J. L. Brown
A. J. Worsley

Problems with defining barycentric coordinates for the sphere


<http://www.numdam.org/item?id=M2AN_1992__26_1_37_0>
PROBLEMS WITH DEFINING BARYCENTRIC COORDINATES
FOR THE SPHERE

by J. L. BROWN (1) and A. J. WORSEY (2)

Abstract. — We consider the problem of assigning barycentric coordinates for triangles on the sphere $S^2$. We show that a direct generalization of barycentric coordinates in the plane to geodesic triangles on $S^2$ is not possible. Geodesic triangles are the « natural » choice for the sphere and our results indicate that the techniques for triangular Bézier patches over the plane do not generalize to $S^2$. If we relax the condition that the spherical triangles have edges that are geodesics, then it is possible to define barycentric coordinate systems on the sphere. This is done by constructing area preserving maps from $R^2$ to $S^2$. However, the triangles so generated are inevitably distorted, as shown by examples.

1. INTRODUCTION

The problem of defining surfaces on surfaces is one of some importance in Computer Aided Geometric Design (CAGD), as illustrated by the papers of Barnhill (1985) and Barnhill & Ou (1990), only two of the many references to this whole area. In particular, the problem of defining curves and surfaces over the sphere is clearly pertinent, since it allows us to address issues where the need is to model phenomena using data taken from the surface of the Earth. Interpolation over the sphere is therefore clearly

(1) Department of Mathematical Sciences, University of North Carolina at Wilmington, Wilmington, NC 28403, U.S.A.

(2) Research and Development Center, General Electric Company, P.O. Box 8, Schenectady, NY 12301, U.S.A.

M2 AN Modélisation mathématique et Analyse numérique 0764-583X/92/01/37/13/$ 3.30
Mathematical Modelling and Numerical Analysis © AFCET Gauthier-Villars
important and it is a problem that has been considered by, amongst many others, Lawson (1984), Renka (1984), Nielson and Ramaraj (1987), Nielson (1989), Pottmann and Eck (1990), and Foley (1990a, b).

Nielson (1989) considered the problem of generalizing the geometric construction of the de Casteljau algorithm for Bézier curves ; Boehm et al. (1984, p. 8) (and, analogously, the knot insertion algorithm for B-spline curves ; Boehm (1980)), for constructing smooth piecewise curves over the sphere $S^2$. The idea is to define a polygon composed of geodesic line segments on the sphere and generate a curve from this polygon by repeated geodesic interpolation at a parameter value $t$. This approach is, arguably, the natural one for defining « Bézier curves on $S^2»$.

At the time, Nielson (1989) also addressed the problem of extending his approach for curves on $S^2$ to define triangular patches over the sphere. Triangular Bézier patches are a fundamental tool in CAGD ; see Boehm et al. (1984), Farin (1986, 1989), offering as they do, elegant geometric constructions for surfaces. It is therefore entirely reasonable to want to generalize them to the case where $S^2$ is the domain, given the numerous applications to surface on surface problems.

For surfaces defined over $n$-simplices in $\mathbb{R}^n$, Bernstein-Bézier techniques hinge fundamentally upon the use of barycentric coordinates. It follows that as a precursor to defining triangular Bézier patches over the sphere, we need to be able to define such a coordinate system for domain triangles on $S^2$. In particular, since geodesic triangles are the only ones that are intrinsic to the sphere, we would like to define a barycentric coordinate system for such triangles. That is, triangles whose vertices $p_1$, $p_2$, $p_3$ on $S^2$ (not all lying on the same great circle) are pairwise connected by three geodesics, that is, sections of great circles.

The main point to this paper is to prove, in Section 2, that it is impossible to define barycentric coordinate systems for geodesic triangles on $S^2$ that are consistent (see Definition 2.1), and which reduce, on edges, to Nielson’s (1989) definition of comparing ratios of geodesic lengths. Both of these (very mild) conditions are defined in Section 2 and are drawn directly from fundamental properties of barycentric coordinates in the plane. They merely serve to give some minimal structure to the problem of defining barycentric coordinates for geodesic triangles.

