Exact Representation and Operations on Spherical Maps

Marcus Vinícius A. Andrade[1 2], Jorge Stolfi[1]

[1]Instituto de Computação - Unicamp, Caixa Postal 6176 - 13083-970 - Campinas, SP, Brasil
marcus,stolfi@dcc.unicamp.br
[2]Depto. de Informática - Universidade Fed. de Viçosa, - Viçosa, MG, Brasil
marcus@dpi.ufv.br

Abstract. We develop exact algorithms for geometric operations on general circles and circular arcs on the sphere, using integer homogeneous coordinates. The algorithms include testing a point against a circle, computing the intersection of two circles, and ordering three arcs out of the same point. These operations allow robust manipulation of maps on the sphere, providing a reliable framework for GIS, robotics, and other geometric applications.

1 Introduction

A *spherical map* is a partition of the sphere's surface in three elements: vertices (points), edges (circles and arcs of circles), and faces (open regions). Most applications in geographical information systems (GIS) involve maps of this type: they arise from geodetic lines, latitude-longitude grids, stereographic projection of plane polygons, satellite images, etc [4].

Our aim is to develop representations and algorithms for spherical maps which are free from round-off errors, and therefore *robust* in the sense of Hoffman and Yap[3, 6]. The central part of our solution is an exact representation of circles on the sphere, and exact algorithms for computing the intersection of two circles, locating a point with respect to a circle, ordering circular arcs around a point, and other geometric operations.

Exactness may seem a pointless luxury in GIS applications, since GIS data is by its nature approximate, and approximate results are sufficient for all practical purposes. However, all geometric algorithms used in GIS, such as point location and map overlay, become much more complex and prone to failure if their basic operations are subject to rounding errors, no matter how small. Consider for example a distributed application that cuts a map into smaller submaps, handles each piece to a separate processor, and combines the partial results into a single map. If the cutting step is exact, the final step needs only to identify common boundary edges between the partial results, and remove them. In contrast, if the cutting step is affected by rounding error, the task becomes much harder: the partial results may overlap, or may be separated by gaps of nonzero width. The pasting operation is almost impossible to specify, let alone to program.

2 Oriented projective geometry

The geometry of Cartesian three-space $\mathbf{R}^3$ is greatly simplified if we view it as a subset of a larger *projective space* $\mathbf{P}^3$ [1, 5]. Besides the ordinary points of $\mathbf{R}^3$, projective space includes *points at infinity*, where parallel lines are assumed to meet. This extension removes many special cases and allows us to unify many algorithms that seem unrelated in the Cartesian model.

Projective space has one drawback, though. Many algorithms for Cartesian geometry are based on the operation of testing which of the two half-spaces defined by a given plane contains a given point. In projective geometry, however, this test is meaningless, since the two half-spaces are connected through the points at infinity.

The solution is to work in an even larger domain, the *oriented projective space* $\mathbf{T}^3$ [5], which consists of two separate copies of $\mathbf{R}^3$, plus the points at infinity. This space retains all the nice properties of $\mathbf{P}^3$, without losing the orientation and separation properties of Cartesian geometry.

2.1 Points

In $\mathbf{T}^3$, a point p is by definition a non-zero quadruplet of numbers $[w, x, y, z]$, its *homogeneous coordinates*, with positive scalars multiples identified—i.e., $[w, x, y, z]$ and $[\lambda w, \lambda x, \lambda y, \lambda z]$ are the same point, for all $\lambda > 0$. Note that $p = [w, x, y, z]$ and $q = [-w, -x, -y, -z]$ are distinct points; we say that each is the *antipode* of the other, and write $p = \neg q$.

By definition, the Cartesian coordinates of a point $p = [w, x, y, z]$ of $\mathbf{T}^3$, with $w \neq 0$, are $(x/w, y/w, z/w)$. In this case, we say that p is a *finite point*; if $w > 0$ the point is said to be in the *front range* of $\mathbf{T}^3$, else it is in the *back range*. Note that p and $\neg p$ have the same Cartesian coordinates but are distinct points of $\mathbf{T}^3$. Conversely, there are two points of $\mathbf{T}^3$ with Cartesian coordinates (x, y, z), namely $[1, x, y, z]$ and $[-1, -x, -y, -z]$. Thus, the front and back ranges of $\mathbf{T}^3$ can be viewed as two copies of $\mathbf{R}^3$ contained in $\mathbf{T}^3$.

A point of $\mathbf{T}^3$ of the form $[0, x, y, z]$ is assumed to be *at infinity*, in the direction of the Cartesian vector (x, y, z), as seen from any point on the front range. Viewed from any point on the back range, the same point is assumed to be at infinity in the direction $(-x, -y, -z)$. For more details, see [5].

2.2 Planes

A plane α in $\mathbf{T}^3$ is represented by a non-zero quadruplet $\langle a_0, a_1, a_2, a_3 \rangle$, its *homogeneous coefficients*. By definition, $\langle a_0, a_1, a_2, a_3 \rangle$ and $\langle \lambda a_0, \lambda a_1, \lambda a_2, \lambda a_3 \rangle$ are the same planes for all $\lambda > 0$. Note that $\langle a_0, a_1, a_2, a_3 \rangle$ and $\beta = \langle -a_0, -a_1, -a_2, -a_3 \rangle$ are distinct planes; we say that each one is the *opposite* of the other, and write $\alpha = \neg\beta$.

By definition, the plane α is incident to every point $[w, x, y, z]$ of $\mathbf{T}^3$ such that $a_0 w + a_1 x + a_2 y + a_3 z = 0$. In general, the set of points that are incident to a plane consists of two copies of the same Euclidean plane, one on each range of $\mathbf{T}^3$, and all points at infinity in directions parallel to that plane. The exceptions are the plane $\Omega_2 = \langle 1, 0, 0, 0 \rangle$, called the *plane at infinity*, and its opposite $\neg\Omega_2$, which are incident to all points at infinity.

