the tag is inscribed with a 1 and an upside-down \mathbf{j} . Its bottom side is inscribed with a \mathbf{k} and an upside-down \mathbf{i} , in the fashion suggested by Figure 1. (If you wish to follow along using your right arm, your right hand will be the tag. Your fingerprints are 1; your palm is \mathbf{j} ; your fingernails are \mathbf{k} ; and the back of your hand is \mathbf{i} . As a tag, your hand will represent the quaternion associated with the part of your hand facing up and toward the same direction you are facing.)

First we orient the quaternion coordinate frame such that the imaginaries form a right-handed coordinate system where **i** points right, **j** up, and **k** toward the viewer. The belt of the quaternion demonstrator is embedded in the **i** - **k** plane centered along the **k** axis with the fixed end at the origin and the tag at + **k**. The eyepoint is assumed to be somewhere in the positive **j** - **k**



Fig. 1. The quaternion demonstrator in the 1 state (top and bottom views). This figure uses the computer simulation of the quaternion demonstrator, which does not use a tag. Instead it superimposes the 1, \mathbf{i} , \mathbf{j} , and \mathbf{k} labels directly onto the tag end of the belt.

quadrant. (In your arm's coordinate system, the origin is your right shoulder; the **i** axis points toward your left shoulder; the **j** axis points up; and the **k** axis points in the direction you are facing.)

The canonical state of the demonstrator consists of the untwisted belt in this configuration. This state represents the value 1, and is shown in Figure 1. (Extending your arm out in front of you with your palm up puts your arm in the 1 state.)

Rotating the tag by 180 degrees with respect to the **i** axis puts the tag underneath the belt. This state represents the value **i** as can be read on the tag end of the belt (Figure 2 red). Call this rotation a "flip." (Keeping your arm and wrist straight, rotate 180 degrees about your shoulder's axis by dropping your arm to your waist and raise it back up behind you such that your palm is facing down. The back of your hand is facing up and toward the front, so your arm now represents the quaternion **i**.)

Reset the system to the canonical state 1. Rotating the tag by 180 degrees with respect to the \mathbf{j} axis puts the tag to the right of the fixed end of the belt. This state represents the value \mathbf{j} as can also be read on the tag end of the belt (Figure 2 green). Call this rotation a "spin." (From the 1 state, rotate your arm horizontally until you touch your chest with your fingertips. Your palm faces up and toward the front, and your arm now represents the quaternion \mathbf{j} .)

Resetting to 1 again and rotating the tag -180 degrees with respect to the **k** axis flips the tag over. This state represents the value **k**, as indicated by the upright **k** on the tag end of the belt (Figure 2 blue). Call this rotation a "turn." (From the 1 state, turn your wrist about your arm's axis until your



Fig. 2. The quaternion demonstrator multiplication by each of **i** (red), **j** (green), and **k** (blue).

palm is facing down. Your fingernails face up and to the front, and your arm now represents the quaternion \mathbf{k} .)

The fact that positive \mathbf{k} is obtained by a negative rotation in the righthanded coordinate system is an artifact of the belt-centered coordinate system. Centering a right-handed coordinate system on the tag with \mathbf{k} extending along the belt, \mathbf{j} up, and \mathbf{i} to the right yields rotations consistent with the right-handed coordinate system. (For the quaternion demonstrator's task of teaching quaternion multiplication the simplicity of the belt-centered coordinate system outweighs the familiarity of the right-handed rotation rules of the tag-centered coordinate system.)

The negative quaternion imaginaries are likewise produced by the opposite flips, twists, and turns, respectively. Although the labels on the tag have no signs, we can tell a positive imaginary from a negative imaginary by the direction of the twists in the belt. (The quaternion $-\mathbf{i}$ is represented from the 1 state by bending your arm at the elbow -180 degrees about your shoulder's axis, with your palm facing down as if to pat yourself on the back. The quaternion $-\mathbf{j}$ is represented by spinning your hand horizontally -180 degrees about the up axis into the position a waiter would use to hold a tray. The quaternion $-\mathbf{k}$ is physiologically impossible to represent in the arm coordinate system.)

In this system, multiplication is represented by the composition of corresponding rotations of the tag. For example, to demonstrate $\mathbf{ij} = \mathbf{k}$, we find that a flip followed by a spin is equivalent to a turn (after a little translation of the tag—moving the tag does not rotate it and does not change the state of the demonstrator). A reverse turn (multiplication by $-\mathbf{k}$) returns the system to its original state. (From the 1 state, flip 180 degrees around your shoulder's axis by dropping your arm to your waist and raising it up again behind you with your palm facing down into the \mathbf{i} state. Using your shoulder, spin about \mathbf{j}



Fig. 3. The belt trick equation: $\mathbf{k}^2 = (-\mathbf{k})^2$ proven visually using a quaternion demonstrator. The ends of the belt remain parallel through the transformation. Each step of the transformation represents the value - 1.