Given the main (negative) result of Section 2, we focus, in Section 3, on the question of defining barycentric coordinates for triangles on $S^2$ obtained by projections. That is, by considering an area preserving map $r : \mathbb{R}^2 \to S^2$, which imposes the barycentric coordinate system for a triangle in the plane onto a triangle $T_p$ on $S^2$. Such maps have been used, for example, by Foley (1990a), when considering interpolation problems on the sphere. For this reason, we consider the nature and properties of such a map $r$ and concentrate, in particular, on two special cases.
Of course, the problem with defining barycentric coordinates by projection is that the resulting triangle $T_p$ on $S^2$ cannot, from the results of Section 2, be a geodesic triangle and any projection scheme must produce seriously « deformed » triangles somewhere on $S^2$. Precisely where, depends largely upon the ab initio choices for a north pole, south pole, and a line of longitude connecting them — an « international date line » (i.d.l.). We conclude, in Section 4, by examining this problem for the two specific projections considered in Section 3, presenting some graphical results which show that defining barycentric coordinates via projection is an unreasonable approach to the problem.

2. BARYCENTRIC COORDINATE SYSTEMS FOR GEODESIC TRIANGLES ON $S^2$

In this section we show that there is no way of assigning barycentric coordinates to geodesic triangles on the sphere $S^2$, which inherit the properties of barycentric coordinates for planar triangles. To this end, let $p_1$, $p_2$, and $p_3$ be the vertices of a unique, non degenerate, geodesic triangle $T$ on $S^2$, so that the edge, $C_{ij}$ of $T$ opposite vertex $p_i$; $i = 1, 2, 3$, is a geodesic (part of a great circle) connecting $p_j$ and $p_k$, $i, j, k$, distinct; see figure 2.1.

Figure 2.1. — $p_1$, $p_2$, and $p_3$, are the vertices of a geodesic triangle $T$ on the sphere $S^2$. The edge, $C_{ij}$, opposite vertex $p_i$ is the geodesic connecting the other two vertices.

In order to consider the problem of defining barycentric coordinates on $T$ we must first of all define precisely what that means.

**Definition 2.1**: A system of barycentric coordinates on $T$ is a one-to-one correspondence between points $p \in T$ and ordered triples of real numbers $(b_1, b_2, b_3)$ such that:

a) $0 \leq b_i \leq 1$; $i = 1, 2, 3$, and $b_1 + b_2 + b_3 = 1$,

and

b) if $p \in C_i$ (so that $b_i = 0$), $b_j = d(p_k, p) / d(p_k, p_j)$; $i, j, k \in \{1, 2, 3\}$
all distinct, where \( d(p_k, q) \) is the geodesic distance from \( p_k \) to \( q \), measured along the geodesic \( C_r \).

Note: Condition \( b) \) in Definition 2.1 ensures that on the edge of a geodesic triangle, barycentric coordinates reduce to ratios of geodesic lengths. This is totally analogous to the situation for a triangle in the plane and is consistent with Nielson’s (1989) requirement for constructing smooth curves on the sphere.

This definition takes into account only one geodesic triangle \( T \). From the point of view of trying to define smooth surfaces over the sphere, we would need to consider a domain composed of several geodesic triangles. Therefore, in order for a barycentric coordinate system on \( T \) to be useful, we need to be able to assign coordinate systems to subtriangles of \( T \) that are « consistent » with the coordinates on \( T \) itself. More precisely, we have the following definition.