Every plane α divides $\mathbf{T}^3$ in two half-spaces, its the *positive* and *negative sides*. By definition, the positive side consists of all points $[w, x, y, z]$ of $\mathbf{T}^3$ such that $a_0 w + a_1 x + a_2 y + a_3 z > 0$. We define the *position* of a point p *relative to* α as

$$p \diamond \alpha = \operatorname{sign}(a_0 p_0 + a_1 p_1 + a_2 p_2 + a_3)$$

Note that $p \diamond \alpha = +1$ if and only if $(\neg p) \diamond \alpha = -1$. Therefore, the positive side of α consists of a Cartesian half-space of the front range, and the complementary Cartesian half-space of the back range (plus a subset of the points at infinity).

The *external orientation* of a plane, defined by '$\diamond$', can be visualized as an arrow pointing from the negative half-space into the positive one. A plane α also has an *internal orientation*, which can be visualized as a small circular arrow drawn on it. See figure 1. By sliding the circular arrow over α, we can tell whether a turn at any point on the plane is *positive* (agreeing with the arrow) or *negative*; or whether any three non-collinear points p, q and r on α form a *positive* or *negative* triangle. Formally, the orientation of the triangle p, q, r on α is given by

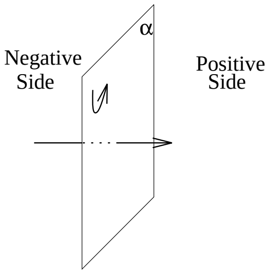

Figure 1: Orientation of a plane.

$$\Delta(p, q, r, \alpha) = \operatorname{sign} \begin{vmatrix} p_0 & p_1 & p_2 & p_3 \\ q_0 & q_1 & q_2 & q_3 \\ r_0 & r_1 & r_2 & r_3 \\ a_0 & a_1 & a_2 & a_3 \end{vmatrix} \tag{1}$$

Note that sliding the arrow across Ω_2 reverses its apparent sense of rotation; i.e. the same turn will have opposite signs on the front and back parts of the plane. Note also that α and $\neg\alpha$ have opposite orientations, both internal and external.

2.3 Lines

A *finite line* of $\mathbf{T}^3$ consists of two copies of the same Cartesian line of $\mathbf{R}^3$, one on the front range and one on the back range, plus the two points at infinity in the directions parallel to those lines. A *line at infinity* consists of all the points at infinity in the directions parallel to some plane.

A line in $\mathbf{T}^3$ has both an *internal* orientation, which defines the positive sense of travel along the line, and an *external* orientation, which defines the positive sense of turning around the line. These components can be visualized as a straight arrow placed on the line, and a circular arrow surrounding the line, respectively. By displacing the arrows across Ω_2 we can check that the front and back parts of the line have the same orientation, internal and external.

2.4 Join and meet

The basic geometric operations of $\mathbf{P}^3$ are *join* and *meet*. The join operation returns the line $p \vee q$ connecting two points p, q, or the plane $p \vee l$ containing a point p and a line l. The meet operation returns the line $\alpha \wedge \beta$ where two planes α, β intersect, or the point $l \wedge \alpha$ where a line l intersects a plane α.

In oriented projective space $\mathbf{T}^3$, the results of join and meet have specific orientations. By definition, the line $p \vee q$ is oriented so as to go from p to q by the route that does not include their antipodes.

To visualize the orientation of the line $l = \alpha \wedge \beta$, we must imagine the plane α turning towards β

around l, by the smallest angle that makes the two planes coincide in position and orientation. The sense of rotation defines the external orientation of l; and the internal orientation can be derived by the right-hand rule. See figure 2.

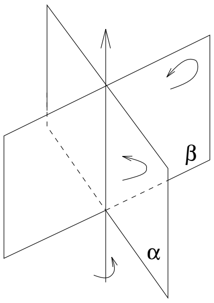

Figure 2: Meet of two planes

Note that $\alpha \wedge \beta = \neg(\beta \wedge \alpha)$; that is, in $\mathbf{T}^3$ the intersection of two planes is anticommutative.

In general, a line l intersects a plane α in two antipodal points. By definition, $l \wedge \alpha$ is the one where l enters the positive side of α. For consistency, it is best to define $l \wedge \alpha$ as equal to $\alpha \wedge l$; i.e., the intersection of a line and plane is commutative.

2.5 Plücker coefficients of a line

It can be proved [5, 2] that the line $l = \alpha \wedge \beta$ is determined by the six coefficients $l_{01}, l_{02}, l_{12}, l_{03}, l_{13}, l_{23}$ where

$$l_{ij} = \begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix} \tag{2}$$

These six numbers are called the *Plücker coefficients* of the line l. We write $l = \langle l_0, l_1, l_2, l_3, l_4, l_5 \rangle$, where the l_i is the ith determinants 2 in the fixed sequence $l_{01}, l_{02}, l_{12}, l_{03}, l_{13}, l_{23}$.

The six Plücker coefficients are not independent: a non-zero sextuple $\langle l_0, .. l_5 \rangle$ represents a line of $\mathbf{T}^3$ if and only if $l_0 l_5 - l_1 l_4 + l_2 l_3 = 0$.

Observe that the Plücker coefficients of a line are homogeneous, that is, $\langle l_0, .. l_5 \rangle$ and $\langle \lambda l_0, .. \lambda l_5 \rangle$ are the same line, for any $\lambda > 0$. On the other hand, $l = \langle l_0, .. l_5 \rangle$ and $m = \langle -l_0, .. -l_5 \rangle$ are distinct lines: although they are incident to the same points, they have different orientations. We say that m is the *opposite* of l, denoted by $m = \neg l$.