180 degrees by swinging your arm out and back to the front with your palm down. Your arm is now in front of you, palm down, representing \mathbf{k} .)

By (1), the -1 state is achieved by two flips (i^2), two spins (j^2), two turns (k^2), or a flip-spin-turn (**ijk**). Each of these operations returns the tag to its original orientation but puts a 360 degree twist in the belt. (From the 1 state, twist your wrist 360 degrees about the arm's axis until your palm faces up again (k^2). Your fingerprints are facing up and in front, but with a full 360 degree twist in your arm, which now represents -1.)

We can also create -1 by two reverse flips $((-i)^2)$ which also returns the tag to its original orientation but puts the opposite (-360 degrees) twist in the belt (more on this in the next section.)

3. THE BELT TRICK

In the quaternion demonstrator, -1 can be represented as \mathbf{k}^2 by two turns which return the tag to its original state but cause a 360 degree twist in the belt. Two reverse turns represents -1 also, as $(-\mathbf{k})^2$, but cause a -360 degree twist in the belt. The quaternion demonstrator has (at least) two distinct representations for -1.

These two representations are equivalent, however. Consider the demonstrator in the \mathbf{k}^2 state, after two turns. Without rotating the tag, move the tag in a positive 360 degree arc around the fixed end of the belt. This belt trick changes the 360 degree twist in the belt into a -360 degree twist, and proves the Belt Trick Equation shown in Figure 3.

(You can perform a variation of the belt trick called the plate trick, shown live in Sandin et al. [1993], with your arm by representing $j^4 = 1$. From the 1

state, multiply by **j** by touching your chest with your fingertips. Multiply by **j** again by moving your hand under your arm and back out, always keeping the palm up, into the -1 state. Multiply by **j** again by swinging your hand into the $-\mathbf{j}$ state, the waiter position, again always keeping your palm up. Multiply by **j** one final time back into the 1 state. You have rotated your hand 720 degrees without incurring the associated, and painful, 720 degree twist in your arm. Section 3.2 explains how this is possible.)

Discussions of the belt trick can be found in Misner et al. [1973], Bolker [1973], Kauffman [1987], Francis [1987], Kauffman [1991], and Francis and Kauffman [1994]. The rest of this section follows Kauffman [1991] and Francis and Kauffman [1994], developing a globally specified deformation for simulating the belt trick topologically, treating the belt more like a rubber band than a ribbon. Thus we find that the quaternions are neatly represented by a combination of rotational mechanism and appropriate topology.

3.1 The Belt Trick Deformation

Whereas Figure 3 demonstrates the belt trick using the quaternion demonstrator, where the belt is fixed at one end and free at the tag, the following discussion uses an equivalent but alternate construction. This new system consists of a belt connecting two concentric spheres, and is described quantitatively as follows.

Let

$$S(r) = r S^2 = {\mathbf{x} : ||\mathbf{x}|| = r}$$

specify a sphere of radius r centered at the origin. The hollow ball

$$\mathbb{HB} = \bigcup_{r=r_0}^1 S(r),$$

where $0 < r_0 < 1$ is the radius of the hollowed-out part, will serve to define the space in which the belt performs its trick.

The spheres S(1), $S(r_0)$ are called the outer sphere and the inner sphere, respectively. The outer sphere will represent the fixed end of the belt, and the inner sphere will represent the tag of the quaternion demonstrator. We can use intervals to represent a belt connecting the inner sphere to the outer sphere as

$$\mathscr{B} = (\llbracket -
ho,
ho
brace, 0, \llbracket 0, 1
brace) \cap \mathbb{HB}$$

where $\rho < r_0$ is half the width of the belt.

The belt-trick can now be illustrated as a global deformation $B_t : \mathbb{R}^3 \to \mathbb{R}^3$ parameterized by time $t \in [0, 1]$. At t = 0, $B_t(\mathscr{B})$ deforms the belt \mathscr{B} , giving it a 360 degree twist. As $t \to 1$, the belt will continuously deform into a belt with a -360 degree twist without rotating the inner or outer spheres—keeping the ends of the belt fixed.



Fig. 4. Unit quaternion paths corresponding to various stages of the belt trick deformation. The sphere S^2 is a slice of the hypersphere S^3 , consisting of unit quaternions whose **j** components is zero (similar to the figures at the end of Barr et al. [1992]).