**Definition 2.2:** Let \( T^1 \subseteq T \) be a geodesic triangle on \( S^2 \) with vertices \( p_1^1, p_2^1, p_3^1 \), where, with respect to \( T \), \( p_i^1 \) has barycentric coordinates \( (b_{1,i}, b_{2,i}, b_{3,i}) \). A barycentric coordinate system on \( T^1 \) is consistent with that on \( T \) if for all \( p \in T^1 \)

\[
b_i = b_{1,i} + b_{2,i} + b_{3,i}; \quad i = 1, 2, 3,
\]

where \( (b_1, b_2, b_3) \) and \( (b_{1,i}, b_{2,i}, b_{3,i}) \) are, respectively, the barycentric coordinates of \( p \) with respect to the triangles \( T \) and \( T^1 \).

Given the (minimal) structure for a barycentric coordinate system on geodesic triangles on \( S^2 \) imposed by Definitions 2.1 and 2.2, we are in a position to present the following result, the main one of this paper.

**Theorem 2.3:** Let \( 0 \) be any open subset of \( S^2 \). It is impossible to define a scheme for assigning a barycentric coordinate system for every geodesic triangle \( T \in 0 \), so that if \( T^1 \) and \( T^2 \) are two such triangles, the barycentric coordinate systems on each are consistent.

**Proof:** Let \( 0 \) be an open subset of \( S^2 \) and suppose that there does exist an atlas of barycentric coordinate systems on \( 0 \). Further, let \( T \) be a non degenerate geodesic triangle in \( 0 \) with vertices \( p_1, p_2, p_3 \), and its own system of barycentric coordinates \( (b_1, b_2, b_3) \).

We claim that \( \{ p \in T : p = (b_1, 0.5, b_3) \} \) is the geodesic \( L \in T \) connecting the two points \( x = (0.5, 0.5, 0) \) and \( y = (0, 0.5, 0.5) \); see figure 2.2. To justify this claim, let \( T^1 \subseteq T \) denote the geodesic triangle with vertices \( x, p_2, y \). By assumption, this subtriangle has its own barycentric coordinate system \( (b_{1,i}, b_{2,i}, b_{3,i}) \) with respect to these vertices and in terms of this system \( L = \{ (b_{1,i}, 0, b_{3,i}) \} \).
Figure 2.2. — $p_1$, $p_2$, and $p_3$, are the vertices of a geodesic triangle, $x$ and $y$ are the midpoints of edges $C_3$ and $C_1$ respectively, and $L$ is the geodesic joining them.

From the consistency condition, it follows that

\[ b_1 = 0.5 b_1^{\perp}, \]
\[ b_2 = 0.5 b_1^{\perp} + b_2^{\perp} + 0.5 b_3^{\perp}, \]
\[ b_3 = 0.5 b_3^{\perp}. \]

Therefore, in terms of the barycentric coordinate system $(b_1, b_2, b_3)$ for $T$, a point $p$ is on $L$ if and only if $b_2 = 0.5 b_1^{\perp} + 0.5 b_3^{\perp} = b_1 + b_3$, since, on $L$, $b_2^{\perp} = 0$. However, since $b_1 + b_2 + b_3 = 1$, the condition $b_2 = b_1 + b_3$ is equivalent to $b_2 = 0.5$. Therefore, \{\(p \equiv (b_1, b_2, b_3) \in T: b_2 = 0.5\}\} is the geodesic $L$ as claimed.

Now, let $z$ be a point on $C_2$, the edge of the triangle $T$ connecting $p_1$ and $p_3$ and let $p$ be the intersection of the geodesic from $p_2$ to $z$ with the geodesic $L$; see figure 2.3.

Figure 2.3. — Given the setup of figure 2.2, $z$ is a fixed yet arbitrary point on the edge $C_2$ opposite vertex $p_2$. The geodesic joining $p_2$ and $z$ intersects the geodesic $L$ at the point $p$.