Let $p = [p_0, .. p_3]$ and $q = [q_0, .. q_3]$ be two points, $\alpha = \langle a_0, a_1, a_2, a_3 \rangle$ and $\beta = \langle b_0, b_1, b_2, b_3 \rangle$ be two planes and $l = \langle l_0, .. l_5 \rangle$ be a line of $\mathbf{T}^3$. The basic geometric operations of $\mathbf{T}^3$ are given by these formulas:

$$\begin{aligned}
p \vee q = {} & \langle p_2 q_3 - p_3 q_2,\ p_3 q_1 - p_1 q_3,\ p_0 q_3 - p_3 q_0, \\
& p_1 q_2 - p_2 q_1,\ p_2 q_0 - p_0 q_2,\ p_0 q_1 - p_1 q_0 \ \rangle \\
p \vee l = {} & \langle \ \ l_0 p_1 + l_1 p_2 + l_3 p_3, -l_0 p_0 + l_2 p_2 + l_4 p_3, \\
& -l_1 p_0 - l_2 p_1 + l_5 p_3, -l_3 p_0 - l_4 p_1 - l_5 p_2 \ \rangle \\
\alpha \wedge \beta = {} & \langle a_0 b_1 - a_1 b_0, a_0 b_2 - a_2 b_0, a_1 b_2 - a_2 b_1, \\
& a_0 b_3 - a_3 b_0, a_1 b_3 - a_3 b_1, a_2 b_3 - a_3 b_2 \ \rangle \\
l \wedge \alpha = {} & [-l_2 a_3 + l_4 a_2 - l_5 a_1, l_1 a_3 - l_3 a_2 + l_5 a_0, \\
& -l_0 a_3 + l_3 a_1 - l_4 a_0, l_0 a_2 - l_1 a_1 + l_2 a_0 \ ]
\end{aligned}$$

These formulas return all-zero tuples when the geometric object is not defined. In particular, the line $p \vee q$ is undefined if $p = q$ or $p = \neg q$, and the plane $p \vee l$ is undefined when p lies on l. Similarly, the line $\alpha \wedge \beta$ is undefined if $\alpha = \beta$ or $\alpha = \neg\beta$, and the point $l \wedge \alpha$ is undefined if l lies on α.

If l is not a line at infinity, its *direction* $\mathrm{dir}(l)$ is the point at infinity where l intersects Ω_2, that is

$$\mathrm{dir}(l) = l \wedge \Omega_2 = [0, l_5, -l_4, l_2] \tag{3}$$

Note that two lines l and m not on Ω_2 are parallel if $\mathrm{dir}(l) = \mathrm{dir}(m)$ or $\mathrm{dir}(l) = \neg\,\mathrm{dir}(m)$.

The Euclidean distance between l and the origin $O = [1, 0, 0, 0]$ of $\mathbf{T}^3$ is

$$\mathrm{dist}(l, O) = \sqrt{\frac{l_0^2 + l_1^2 + l_3^2}{l_2^2 + l_4^2 + l_5^2}} \tag{4}$$

It follows that l passes through O if and only if $l_0 = l_1 = l_3 = 0$. The point of l closest to O is

$$\begin{aligned}
\mathrm{mid}(l) = {} & [\ l_2^2 + l_4^2 + l_5^2, -l_1 l_2 - l_3 l_4 \\
& l_0 l_2 - l_3 l_5, \quad l_0 l_4 + l_1 l_5\ ]
\end{aligned} \tag{5}$$

Two lines $l = \langle l_0, .. l_5 \rangle$ and $m = \langle m_0, .. m_5 \rangle$ intersect if and only if

$$l_0 m_5 - l_1 m_4 + l_2 m_3 + l_3 m_2 - l_4 m_1 + l_5 m_0 = 0$$

In that case, if the lines are not parallel, the intersection consists of two antipodal finite points. We define *the front intersection* point $l \underset{+}{\wedge} m$ as being the one in the front range, which is given by

$$\begin{aligned}
l \underset{+}{\wedge} m = [& \kappa_{24}^2 + \kappa_{25}^2 + \kappa_{45}^2, \\
& (l_3 m_2 - l_1 m_4 + l_5 m_0)\kappa_{24} - \kappa_{15}\kappa_{25} - \kappa_{35}\kappa_{45}, \\
& (l_0 m_5 + l_3 m_2 - l_4 m_1)\kappa_{25} + \kappa_{04}\kappa_{24} + \kappa_{34}\kappa_{45}, \\
& (l_0 m_5 - l_1 m_4 + l_2 m_3)\kappa_{45} - \kappa_{02}\kappa_{24} - \kappa_{12}\kappa_{25}]
\end{aligned} \tag{6}$$

where

$$\kappa_{ij} = \begin{vmatrix} l_i & l_j \\ m_i & m_j \end{vmatrix}$$

3 Spherical geometry

3.1 The unit sphere

In $\mathbf{T}^3$, the *unit sphere* $\mathbf{S}^2$ is the sphere with unit radius centered at origin, that is,

$$\mathbf{S}^2 = \{[w, x, y, z] \,|\, x^2 + y^2 + z^2 - w^2 = 0\}$$

Note that $\mathbf{S}^2$ corresponds to two copies of the unit sphere of $\mathbf{R}^3$: one in the front range and another one in the back range.

The *position* of a point $p = [p_0, p_1, p_2, p_3]$ *relative to* $\mathbf{S}^2$ is

$$p \circ \mathbf{S}^2 = \operatorname{sign}(p_1^2 + p_2^2 + p_3^2 - p_0^2)$$

Note that the negative side of $\mathbf{S}^2$ (the points with $p \circ \mathbf{S}^2 = -1$) consists of two separate unit balls.