The belt trick deformation shears the belt, rotating each increasingly larger spherical "shell" in \mathbb{HB} by angles of increasingly larger value about an axis that changes over time. The rotation angle function is

$$\theta(\mathbf{x}) = 2\pi \frac{1 - \|\mathbf{x}\|}{1 - r_0} \tag{8}$$

whereas the rotation axis function is

$$\mathbf{u}(t) = e^{\pi t \mathbf{j}} \mathbf{k} = (\sin \pi t, 0, \cos \pi t).$$
(9)

The function R, from (7), specifies rotation about an arbitrary axis, and is used to define the belt trick deformation

$$B_t(\mathbf{x}) = R(e^{\frac{1}{2}\theta(\mathbf{x})\mathbf{u}(t)})\mathbf{x}.$$
(10)

Consider the unit quaternions used for rotations in the belt trick deformation. These are plotted in Figure 4. At t = 0, as $\|\mathbf{x}\|$ ranges from ρ_0 to 1, the unit quaternions form an arc from 1 through \mathbf{k} to -1. As $t: 0 \rightarrow 1$ this arc rotates about \mathbb{S}^3 from one side to the other. At t = 1/2 this arc extends from 1 through \mathbf{i} to -1, and at t = 1 this arc extends from 1 through $-\mathbf{k}$ to -1. Since the arcs all begin at 1, the orientation of the inner sphere, and end at -1, the orientation of the outer sphere, the inner and outer spheres do not rotate during the belt trick.

3.2 The Unfurling Deformation

Composing the belt trick with a twist along the k axis produces a belt trick that takes the 720 degree twisted belt into an untwisted belt, and is given in Francis and Kauffman [1994] as

$$B_{s,t}(\mathbf{x}) = R(e^{\frac{1}{2}\theta(\mathbf{x})\mathbf{u}(t)})R(e^{\frac{1}{2}s^{2}\pi\mathbf{k}})\mathbf{x}$$
$$= R(e^{\frac{1}{2}\theta(\mathbf{x})\mathbf{u}(t)}e^{\frac{1}{2}s^{2}\pi\mathbf{k}})\mathbf{x}$$
(11)

where $s \in [-1, 1]$ is used to twist the belt. At s = -1 and t = 0, the deformation $B_{s,t}$ leaves the belt \mathscr{B} unchanged. As $s \to 1$ while t = 0 the deformation $B_{s,t}$ rotates the inner spheres -720 degrees, which puts a -720 degree twist in the belt and returns the sphere to its original orientation. Then, while s = 1 as $t \to 1$, the -720 degree twist unfurls around the inner sphere, returning the system to its original state with an untwisted belt. This process is illustrated by Figure 5. (The Appendix describes the ray-tracing technique used to render this figure.)

The unit quaternions used for rotations in the unfurling deformation are plotted in Figure 6. At s = 1, t = 0, as $||\mathbf{x}||$ ranges from ρ_0 to 1, the unit quaternions form a circle from 1 through \mathbf{k} through -1 through $-\mathbf{k}$ and back to 1. As $t: 0 \rightarrow 1$ this circle contracts around \mathbb{S}^3 , always intersecting 1. At t = 1/2 the circle extends from 1 through \mathbf{i} and back to 1, and at t = 1 circle degenerates to the point 1. Since the circles all begin and end at 1, as with the belt trick, the inner and outer spheres do not rotate during the unfurling.

4. SIMULATING THE QUATERNION DEMONSTRATOR

The belt trick and unfurling deformations are global deformations. They maintain the belt's volume (for the same reason the twist deformation [Barr 1984] preserves volume) but stretch the length of the belt like a rubber band such that the inner and outer spheres remain centered about the origin. The belt in the quaternion demonstrator maintains a constant length, but the tag is free to move about. This suggests that a local deformation should be used to simulate the belt in the quaternion demonstrator. Furthermore, this local deformation should minimize the twisting of the belt.

First, the orientation of the tag is represented by a quaternion q. Then a geodesic (a great arc on S^3) of unit quaternions interpolates the orientations along the belt from the orientation of the fixed end of the belt, 1, to the orientation of the tag, q. Finally, changes to this geodesic resulting from rotations of the tag must be carefully monitored to prevent drastic changes in the shape of the belt.

4.1 Simulating the Tag

In the physical quaternion demonstrator, the tag is a small rectangle attached to the free end of the belt. In the simulation, we consider the tag to be the edge of the free end of the belt.

Let \mathbf{v}_F represent the orientation of the edge of the belt's fixed end as a vector from one corner to the other. Let \mathbf{v}_T denote the orientation of the tag,



Fig. 5. The unfurling, using a global deformation to show how to remove a 720 degree twist from a ribbon without moving either end.