If $T^2$ is the geodesic triangle with vertices $z$, $p_2$, $p_3$, then, with respect to these vertices, $p$ has barycentric coordinates $(b_1^2, b_2^2, b_3^2)$, whilst with respect

vol. 26, n° 1, 1992
to the vertices $p_1$, $p_2$, $p_3$, of the macrotriangle $T$, $p$ has barycentric coordinates $(b_1, 0.5, b_3)$. Looking at the second barycentric coordinate, we have, using the consistency condition for $T^2 \subset T$,

$$0.5 = b_1(t)(0) + b_2(t)(1) + b_3(t)(0).$$  (2.1)

Therefore, in terms of the triangle $T^2$, the geodesic $L$ is characterized by the equation

$$b_2 = 0.5.$$  (2.2)

From Definition 2.1 this means that $p$ must be the midpoint of the geodesic from $p_2$ to $z$. However, this is not possible since $L$ and $C_2$ are geodesics and this contradiction completes the argument.

This final assertion is obvious in the case $p_1$ and $p_3$ are on the equator, $p_2$ is the north pole, and the points $x, y$ are both on the 45° N line of latitude, because then the line

$$L = \{(t, 0.5, 0.5 - t): 0 \leq t \leq 0.5\}$$

must be on the 45° N line of latitude since all points of $L$ must be equidistant from the north pole and the equator. However, that line is clearly not a geodesic on $S^2$, contradicting an earlier conclusion about $L$. In the general case for a triangle $T$ with vertices $p_1, p_2, p_3$, in arbitrary position, the proof of the assertion is a straightforward, but tedious, algebraic calculation. It is omitted for the sake of brevity. Q.E.D.

This negative result is the main one of this paper. Since geodesic triangles are the only ones intrinsic to the sphere, we feel that there is very little chance that the methods for triangular Bézier patches can be reasonably extended from $R^2$ to $S^2$. That said, it is worth examining the possibility of generating coordinate systems on $S^2$ by other means. In particular, by looking at barycentric coordinates for a planar triangle and seeing how these are affected by projections from $R^2$ to $S^2$. This is the question we consider in the remainder of the paper.

3. BARYCENTRIC COORDINATE SYSTEMS FOR $S^2$ USING PROJECTIONS

Given three point on $S^2$ that are linearly independent as vectors in $R^3$, we would like to describe the points of the resulting geodesic triangle in terms of barycentric coordinates. The result of Theorem 2.3 states that this cannot be done in a consistent way, but there are alternative methods for defining barycentric coordinate systems for « triangles » with edges that are not geodesics.
DEFINING BARYCENTRIC COORDINATES FOR THE SPHERE

Specifically, one approach is to choose a map

$$r : R^2 \rightarrow S^2$$

(3.1)

so that the vertices of a triangle $T \in R^2$ map to the given points on $S^2$, and use the natural barycentric coordinate system on $T$ to describe points of $r(T)$ on the sphere.

The problem with such an approach is that in general the edges of $r(T)$ are not geodesic, so that the « triangle » $r(T)$ looks distorted. In this Section we discuss the properties of two particular choices of the map $r$ in (3.1). In both cases, there is distortion in the « triangle » $r(T)$, which in certain situations is unacceptably severe.

3.1. Area Preserving Maps from $R^2$ to $S^2$

It is not possible to preserve length when mapping $R^2$ to $S^2$. In fact, if $0$ is any open set in $R^2$, then there is no length preserving map $r : 0 \rightarrow S^2$; see Berger (1987; section 18.4.4, p. 280). However, it is possible to impose the constraint that a map $r : R^2 \rightarrow S^2$ be area preserving. From the point of view of trying to define a barycentric coordinate system for triangles on the sphere, it is entirely reasonable to restrict our attention to maps with this property.

Let $f = f(u, v)$ and $g = g(u, v)$ be functions of the variables $u$ and $v$ and let $r$ map the $uv$ plane to $S^2$ by

$$r = (\sin f \cos g, \sin f \sin g, \cos f) .$$

(3.2)

If $A$ is a region in the $uv$ plane, then the area of $r(A)$ is

$$\iint_A \left| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right| du \, dv$$

(3.3)

so $r$ preserves area if

$$\left| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right| = 1 .$$

(3.4)

After some simplification we have

$$\left| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right| = \left| \frac{\partial f}{\partial u} \frac{\partial g}{\partial v} - \frac{\partial g}{\partial u} \frac{\partial f}{\partial v} \right| \sin f .$$