3.2 Circles on the sphere

In $\mathbf{T}^3$, the intersection of a plane α and $\mathbf{S}^2$ consists of a pair of antipodal circles c, one in the front range and another one in the back range. We will say that c is the *S-circle* whose *supporting plane* is α, and write $c = \operatorname{scrc}(\alpha)$, $\alpha = \operatorname{spln}(c)$. By definition, the *positive sense of travel* along $\operatorname{scrc}(\alpha)$ is the one that agrees with the internal orientation of α. Informally, the positive sense of travel on the $\operatorname{scrc}(\alpha)$ is counterclockwise as seen by an observer on the front range of $\mathbf{T}^3$ and on the positive side of α (assuming the axes of $\mathbf{R}^3$ are in their standard position). See figure 3.

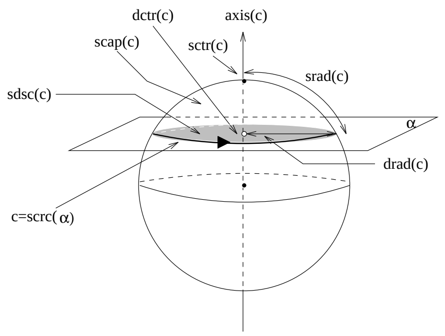

Figure 3: Elements of an S-circle.

The S-circle c defined by $\alpha = \langle a_0, a_1, a_2, a_3 \rangle$ will be denoted by $(\!(a_0, a_1, a_2, a_3)\!)$. Note that α defines an S-circle (i.e., α intersects $\mathbf{S}^2$) if and only if $a_0^2 \leq a_1^2 + a_2^2 + a_3^2$. If equality holds, α is tangent to the sphere, and c reduces to a pair of antipodal points.

Note that an S-circle c divides each component of $\mathbf{S}^2$, front and back, into two regions (one of which may be empty). By definition, the *S-cap* or *front positive side* of c, denoted by $\operatorname{scap}(c)$, is the region of $\mathbf{S}^2$ that lies on the front range of $\mathbf{T}^3$ and on the positive side of the supporting plane α. The center of $\operatorname{scap}(c)$ is called the *S-center* of c, and denoted by $\operatorname{sctr}(c)$. Its coordinates are

$$\operatorname{sctr}(c) = \left[\sqrt{a_1^2 + a_2^2 + a_3^2}, a_1, a_2, a_3\right]$$

The front part of $\operatorname{spln}(c)$ bounded by the circle c is called the *S-disk* of c, and denoted $\operatorname{sdsk}(c)$. Its center is the *D-center* of c, or $\operatorname{dctr}(c)$ for short, whose coordinates are

$$\operatorname{dctr}(c) = \left[a_1^2 + a_2^2 + a_3^2, -a_0 a_1, -a_0 a_2, -a_0 a_3\right]$$

In particular, if the circle c is a single point p, its S-cap and S-disk are empty, and $\operatorname{sctr}(c) = p = \operatorname{dctr}(c)$.

The *axis* of an S-circle c is the line $\operatorname{axis}(c) = O \vee \operatorname{sctr}(c)$, where O is the front origin of $\mathbf{T}^3$. The direction of $\operatorname{axis}(c)$, namely $\operatorname{axis}(c) \wedge \Omega_2$, is the *normal* of c, denoted by $\operatorname{snrm}(c)$, that is

$$\begin{aligned} \operatorname{axis}(c) &= \langle 0, 0, a_3, 0, -a_2, a_1 \rangle \\ \operatorname{snrm}(c) &= [0, a_1, a_2, a_3] \end{aligned}$$

The length of the spherical arc that joins $\operatorname{sctr}(c)$ to any point on c is the *S-radius* of c, denoted by $\operatorname{srad}(c)$. The distance from $\operatorname{dctr}(c)$ to any point on c is the *D-radius* of c, denoted by $\operatorname{drad}(c)$. The S-circle c whose S-center is $p = [p_0, p_1, p_2, p_3]$ and whose S-radius is $\theta \in [0 \,\_\_\, \pi]$ is

$$c = (\!(-p_0 \cos\theta, p_1, p_2, p_3)\!)$$

3.3 Spherical Polarity

Consider two quadruplets of numbers a_0, a_1, a_2, a_3 and b_0, b_1, b_2, b_3 such that $a_0 b_0 + a_1 b_1 + a_2 b_2 + a_3 b_3 = 0$. We can interpret this equation as saying that the point $[a_0, a_1, a_2, a_3]$ is on the plane $\langle b_0, b_1, b_2, b_3 \rangle$; or, conversely, that the point $[b_0, b_1, b_2, b_3]$ is on the plane $[a_0, a_1, a_2, a_3]$. Thus, the roles of point and plane in the equation are interchangeable. This symmetry is an instance of a fundamental property of projective space, the *principle of duality* [1].

In general terms, projective duality is formalized by the concept of *duomorphism*, a correspondence between two spaces that maps planes to points, lines to lines and points to planes, while preserving the relative position predicate '$\diamond$.' It can be shown that a duomorphism exchanges join and meet, that is, the meet of two planes becomes the join of their dual points, and vice-versa. For our purposes, the following duomorphism is particularly useful:

Definition 1 *The* polar complement *relative to* $\mathbf{S}^2$*, denoted by '$*$', is given by:*

$$[p_0, p_1, p_2, p_3]^* = \langle -p_0, p_1, p_2, p_3 \rangle$$

$$\begin{aligned}\langle a_0, a_1, a_2, a_3\rangle^* &= [-a_0, a_1, a_2, a_3]\\ \langle l_0, l_1, l_2, l_3, l_4, l_5\rangle^* &= \langle -l_5, l_4, l_3, -l_2, -l_1, l_0\rangle\end{aligned}$$

Geometrically, if p is a point outside $\mathbf{S}^2$, let c be the circle where the cone tangent to $\mathbf{S}^2$ with vertex in p meets $\mathbf{S}^2$. See figure 4. Then p^* is the supporting plane of c, oriented so that p is on its positive side.