Fig. 6. Unit quaternion paths corresponding to various stages of the unfurling deformation

as the vector connecting the corresponding corners at the edge of the belt's free end. Let $q \in S^3$ be a unit quaternion denoting the state of the quaternion demonstrator. Then the orientation of the tag with respect to the orientation of the fixed end of the belt is given by

$$\mathbf{v}_T = R(q)\mathbf{v}_F.$$

Multiplication of q by an imaginary rotates the tag 180 degrees discretely. In the simulation, these rotations are performed incrementally and appear ACM Transactions on Graphics, Vol 13, No 3, July 1994 Fig. 7. The belt as the union of rotated line segments (left). The orientation of \mathbf{v}_F , the fixed edge of the belt, corresponds to one whereas the orientation of \mathbf{v}_T , the tag edge of the belt, corresponds to q. The geodesic on \mathbb{S}^3 (right) connects 1 to q, and specifies the orientations of the segments interpolating \mathbf{v}_F to \mathbf{v}_T .



continuous. With the physical quaternion demonstrator, multiplication by \mathbf{i} is performed by flipping the tag. In the simulation, multiplication by \mathbf{i} is accomplished by pressing and holding the "i" key and watching the tag slowly flip.

Let q_0 be a unit quaternion denoting the current orientation of the tag. Then incremental multiplication of the tag is simulated by

$$q_1 = q_0 e^{\epsilon \mathbf{u}},\tag{12}$$

where unit quaternion q_1 specifies the new tag orientation; **u** is one of **i**, **j**, or **k**, and ϵ is a small rotation angle. In our implementation, setting $\epsilon = 0.02$ radians resulted in a pleasing ribbon animation speed.

4.2 Simulating the Belt

In Section 3, the belt was sheared by a family of rotating concentric spheres. Here, the belt is best represented by a family of rotating line segments. As before, let \mathbf{v}_F represent the vector at the edge of the fixed end of the belt, and let \mathbf{v}_T represent the vector at the tag.

The orientation of the fixed end of the belt \mathbf{v}_F corresponds to the unit quaternion 1 whereas the orientation of the tag end of the belt \mathbf{v}_T corresponds to q. Let $\Gamma \subset \mathbb{S}^3$ be the geodesic connecting 1 to q. Then the belt consists of the union of line segments whose orientations interpolate the orientation of \mathbf{v}_F into the orientation of \mathbf{v}_T , specifically the orientations represented by points along the geodesic Γ , as shown in Figure 7.

We describe the local deformation of the belt by applying the rotations represented by the quaternions along the geodesic Γ to the tangent space of the belt. The deformed belt is then constructed as an initial-value problem by integrating the belt over these deformed tangent vectors.

First, decompose q into exponential form as

$$q = r + \mathbf{v},$$

$$\theta = 2\cos^{-1} r,$$

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}}{\sin\frac{1}{2}\theta},$$

$$q = e^{\frac{1}{2}\theta \mathbf{u}},$$



Fig. 8. The Frenet frame along the spine of the belt.

where \cos^{-1} always returns values in the range $(-\pi, \pi]$. This form reveals the rotation q represents. The geodesic Γ extending from 1 to q on \mathbb{S}^3 is parameterized by the function $\gamma(s) \in \Gamma$ for $s \in [0, 1]$ as

$$\gamma(s) = e^{\frac{1}{2}s\theta \mathbf{u}}.$$
 (13)

The unit quaternion function $\gamma(s)$ specifies the orientations the belt twists through on its path from its fixed end to the tag. (Quaternion exponentiation maps the line segment in \mathbb{R}^3 connecting the origin to $(1/2)\theta \mathbf{u}$ to the geodesic in \mathbb{S}^3 connecting 1 to q.)

Let $\mathbf{x}_0 = \mathbf{0}$ be the position, and $\mathbf{t}_0 = \mathbf{k}$, $\mathbf{b}_0 = \mathbf{i}$, and $\mathbf{n}_0 = \mathbf{j}$ be the Frenet frame (tangent \mathbf{t} , binormal \mathbf{b} , and normal \mathbf{n}) of the center of the fixed end of the belt, as in Figure 8.

The local description of points along the spine of the belt is given by the Frenet frame

$$\mathbf{t}(s) = R(\gamma(s))\mathbf{t}_0,$$

$$\mathbf{b}(s) = R(\gamma(s))\mathbf{b}_0,$$

$$\mathbf{n}(s) = R(\gamma(s))\mathbf{n}_0.$$

Integrating the tangent t(s) produces points along the spine of the belt

$$\mathbf{x}(s) = \mathbf{x}_0 + \int_0^s \mathbf{t}(\sigma) \, d\sigma.$$
 (14)

The belt is formed as a ruled surface consisting of line segments connecting the vertices

$$\mathbf{x}(s) \pm \rho \mathbf{b}(s),$$

where, as before, ρ is one-half the width of the belt. The twisting of the belt visualizes the interpolation of orientations of a line segment from its fixed end to the tag end.