(3.5)

Therefore, choosing the functions $f$ and $g$ to satisfy the equation

$$\left| \frac{\partial f}{\partial u} \frac{\partial g}{\partial v} - \frac{\partial g}{\partial u} \frac{\partial f}{\partial v} \right| \sin f = 1$$

(3.6)

makes the map (3.2) area preserving onto the sphere $S^2$. 

vol. 26, n° 1, 1992
We now consider two particular choices for the map \( r \) in (3.2) for which (3.6) is satisfied.

### 3.2. Two specific area preserving maps

Let \( f = \cos^{-1} v \) and \( g = u \), so that

\[
  r_1(u, v) = \left( \sqrt{1 - v^2} \cos u, \sqrt{1 - v^2} \sin u, v \right)
\]

is an area preserving map from the rectangle \( \{ (u, v) : |u| \leq \pi, |v| \leq 1 \} \), onto the sphere \( S^2 \).

One problem with this map is that « triangles » on the sphere near the poles can be very distorted. Another difficulty is that the map is not locally invertible at either pole because the preimages of the poles are not discrete. The entire segment in the domain given by \( v = 1 \) maps onto the « north » pole and, similarly, the line segment \( v = -1 \) maps onto the opposite pole.

It is possible to decrease the distortion at the poles in the following way. In (3.7) replace \( v \) by \( \phi(v) \) and \( u \) by \( u/\phi'(v) \), where \( \phi(v) \) is any \( C^1 \) function such that:

a) \( \phi : [-1, 1] \rightarrow [-1, 1] \),

and

b) \( \phi'(v) \neq 0 ; v \in (-1, 1) \).

With these changes the map (3.7) becomes

\[
  r_2(u, v) = \left( \sqrt{1 - \phi^2} \cos \left( u/\phi' \right), \sqrt{1 - \phi^2} \sin \left( u/\phi' \right), \phi(v) \right).
\]

This map is still area preserving. It is possible to choose \( \phi \) appropriately so as to produce less distortion at the poles than (3.7). In particular, let \( \phi(v) = \sin v \), to give

\[
  r_2(u, v) = \left( \cos v \cos \left( u/\cos v \right), \cos v \sin \left( u/\cos v \right), \sin v \right),
\]

which maps the region \( \{ (u, v) : |u| \leq \pi \cos v, |v| < \pi/2 \} \) onto \( S^2 \)-{poles}. However, since the limits are well defined we define \( r_2(0, \pi/2) \) and \( r_2(0, -\pi/2) \) to be the north and south poles respectively.

The « inverse » of the map \( r_2 \) is the Samson-Flamsteed (Sinusoidal) projection; Berger (1987, p. 271). The distortion in « triangles » near the poles is not nearly as severe as that in the map (3.7). However, there is a considerable amount of distortion near the « international date line » (i.d.l.), that is, \( r_2(\pi \cos v, v) \); \( -\pi/2 < v < \pi/2 \).

In order to analyze the possible usefulness of projection schemes for defining barycentric coordinate systems on \( S^2 \), we conclude with some examples illustrating the types of triangles that the maps (3.7) and (3.8) produce.
4. EXAMPLES AND CONCLUSIONS

To use any projection method for defining barycentric coordinates on \( S^2 \), it is first of all necessary to prescribe north and south poles as well as an i.d.l. This is needed in order to orient the sphere and define any map of the form (3.7a). This pre-processing step is, in itself, a limitation of the approach since it imposes, a priori, an orientation of the sphere. That aside, we now consider some of the relevant properties of the map (3.7a) for our purposes.

The issue that arises when using a projection method is basically that of assessing how triangles in the planar domain are distorted when mapped onto the sphere. More precisely, given three points on the sphere (not all lying on one great circle) how sensitive is the shape of the triangle \( T_p \) on the sphere, determined by this projection, relative to the \( a \) priori choice of the north and south poles, and the i.d.l.? Ideally, we would obviously like the boundaries of \( T_p \) to be as close to geodesics as possible, since only geodesic triangles are intrinsic to \( S^2 \) and, therefore, independent of the choice of the date line.