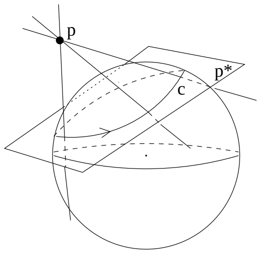

Figure 4: Point-Plane spherical polarity

More generally, the polar of a finite point $p = [p_0, p_1, p_2, p_3]$ (inside or outside the sphere) is the plane that cuts the ray from O through p perpendicularly at distance $1/\operatorname{dist}(p, O)$ from O, and oriented so that $p \diamond p^* = p \circ \mathbf{S}^2$. In particular, $O^* = \Omega_2$.

It turns out that '$*$' is the composition of Stolfi's *right polar complement* '$\vdash$' [5, page 85] and the projective map $M = [p_0, p_1, p_2, p_3] \mapsto [-p_0, p_1, p_2, p_3]$. Strictly speaking, M is an isomorphism between $\mathbf{T}^3$ and $\neg\mathbf{T}^3$, a copy of $\mathbf{T}^3$ where the meet of two planes is defined with opposite orientation as in $\mathbf{T}^3$. Therefore, the effect of '$*$' on meet and join sometimes requires orientation reversal. More precisely, for any points, lines, or planes u, v of $\mathbf{T}^3$, we have

$$\begin{aligned}(u \vee v)^* &= u^* \wedge v^*\\ (u \wedge v)^* &= \neg(u^* \vee v^*)\\ (u^*)^* &= \neg^d u\end{aligned}$$

On the other hand, '$*$' preserves the relative position of points and planes; that is, $\alpha^* \diamond p^* = p \diamond \alpha$ for any plane α and any point p.

3.4 Tangent lines and planes

It's easy to see that, if p is a point on the sphere, then p^* is a plane that is tangent to the sphere at p. Moreover, if p is a front range point, the plane p^* will be oriented so that the front origin O is on its negative side; and $\operatorname{snrm}(p^*)$ is the direction of the line $O \vee p$, that is, $\operatorname{snrm}(p^*) = [0, p_1, p_2, p_3]$.

Also, let p be a front point on an S-circle c. The line $\operatorname{spln}(c) \wedge p^*$ is tangent to c at p, and oriented so as to agree with the orientation of c at that point.

3.5 Arcs of S-circles

Let p and q be two distinct front range points on an S-circle $c = \operatorname{scrc}(\alpha)$. The points divide the front part of c into two connected parts. By definition, *the S-arc of c from p to q*, denoted by $\operatorname{sarc}(p, q, c)$, is the set of points encountered as we move from p to q on c, along its positive sense of travel.

The arc can be described geometrically as the part of c that lies on the front range and on the positive side of the plane $\beta = p \vee q \vee \alpha^*$. In fact, we can let β be any plane such that $\alpha \wedge \beta = p \vee q$.

3.6 Intersection of S-circles

Let a and b be two S-circles and let l be the line $\operatorname{spln}(a) \wedge \operatorname{spln}(b)$. It's easy to see that $a \cap b = l \cap \mathbf{S}^2$. It follows from equation (4) that, depending on whether $(l_2^2 + l_4^2 + l_5^2) - (l_0^2 + l_1^2 + l_3^2)$ is negative, zero, or positive, the intersection of two S-circles consists of respectively zero, one or two pairs of antipodal points.

We define *the front meeting point* $a \mathbin{\hat{+}} b$ of two S-circles a and b as being the point where a arrives at b from its positive side, or leaves b into its negative side. See figure 5. Another way of identifying $a \mathbin{\hat{+}} b$ is to observe that the unit vectors of $\mathbf{R}^3$ corresponding to $\operatorname{sctr}(a)$, $\operatorname{sctr}(b)$ and $a \mathbin{\hat{+}} b$, in that order, form a positive basis of $\mathbf{R}^3$.

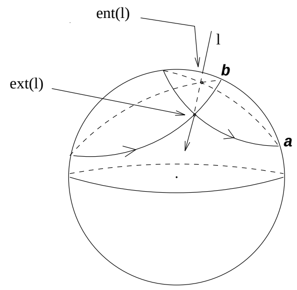

Figure 5: Intersection of two S-circles.

3.7 Enter and exit points of a line

In general, if l is a line that meets $\mathbf{S}^2$, we can distinguish the intersection points using the orientation of l. We define $\operatorname{ent}(l)$ and $\operatorname{ext}(l)$ to be the points on the front range where l enters and leaves the positive side of $\mathbf{S}^2$, respectively. See figure 5. Then, for two S-circles a and b, we can write $a \mathbin{\hat{+}} b = \operatorname{ext}(\operatorname{spln}(a) \wedge \operatorname{spln}(b))$.

As we can prove,

$$\begin{aligned}\mathrm{ent}(l) &= [\ \mu\ ,\ -l_1 l_2 - l_3 l_4 - l_5\sqrt{\delta},\\ &\qquad l_0 l_2 - l_3 l_5 + l_4\sqrt{\delta},\\ &\qquad l_0 l_4 + l_1 l_5 - l_2\sqrt{\delta}\ ]\\ \mathrm{ext}(l) &= [\ \mu\ ,\ -l_1 l_2 - l_3 l_4 + l_5\sqrt{\delta},\\ &\qquad l_0 l_2 - l_3 l_5 - l_4\sqrt{\delta},\\ &\qquad l_0 l_4 + l_1 l_5 + l_2\sqrt{\delta}\ ]\end{aligned} \tag{7}$$

where $\mu = l_2^2 + l_4^2 + l_5^2$ and $\delta = \mu - (l_0^2 + l_1^2 + l_3^2)$. Note that $\mathrm{ent}(l) = \mathrm{ext}(l)$ if and only if l is tangent to $\mathbf{S}^2$, that is $(l_2^2 + l_4^2 + l_5^2) = (l_0^2 + l_1^2 + l_3^2)$.