Since the Frenet frame is just rotated by these functions, the length of the belt remains unchanged under the deformation. (By the way, if the locally specified deformation were not simply a rotation, then the normal vector

transformation rule [Barr 1984] would apply. In such a case, the transformed binormal would be found via the inverse transpose of the Jacobian matrix of R(), and the normal would be constructed from the cross product of the tangent vector with the binormal vector.)

4.3 Limiting Belt Velocity

The orientation interpolation geodesic on \mathbb{S}^3 is an arc extending from 1 (the north pole) to q. As q passes by -1 (the south pole) the geodesic generated by (13) will move from one side of \mathbb{S}^3 to the other (since the range of \cos^{-1} is $(-\pi, \pi]$).

This movement keeps the belt from accumulating unnecessary twists. Since the geodesic connecting 1 to q is the shortest path on \mathbb{S}^3 , belt tricks occur naturally as the tag is rotated to avoid twists of greater than 2π in the belt. This movement has one disadvantage in that certain small movements of qnear the south pole cause the geodesic to swing around quickly to the other side of \mathbb{S}^3 resulting in a belt trick that is too fast for the user to follow. In fact, if the tag rotates directly through -1, the geodesic snaps from one side of \mathbb{S}^3 to the other, causing an instantaneous belt trick (an instant reverse of the twist in the belt).

For example, turning the tag about the **k** axis from the initial 1 state produces eventually a 2π twist in the ribbon about the **k** axis. Turning the tag slightly causes an instantaneous belt trick, snapping the belt from a 2π twist to a -2π twist. Although both configurations are nearly equivalent, representing nearby quaternion values, their appearance to the user is quite different.

There are two remedies for handling instantaneous belt tricks. The first remedy senses when the tag orientation path crosses -1, or nearly misses it. When it does, this remedy assumes control of the demonstrator from the user and performs an explicit animated belt trick to remove the excess twist in the belt.

The second remedy capitalizes on numerical error to perform belt tricks automatically, as necessary. As (12) rotates the tag incrementally, small numerical errors will accumulate in the quaternion representation of the tag's orientation. In other words, turning the tag about the **k** axis will introduce slight rotations about the **i** and **j** axes as well. By the time the tag's turning has twisted the belt by 2π and more, these perturbations will cause the tag's orientation quaternion q to miss the south pole. The resulting geodesics will quickly swing across S^3 producing a belt trick, although possibly at a very fast rate.

We chose to implement the second remedy in the simulation of the quaternion demonstrator for its elegance and because it never assumes control of the demonstrator away from the user. This elegance may give the impression that the second remeby avoids belt tricks, which are an essential point of the quaternion demonstrator. To the contrary, belt tricks resulting from near misses of the south pole are indiscernable from the belt trick required to simulate the belt continuously through a direct hit. The dependence of this remedy on numerical noise affects its robustness in that a direct hit on the

south pole would lock up the quaternion demonstrator simulation. Such a direct hit is highly unlikely and has never occured in our experience. The only remaining task is to regulate the rate at which the quaternion demonstrator performs automatic belt tricks.

The speed of the automatic belt trick is regulated by controlling the speed at which the geodesic Γ flips around \mathbb{S}^3 , which, in turn, is controlled from the rate of rotation of the tag by regulating the ϵ in (12). The rest of this section is devoted to deriving the amount of regulation of ϵ necessary to control the speed of the automatic belt tricks when they occur.

Following Misner et al. [1973], we can describe a unit quaternion $r + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ in spherical coordinates with three angles α , ϕ , and θ as

 $x = \sin \alpha \sin \phi \cos \theta,$ $y = \sin \alpha \sin \phi \sin \theta,$ $z = \sin \alpha \cos \phi,$ $r = \cos \alpha.$

The inverse is computed as

$$\alpha = \cos^{-1} r,$$

$$\phi = \cos^{-1} \frac{z}{\sin \alpha},$$

$$\theta = \cos^{-1} \frac{x}{\sin \alpha \sin \theta}$$

$$= \sin^{-1} \frac{y}{\sin \alpha \sin \theta}$$

Also from Misner et al. [1973], the differential of geodesic length is given by

$$ds^{2} = d\alpha^{2} + \sin^{2}(\alpha)(d\phi^{2} + \sin^{2}(\phi) d\theta^{2}).$$
 (15)

Geodesics extending from the north pole to q have fixed ϕ , θ , and an α that ranges from 0 at the north pole to a positive value at q.