In order to examine the question we consider five examples for triangles on the sphere determined by the projections (3.7) and (3.8). In each case we choose three points, \( p_1, p_2, p_3 \), on the sphere and produce the triangles determined by these vertices and these projections. More precisely, for each projection we produce the triangle on \( S^2 \) with vertices \( p_1, p_2, p_3 \) that is the image of the planar triangle in the domain whose vertices are the preimages of the \( p_1, p_2, p_3 \). The five examples cover cases where the points \( p_1, p_2, p_3 \) are chosen to be:

1) near a pole,
2) away from the poles but near the i.d.l.,
3) at middle latitudes, away from the i.d.l.,
4) near the equator, away from the i.d.l.,
5) the vertices of a geodesic triangle whose interior contains a pole.

These cases highlight all the interesting properties of the projections (3.7) and (3.8).

As figures 4.1a)-4.5b) indicate, the map (3.7) produces distortion near the poles, whilst (3.8) gives rise to distorted triangles near the i.d.l. In choosing different functions \( \phi(\nu) \) in (3.7a), we only succeed in repositioning the region of the sphere where the projection method produces badly distorted triangles. \textit{It is not possible to define a projection method of the form (3.7a) which is free of such distortion over the entire sphere}, and this is a severe limitation to the approach of using projection methods.
Figure 4.1a). — The distortion in the triangle under the map $r_1$ defined by (3.7) is due to the proximity to the pole. The axis of the sphere as well as the « international date line » are shown to orient the image.

Figure 4.1b). — The same three vertices on the sphere as in figure 4.1a) but in this case the triangle is generated by the sinusoidal map $r_2$ in (3.8).

Figure 4.2a). — The three vertices on the sphere are away from the poles and the map $r_1$ produces a « reasonable » triangle. It is similar to the geodesic triangle defined by the same three vertices.

Figure 4.2b). — The same three vertices as in figure 4.2a) produce a severely distorted triangle when using the sinusoidal map (3.8), because of the close proximity of the vertices to the « international date line ».

$M^2$ AN Modélisation mathématique et Analyse numérique
Mathematical Modelling and Numerical Analysis
Figure 4.3a). — The triangle vertices are away from the poles and the map (3.7) produces a reasonable triangle in this case.

Figure 4.3b). — For the same three vertices as in figure 4.3a), the sinusoidal map (3.8) also produces a reasonable triangle, since the vertices are not close to the « international date line ».

Figure 4.4a). — As we move further away from the poles, the map (3.7) produces triangles with less distortion. The distortion is minimized at the equator.

Figure 4.4b). — The vertices in figure 4.4a) are also well away from the date line. This means that the sinusoidal map (3.8) also performs well and in fact, the triangle shown in this case is very similar to that in figure 4.4a).
Figure 4.5a). — The map (3.7) does not produce a triangle whose interior contains the pole. The geodesic triangle for the same three vertices would, so that the map gives unacceptable distortion.

Figure 4.5b). — Using the same vertices as in figure 4.5a) produces a comparable level of distortion in the triangle generated by the sinusoidal map (3.8).

Based on these and other examples, we feel that this method for defining barycentric coordinates on \( S^2 \) has no practical value. These results, together with those of Section 2, suggest that it is unlikely that the theory of triangular Bézier patches can be generalized directly to the sphere.

ACKNOWLEDGMENTS

The work of A. Worsey was supported by the National Science Foundation grant DMS-8803257. That of J. Brown was supported in part by the North Carolina Board of Science and Technology grant 89SE09. The authors thank Andy Groome for his help in creating some of the figures, and the referees for their comments.

REFERENCES


M² AN Modélisation mathématique et Analyse numérique
Mathematical Modelling and Numerical Analysis