If p is a front point of $\mathbf{S}^2$, any line l such that $p = \mathrm{ext}(l)$ is called a *stabbing line* for p.

4 Exact spherical geometry

4.1 Rational points, lines, and planes

A point, line, or plane of $\mathbf{T}^3$ is said to be *rational* if all its homogeneous coordinates or coefficients are rational numbers. Note that the set of rational objects in $\mathbf{T}^3$ is dense in the set of all objects.

Since a point of $\mathbf{T}^3$ does not change when its coordinates are scaled by a positive factor, the set $\mathbf{QT}^3$ of all rational points of $\mathbf{T}^3$ is precisely the set of points with integer homogeneous coordinates. The same holds for rational lines and planes.

4.2 Exactly representable points of $\mathbf{S}^2$

There are two sets of points of $\mathbf{S}^2$ that have obvious exact representations. One is the set $\mathbf{A}$ of those points whose all four homogeneous coordinates are rational. Another is the set $\mathbf{B}$ consisting of the points whose x, y, and z coordinates are rational; since the points lie on $\mathbf{S}^2$, the weight coordinate w is determined by the formula $w = \sqrt{x^2 + y^2 + z^2}$. In other words, $\mathbf{B}$ is the radial projection onto $\mathbf{S}^2$ of all rational points of $\mathbf{R}^3$. Since w does not have to be explicitly stored or computed, it does not matter whether it is rational

$$\begin{aligned}\mathbf{A} &= \{[a_0, .. a_3] \mid a_1^2 + a_2^2 + a_3^2 = a_0^2 \text{ and}\\ &\qquad (a_0, .. a_3) \in \mathbf{Z}^4\}\\ \mathbf{B} &= \{[b_0, .. b_3] \mid b_1^2 + b_2^2 + b_3^2 = b_0^2 \text{ and}\\ &\qquad (b_1, b_2, b_3) \in \mathbf{Z}^3\}\end{aligned}$$

Obviously, $\mathbf{B}$ is a proper superset of $\mathbf{A}$.

4.3 Rational circles

An S-circle is *rational* if it is defined by a rational plane. Our proposition is to consider only rational S-circles and arcs thereof. Again, this is not a significant restriction for practical purposes, since the subset of rational S-circles is dense in the set of all S-circles, in the Hausdorff metric.

Let $a = ((a_0, .. a_3))$ be a rational S-circle, and $l = \langle l_0, .. l_5\rangle$ be a rational line. Then, the elements

$$\begin{array}{ll}\mathrm{sctr}(a) \in \mathbf{B} & \mathrm{mid}(l) \in \mathbf{QT}^3\\ \mathrm{dctr}(a) \in \mathbf{A} & \mathrm{snrm}(a) \in \mathbf{QT}^3\end{array}$$

can be represented exactly.

Unfortunately the front meeting of two rational S-circles may not be a rational point, and not even a point of $\mathbf{B}$. For instance, $((1, 2, 2, 2)) \mathbin{\underset{+}{\wedge}} ((1, 2, -2, 2))$ is the point $p = \left[4\ ,\ -1 + \sqrt{7}\ ,\ 0\ ,\ -1 - \sqrt{7}\right]$. Since the ratio p_3/p_1 is not rational, p is not in $\mathbf{B}$.

On the other hand, the intersection of the two rational supporting planes is a rational line. It follows that the front meeting point of two rational S-circles lies in the set

$$\mathbf{C} = \{\, \mathrm{ext}(l) \mid l \text{ is a rational line}\,\}$$

A point p of $\mathbf{C}$ can be represented exactly by the six integer coefficients of any of its rational stabbing lines. In that case, we denote that line by $\mathrm{lin}(p)$.

Let $l = \langle l_0, .. l_5\rangle$ be a rational line. Equation (7) implies that $\mathrm{ext}(l)$ is in $\mathbf{A}$ if and only if $(l_2^2 + l_4^2 + l_5^2) - (l_0^2 + l_1^2 + l_3^2)$ is a perfect square. The set $\mathbf{A}$ can be characterized as follows:

Theorem 1 *A point p lies in $\mathbf{A}$ if and only if there are two non parallel rational lines l and m such that* $\mathrm{ext}(l) = p = \mathrm{ext}(m)$.

Here is a direct characterization of the set $\mathbf{C}$:

Theorem 2 *A point p of $\mathbf{T}^3$ is in $\mathbf{C}$ if and only if*

$$p = \left[a_0, a_1 + b_1\sqrt{c}, a_2 + b_2\sqrt{c}, a_3 + b_3\sqrt{c}\,\right]$$

where a_0, a_1, a_2, a_3, b_1, b_2, b_3, and c are integers satifying

(i) $a_0 \neq 0$ *and* $b_1^2 + b_2^2 + b_3^2 \neq 0$

(ii) $a_1 b_1 + a_2 b_2 + a_3 b_3 = 0$

(iii) $(a_1^2 + a_2^2 + a_3^2) + (b_1^2 + b_2^2 + b_3^2)c = a_0^2$

In that case, a stabbing line of p is

$$\begin{array}{lll}\langle a_2 b_3 - a_3 b_2, & a_3 b_1 - a_1 b_3, & a_0 b_3,\\ \ a_1 b_2 - a_2 b_1, & -a_0 b_2, & a_0 b_1\rangle\end{array}$$

5 Algorithms for spherical circles

We will now describe some exact algorithms for rational points and circles on $\mathbf{S}^2$. We will assume available the basic geometric operations of $\mathbf{T}^3$, such as '$\neg$', '$\diamond$', '$\wedge$', '$\vee$', and the equality test '$\equiv$' for points, lines, and planes (ignoring homogeneous scaling).