Specify the original tag orientation q_0 in polar form as $(\alpha_0, \phi_0, \theta_0)$ and likewise with the new tag orientation q_1 . The geodesic Γ_0 connecting 1 to q_0 is of the polar form ([0, α_0], ϕ_0, θ_0), and the new geodesic Γ_1 is ([0, α_1], ϕ_1, θ_1). We are only concerned with geodesics that extend from the north pole to near the south pole, at least where $\alpha_0, \alpha_1 > \pi/2$. We also assume, without loss of generality, that $\phi_0 \leq \phi_1$ and $\theta_0 \leq \theta_1$.

Let corresponding points on Γ_0 and Γ_1 be points of equal α . By observation, the maximum distance between corresponding points on Γ_0 and Γ_1 occurs at the equator, where $\alpha = 1/2$. Let the distance geodesic Γ_d denote the geodesic between corresponding equatorial points of Γ_0 , Γ_1 . By regulating the length of Γ_d , we regulate the rate of the belt trick.

Let $\Delta \phi = \phi_1 - \phi_0$ and $\Delta \theta = \theta_1 - \theta_0$. From (15) we can approximate the length of Γ_d with

$$s^2 \approx \Delta \phi^2 + \sin^2 (\phi_0) \Delta \theta^2 \tag{16}$$

since the change in α on Γ_d is zero and $\sin^2(\pi/2) = 1$.

Let s_{\max} be the maximum allowed length of Γ_d —the maximum rate of change between Γ_0 and Γ_1 . If the length of Γ_d exceeds the maximum allowed length s_{\max} then we must reduce the increment of q. Let

$$\lambda = s_{\max}/s \tag{17}$$

be the amount Γ_d needs to be scaled back to meet the maximum allowed length. Then

$$s_{\max}^{2} = \lambda^{2} s^{2}$$

= $\lambda^{2} (\Delta \phi^{2} + \sin^{2} (\phi_{0}) \Delta \theta^{2}),$
= $(\lambda \Delta \phi)^{2} + \sin^{2} (\phi_{0}) (\lambda \Delta \theta)^{2}.$

Changing the polar values ϕ and θ of q_0 by no more than $\lambda\Delta\phi$ and $\lambda\Delta\theta$ prevents the geodesic from rotating too fast around S³—prevents the belt from moving too quickly. Hence, the incremental rotation

$$q_1 = q_0 e^{\lambda \epsilon \mathbf{u}}, \tag{18}$$

where λ is given in (17) as the quotient of the parameter s_{\max} and the "distance" between successive geodesics s, produces a new tag orientation sufficiently close to the original to limit belt movement properly when necessary. In practice, setting $s_{\max} = 0.1$ disciplines the belt into reasonable behavior.

5. CONCLUSION

Using the methods of Section 4, we have constructed a simulated quaternion demonstrator, as described in Section 5.1. The simulated quaternion demonstrator not only demonstrates unit quaternion multiplication, like its physical counterpart, but also illustrates the quaternion interpolation of orientation from one end of the belt to the other. The initial success of this prototype has inspired several ideas for further research in this direction, which are described in Section 5.2.

5.1 Results

Our implementation of the quaternion demonstrator, titled "quatdemo," was developed on an SGI Indigo Elan, and can be obtained via anonymous ftp to the Imaging Research Laboratory at irl.eecs.wsu.edu from the directory /pub/IRL/quatdemo.

It consists of the coordinate axis and a belt. The imaginaries are labeled at the tag end of the belt, in their corresponding orientations. Pressing and holding the "i," "j," or "k" key rotates the tag end of the belt, causing minimal twists in the belt. The current unit quaternion value is represented visually by the configuration of the demonstrator and is verified by a formatted text version of the current unit quaternion value.

Our implementation simulates the spine of the belt discretely with 256 samples, using the Euler method to approximate the integral (14). Euler integration is highly susceptible to accumulated error, but errors in the

position of the tag end of the quaternion demonstrator are of little consequence. The only state where the position of the tag is noticeable is -1, where the tag end of the belt should be coincident with the fixed end of the belt. We are again fortunate in that the numerical noise accumulated in the incremental rotations of the tag, statistically, prevents this state from being represented exactly.

Figure 2 displays the various states of the quaternion demonstrator simulated by the methods discussed in Section 4. As expected, belt tricks occur automatically when necessary to remove excess twists in the belt; the belt never contains more than a full 360 degree twist in any direction. Figure 3 displays an automatically occurring belt trick, which results at the midway point when holding the "k" key down.

After an automatic belt trick, enough error accumulates to cause the spine of the belt to return to a position slightly offset from its original state. In our implementation, pressing the space bar resets the quaternion demonstrator to the 1 state.

5.2 Further Research

The concept of illustrating the track of a rotation through the use of attached belts to objects is basic to the quaternion demonstrator. In this article we have considered the resulting symmetries of a rectangle in three-dimensional space. The same results apply to the symmetry of any object in three-space \mathbb{R}^3 .

Formally, Let $\mathscr{C} \subset \mathbb{R}^3$ be a subset of \mathbb{R}^3 containing the origin. Let SO(3) denote the set of rotations about the origin of \mathbb{R}^3 . Let $Symm(\mathscr{C})$ denote the subgroup of SO(3) consisting in those rotations $g \in SO(3)$ for which $g(\mathscr{C}) = \mathscr{C}$ setwise. Now the three-sphere \mathbb{S}^3 of unit quaternions covers the SO(3) doubly via the map

$$\pi \colon \mathbb{S}^3 \to \mathbb{SO}(3) \colon \mathbf{v} \to R(q)\mathbf{v} = q\mathbf{v}q^{-1}, \tag{19}$$

where **v** is a pure quaternion—hence **v** is a vector in \mathbb{R}^3 . This is an abstract description of our representation of rotations by quaternions.

Now, the set of unit quaternions, $\pi^{-1}(\text{Symm}(\mathscr{O}))$, covering the symmetry group of the object, is a subgroup of \mathbb{S}^3 , called the *binary group* of $\text{Symm}(\mathscr{O})$. If the object \mathscr{O} is the rectangular tag \mathscr{T} , this group is the eight-element quaternion group,

$$\{\pm 1, \pm \mathbf{i}, \pm \mathbf{j}, \pm \mathbf{k}\}.$$
 (20)

This is the abstract description of what the quaternionic demonstrator has demonstrated.

An extension of the demonstrator (by attaching a belt to another object, \mathcal{O}) can illustrate the binary group for the symmetries of any object. These groups, in the case of regular solids such as the tetrahedron, octahedron, or icosahedron, are of great interest both practically and mathematically. This extension of our demonstrator is one of the immediate prospects for further research.

It is also possible to extend our methods to study the structure of rotations of four-space and to the study and illustration of properties of *octonians* [Cayley 1897], which are an eight-dimensional generalization of the quaternions.

APPENDIX

RENDERING

The most straightforward method for rendering a deformed object is to polygonize its surface, apply the deformation to the polygon vertices, then render the resulting polygons. Special-purpose hardware can render polygonal objects in real time, permitting interactive modeling. For example, the local deformation of the belt used in the quaternion demonstrator was rendered in this fashion. Polygonization can be problematic when investigating deformations in that the result of deforming vertices produces a polygon that may be a poor fit when compared to the deformation of the entire polygon, requiring some form of detection and dynamic subdivision.

Alternatively, we can render the deformed object directly as an implicit surface, preserving, at least to pixel precision, the detail of the deformed geometry. Let $f(\mathbf{x})$ be a function implicitly defining the set $A \subset \mathbb{R}^3$ such that

$$f(\mathbf{x}) < 0 \Leftrightarrow \mathbf{x} \in \check{A},\tag{21}$$

$$f(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} \in \partial A, \tag{22}$$

$$f(\mathbf{x}) > 0 \Leftrightarrow \mathbf{x} \in \mathbb{R}^3 \setminus A,\tag{23}$$

and let $D: \mathbb{R}^3 \to \mathbb{R}^3$ denote the deformation function. Then the deformed set D(A) is implicitly defined by the function $f \circ D^{-1}(\mathbf{x})$.

Ì

With few exceptions, ray tracing is the means for direct visualization of implicit surfaces. Some recent ray intersection methods require the Lipschitz constant of the function [Kalra and Barr 1989; Hart 1993]. The Lipschitz constant of a function $f: A \to B$ from metric space (A, d_A) to metric space (B, d_B) is the smallest positive value λ such that

$$d_{R}(f(x), f(y)) \le \lambda d_{A}(x, y)$$
(24)

for all $x, y \in A$. The Lipschitz constant bounds the amount a transformation can expand an object. If $f: \mathbb{R} \to \mathbb{R}$, then the Lipschitz constant of f indicates the steepest slope in the graph of f. One can use the Lipschitz-based ray-tracing method to investigate the "unfurling," the global deformation described in Section 3.2.