5.1 Equality test for points in C

Since a point p of $\mathbf{C}$ can have many stabbing lines, it is not trivial to decide whether two tuples $\langle l_0, .. l_5\rangle$ and $\langle m_0, .. m_5\rangle$ denote the same point. To implement this test, we need the following results:

Theorem 3 *If l is a line that intercepts $\mathbf{S}^2$, then* $\mathrm{ent}(l) \diamond \mathrm{dir}(l)^* \leq 0$, *and* $\mathrm{ext}(l) \diamond \mathrm{dir}(l)^* \geq 0$.

Corollary 4 *If l and m are two coplanar and non-parallel lines, and $q = l \underset{+}{\wedge} m$, then* $\mathrm{ext}(l) = \mathrm{ext}(m)$ *if and only if $q \in \mathbf{S}^2$ and $q \diamond \mathrm{dir}(l)^* = +1$ and $q \diamond \mathrm{dir}(m)^* = +1$.*

procedure CEqual(l, m);
  input: lines $l = \langle l_0, .. l_5\rangle$ and $m = \langle m_0, .. m_5\rangle$.
  output: $\mathrm{ext}(l) \equiv \mathrm{ext}(m)$.
begin
  if ($l \equiv m$) **then return true**
  else if ($l \equiv \neg m$) **then**
    return $l_0^2 + l_1^2 + l_3^2 = l_2^2 + l_4^2 + l_5^2$
  else
    $d \leftarrow l_0 m_5 - l_1 m_4 + l_2 m_3 + l_3 m_2 - l_4 m_1 + l_5 m_0$
    if $d \neq 0$ **then return false** $\{ l \wedge m = \emptyset \}$
    else
      $q \leftarrow l \underset{+}{\wedge} m$;
      return $q \circ \mathbf{S}^2 = 0$ **and**
        $q \diamond \mathrm{dir}(l)^* = +1$ **and**
        $q \diamond \mathrm{dir}(m)^* = +1$
end.

5.2 Relative position of point and S-circle

Another important predicate is the position of a given point $p \in \mathbf{C}$ relative to a given S-circle c. As explained in section 3.2, we use the orientation of c to distinguish the two caps delimited by c. The test relies on the following result:

Theorem 5 *Let c be a rational S-circle and let l be a rational line such that* $\mathrm{ext}(l) \in \mathbf{C} \setminus \mathbf{A}$. *Then* $\mathrm{ext}(l)$ *lies on c if and only if l lies on the plane* $\mathrm{spln}(c)$.

procedure SideCpointAndScircle(l, c);
  input: a line l and an S-circle c.
  output: $\mathrm{ext}(l) \diamond \mathrm{spln}(c)$.
begin
  $q \leftarrow l \wedge \mathrm{spln}(c)$
  if $q = [0, 0, 0, 0]$ **then return** 0
  $\sigma \leftarrow q_1^2 + q_2^2 + q_3^2 - q_0^2$
  $\mu \leftarrow \mathrm{mid}(l) \diamond \mathrm{spln}(c)$; $\delta \leftarrow \mathrm{dir}(l) \diamond \mathrm{spln}(c)$
  if $s < 0$ **then return** δ
  else if $s > 0$ **then return** μ
  else if $\mu = \delta$ **then return** μ
  else return 0
end.

6 Circular order

Many algorithms for Euclidean geometry rely on the ordering of points along a line. When we extend such algorithms to projective space, or to circles on the sphere, we must replace this notion of linear order with that of *circular order*.

6.1 Ahead and behind

The basic predicate $\oplus(p, q, c)$ for circular ordering takes two points p, q on an oriented (topological) circle c, and returns $+1$ if q is *ahead of* p, -1 if q is *behind* p. For this predicate to make sense, we must define what is the "natural" or "shortest" route from p to q. Equivalently, we must define the "diametral opposite" point $\ominus(p, c)$ of each point p on the circle c; then the "shortest" path from p to q is the one that does not go through $\ominus(p, c)$. It follows that the predicate '$\oplus$' is undefined if $q = p$, or $q = \ominus(p, c)$. For consistency, the opposite point must be defined in such a way that

$$\begin{aligned} \oplus(p, q, c) &= \oplus(q, \ominus(p, c), c) \\ &= \oplus(\ominus(p, c), \ominus(q, c), c) \\ &= \oplus(\ominus(q, c), p, c) \end{aligned}$$

In particular, a straight line l of $\mathbf{T}^3$ has the topology of a circle. We can define the ahead/behind predicate $\oplus(p, q, l)$ for two points p,q of l as $+1$ if $p \vee q \equiv l$, -1 if $p \vee q \equiv \neg l$, and 0 if $p \equiv q$ or $p \equiv \neg q$.

Dually, the set of all planes passing through a given line l has the topology of a circle, and is circularly orderd by the external orientation of l. The ahead/behind predicate $\oplus(\alpha, \beta, l)$ for two of those planes is defined as for points, with $\wedge$ substituted for $\vee$.

6.2 Circular order of three points

The concept of circular order can also be formalized by a predicate $\otimes(p, q, r, c)$ that returns $+1$ if the three points p, q, r occur in that order along the oriented circle c; -1 if they occur in the opposite order; and 0 if any two of the points coincide.