By observation, the largest dilation caused by $B_{s,t}$ occurs when s = 1, t = 0. This deformation winds the segment [r, 1]i twice around the origin. Reducing to \mathbb{R}^2 , the parametric equation of this curve takes the x axis from [r, 1] into the twice-winding curve as

$$f_r(x) = x \cos 2\theta(x), \tag{25}$$

$$f_{y}(x) = x \sin 2\theta(x)$$
(26)

with derivatives

$$f'_{x}(x) = \cos 2\theta(x) - 2\theta'(x)\sin 2\theta(x), \qquad (27)$$

$$f'_{y}(x) = \sin 2\theta(x) + 2\theta'(x)\cos 2\theta(x).$$
(28)

$$\theta'(x) = -\frac{2\pi}{1-r_0}$$
 (29)

Where (29) corresponds to (8). The arc length of this double twist is found using

$$ds^{2} = f'_{x}(x)^{2} + f'_{y}(x)^{2}, \qquad (30)$$

$$= 1 + \frac{16\pi^2}{\left(1 - r_0\right)^2}.$$
 (31)

Since ds^2 reaches its maximum (over the proper domain) when x = 1, we have the Lipschitz constant

Lip
$$B_{1,0} = \sqrt{1 + \frac{16\pi^2}{(1-r_0)^2}}$$
. (32)

The deformation $B_{1,0}$ dilates more than any other $B_{s,t}$ for all $s \in [-1,1]$ and $t \in [0,1]$. Equation (32) is an upper bound of the Lipschitz constant for $B_{s,t}$, and under a similar argument, for the belt trick deformation B_t . Hence, (32) is a suitable (though not necessarily optimal for all parameters s and t) Lipschitz bound for ray-tracing the results of the belt trick and unfurling deformations.

This Lipschitz constant was used to render the unfurling demonstration in Sandin et al. [1993], excerpted in Figure 5.

ACKNOWLEDGMENTS

We wish to thank the editor and reviewers for very extensive comments and suggestions on this topic.

REFERENCES

- BARR, A. H., CURRIN, B., GABRIEL, S., AND HUGHES, J. F. 1992. Smooth interpolation of orientations with angular velocity constraints using quaternions. *Comput. Graph.* 26, 2 (July), 313-320.
- BARR, A. H. 1984. Global and local deformations of solid primitives. Comput. Graph. 18, 3 (July), 21-30.
- BOLKER, E. 1973. The spinor spanner. Am. Math. Month. 80, 9 (Nov.), 977-984.
- CAYLEY, A. 1897. The Collected Mathematical Papers of Arthur Cayley. Cambridge University Press, Cambridge, U.K.
- FRANCIS, G. K. 1987. A Topological Picturebook. Springer-Verlag, New York.
- FRANUIS, G. K. AND KAUFFMAN, L. H. 1994. Air on the Dirac strings. In Mathematical Legacy of Wilhelm Magnus. AMS, Providence, R.I.

HAMILTON, W. R. 1866. Elements of Quaternions. Longmans Green, London.

- HANSON, A. 1992. The rolling ball: Applications of a method for controlling three degrees of freedom using two-dimensional input devices. In *Graphics Gems*, D. Kirk, Ed. Vol. 3. Academic Press, San Diego, Calif., 51-60.
- HART, J. C. 1993. Sphere tracing: Simple robust antialiased rendering of distance-based implicit surfaces. Tech. Rep. EECS-93-15, School of EECS, Washington State Univ., Pullman, Wash. Appears in SIGGRAPH '93, Course Notes #25 "Modeling, Visualizing and Animating Implicit Surfaces."

KALRA, D. AND BARR, A. H. 1989. Guaranteed ray intersections with implicit surfaces. *Comput. Graph.* 23, 3 (July), 297-306.

KAUFFMAN, L. H. 1991. Knots and Physics. World Scientific, Teaneck, N.J.

KAUFFMAN, L. H. 1987. On Knots. Princeton University Press, Princeton, N.J.

- MISNER, C. W., THORNE, K. S., AND WHEELER, J. A. 1973. Gravitation. Freeman, San Francisco, Calif.
- SANDIN, D. J., KAUFFMAN, L. H., AND FRANCIS, G. K. 1993. Air on the Dirac strings. SIGGRAPH Video Rev. 93. Animation.
- SHOEMAKE, K. 1992. ARCBALL: A user interface for specifying three-dimensional orientation using a mouse. In *Proceedings of Graphics Interface '92* (May) 151-156.
- SHOEMAKE, K. 1985. Animating rotation with quaternion curves. Comput. Graph. 19, 3 (July), 245-254.

SPIVAK, M. 1965. Calculus on Manifolds. W. A. Benjamin, New York.

Received October 1993; revised July 1994; accepted July 1994

Editor: John Hughes