This approach is more elegant, because it depends only on the topological orientation of c, and does not require the notion of "opposite point" or "shortest route." However, it is often simpler to compute $\otimes(p, q, r, c)$ by reducing it to three calls of the ahead/behind predicate $\oplus$, as shown below:

procedure CircularOrder($p, q, r, c, \oplus$);
  input: three points p, q, r on
    an oriented topological circle c,
    and the ahead/behind predicate $\oplus$ of c.

output: $\otimes(p, q, r, c)$.
begin
$s_{pq} \leftarrow \oplus(p, q, c)$
$s_{qr} \leftarrow \oplus(q, r, c)$
$s_{rp} \leftarrow \oplus(q, r, c)$
return $\mathrm{sign}(s_{pq} + s_{qr} + s_{rp})$
end.

Below we will use this procedure to test the circular order of three planes α, β and γ passing through a common line l, relative to the external orientation of that line.

6.3 Circular ordering of three arcs around a common origin

In the construction of spherical maps, a common subproblem is to compute the circular ordering of three arcs $A = \mathrm{sarc}(v, p, a)$, $B = \mathrm{sarc}(v, q, b)$ and $C = \mathrm{sarc}(v, r, c)$, in a sufficiently small neighbourhood of their common endpoint v. Informally, the result $\otimes(A, B, C, v)$ should be $+1$ if A, B, C leave v in counterclockwise order, as seen from a front range point on the positive side of the tangent plane v^*. The result should be -1 if the arcs leave v in clockwise order, and 0 if any two of them are parts of the same circle. See figure 6.

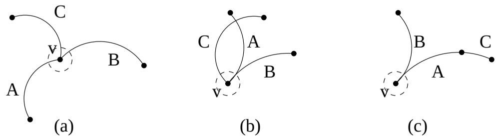

Figure 6: Circular order of three S-arcs.

In the general case, we only need to order the directions in which the three arcs leave the point v, on the line $v^* \wedge \Omega_2$. However, we must overcome two difficulties: the direction vectors may not be rational, and two distinct arcs may leave the point v in the same direction (see figure 6 (b)).

If $v \notin \mathbf{A}$, then, by theorem 1, we know that v has a unique rational stabbing line l (and l is not tangent to the sphere). Therefore, the supporting planes of the three circles pass through the common line l. In that case, the ordering of the arcs is the same as the order of the three planes around l, i.e. $\otimes(\mathrm{spln}(a), \mathrm{spln}(b), \mathrm{spln}(c), l)$, which can be computed exactly.

If $v \in \mathbf{A}$, the three planes may not have a common line. In this case, however, the plane v^* tangent to the sphere at v is rational, so we can compute the oriented tangents at v to the three arcs, and circularly order their directions along the line $v^* \wedge \Omega_2$.

To complete the algorithm, we need to consider the case of two or more coincident tangent lines. In that case, we must break the tie by comparing the curvatures of the arcs—which is again equivalent to ordering their planes around the common tangent line. In particular, if all three tangents are equal, the result is simply the circular order of the three planes around that common line. If only two of the tangents coincide, it suffices to check the ahead/behind order of that two corresponding planes around that common line.

procedure CircOrderArcs(A, B, C, v);
input: three arcs $A = \mathrm{sarc}(v, p, a)$,
$B = \mathrm{sarc}(v, q, b)$, and $C = \mathrm{sarc}(v, r, c)$.
output: $\otimes(A, B, C, v)$.
begin
$\alpha \leftarrow \mathrm{spln}(a)$; $\beta \leftarrow \mathrm{spln}(b)$; $\gamma \leftarrow \mathrm{spln}(c)$
if $v \notin \mathbf{A}$ **then**
{α, β and γ must contain $\mathrm{lin}(v)$,}
{and $\mathrm{lin}(v)$ is not tangent to $\mathbf{S}^2$.}
return $\otimes(\alpha, \beta, \gamma, \mathrm{lin}(v))$
else
{The tangent plane v^* is rational.}
$t_a \leftarrow \alpha \wedge v^*$; $t_b \leftarrow \beta \wedge v^*$; $t_c \leftarrow \gamma \wedge v^*$
if $t_a \not\equiv t_b$ **and** $t_b \not\equiv t_c$ **and** $t_c \not\equiv t_a$ **then**
return $\otimes(\mathrm{dir}(t_a), \mathrm{dir}(t_b), \mathrm{dir}(t_c), v^* \wedge \Omega_2)$
else if $t_a \equiv t_b \equiv t_c$ **then**
return $\otimes(\alpha, \beta, \gamma, t_a)$
else
{Only two tangents are equal.}
while $t_a \not\equiv t_b$ **do**
$(\alpha, \beta, \gamma) \leftarrow (\beta, \gamma, \alpha)$;
$(t_a, t_b, t_c) \leftarrow (t_b, t_c, t_a)$
return $\oplus(\alpha, \beta, t_a)$
end;

References

[1] H. S. M. Coxeter. *The Real Projective Plane.* McGraw-Hill, 1949.

[2] J. Little D. Cox and D. O'Shea. *Ideals, Varieties, and Algorithms.* Springer-Verlag, 1992.

[3] C. M. Hoffmann. The problems of accuracy and robustness in geometric computation. *IEEE Computer*, (22):31–42, 1989.

[4] D. J. Maguire, M. F. Goodchild, and D. Rhind. *Geographical Information Systems - principles and applications.* John Wiley & Sons, 1991.

[5] J. Stolfi. *Oriented Projetive Geometry - A framework for geometric computations.* Academic Press, 1991.

[6] C. Yap and T. Dubé. The exact computation paradigm. Technical report, Courant Inst. of Math. Sciences - New York Univ., 1994.