

On the injectivity of mean value mappings between quadrilaterals

Michael S. Floater* Georg Muntingh†

March 26, 2025

Abstract

Mean value coordinates can be used to map one polygon into another, with application to computer graphics and curve and surface modelling. In this paper we show that if the polygons are quadrilaterals, and if the target quadrilateral is convex, then the mapping is injective.

Keywords: Barycentric mapping, Mean value coordinates, injectivity, Jacobian determinant.

Math Subject Classification: 65D17, 26B10

1 Introduction

Barycentric mapping has emerged as a convenient tool for deforming shapes when modelling and processing geometry. For example, a planar curve can be deformed by enclosing it in a polygon and deforming the polygon. Using some choice of generalized barycentric coordinates (GBCs) to represent each point of the polygon, the polygon can be deformed by moving its vertices. In effect, this defines a mapping $\mathbf{f} : P \rightarrow Q$ from the initial polygon P , the ‘domain’ polygon, to a new polygon Q , the ‘target’ polygon. An early paper proposing such barycentric mapping both in $\mathbb{R}^2$ and higher Euclidean dimension is that

*Department of Mathematics, University of Oslo, PO Box 1053, Blindern, 0316 Oslo, Norway, *email:* michaelf@math.uio.no

†SINTEF, PO Box 124 Blindern, 0314 Oslo, Norway, *email:* Georg.Muntingh@sintef.no

of Warren [23], who used a generalization of Wachspress' rational coordinates for the barycentric representation of points in a convex polygon or polytope. Later, the discovery of mean value coordinates enhanced the practicality of barycentric mapping since these coordinates allow the domain to be non-convex [5, 4, 15, 8, 13, 18, 24].

In many applications of curve deformation we would like to avoid the deformed curve having self-intersections. Thus, we would like some kind of condition on P and Q that guarantees that $\mathbf{f}$ is injective. This was the motivation for the theoretical study made in [9], the main result of which was to show that if P and Q are both convex polygons, then the mapping determined by Wachspress's coordinates is injective. It was also shown there, by way of counterexamples, that, with P and Q again convex, the mean value mapping is not always injective when the number of vertices in P (and in Q) is five or more. This drawback of mean value coordinates was the motivation for the method proposed in [22], in which an injective mapping is constructed as the functional composition of several mean value mappings, each of which is an injective perturbation of the identity mapping.

The results of [9] left open the important case: is a mean value mapping injective when P and Q are convex quadrilaterals? The purpose of this paper is to settle this question. We will show in fact a more general result.

Theorem 1. *If P is any quadrilateral, convex or non-convex, and if Q is convex, then the mean value mapping $\mathbf{f} : P \rightarrow Q$ is injective.*

Figures 1 and 2 illustrate the theorem.

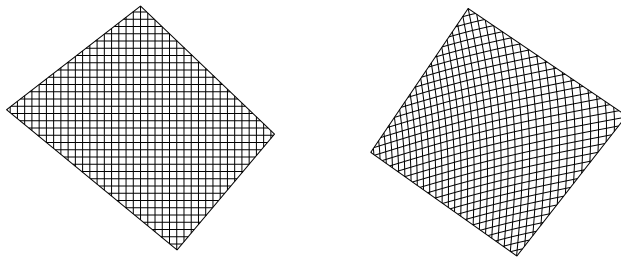


Figure 1: MV mapping, convex to convex.

So in the quadrilateral case, mean value mappings have the same injectivity property as harmonic mappings; see [19, 17, 3].

The essential part of the proof of Theorem 1 is to show that when the vertices of P and Q have the same orientation, the Jacobian (determinant)

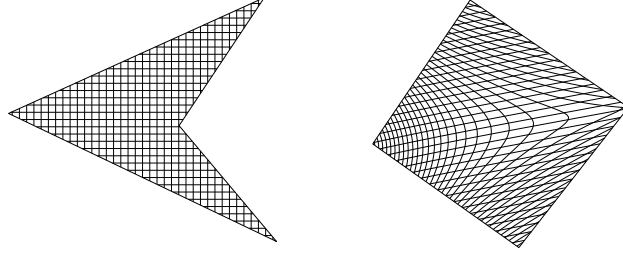


Figure 2: MV mapping, concave to convex.

$J(\mathbf{f})$ is positive both in P and at all boundary points of P except the four vertices. This requires computing the gradients of the mean value coordinates. We will derive in Theorem 2 a new, simple formula for the gradients of the homogeneous form of the coordinates. This formula applies to an arbitrary polygon and might be useful in other applications, for example, to derive bounds on the gradients for finite element applications; see [20, 7, 10].

2 Barycentric mappings

Let $P \subset \mathbb{R}^2$ be a polygon, viewed as an open set, with vertices $\mathbf{p}_1, \dots, \mathbf{p}_n$, ordered anticlockwise. We denote by ∂P the boundary of P and by $\overline{P}$ its closure. Let $\phi_1, \dots, \phi_n : \overline{P} \rightarrow \mathbb{R}$ be continuous functions such that

$$\sum_{i=1}^n \phi_i(\mathbf{x}) = 1, \quad \mathbf{x} \in \overline{P}, \quad (1)$$

$$\sum_{i=1}^n \phi_i(\mathbf{x}) \mathbf{p}_i = \mathbf{x}, \quad \mathbf{x} \in \overline{P}, \quad (2)$$

and such that if

$$\mathbf{x} = (1 - \mu) \mathbf{p}_\ell + \mu \mathbf{p}_{\ell+1} \quad (3)$$

for some ℓ and some $\mu \in [0, 1]$, with indices treated cyclically ($\mathbf{p}_{n+1} := \mathbf{p}_1$ and so on), then

$$\phi_\ell(\mathbf{x}) = 1 - \mu, \quad \phi_{\ell+1}(\mathbf{x}) = \mu, \quad \text{and} \quad \phi_i(\mathbf{x}) = 0, \quad i \neq \ell, \ell + 1. \quad (4)$$

Then $\phi_1, \dots, \phi_n$ are *generalized barycentric coordinates* (GBCs) for P .

Next suppose $Q \subset \mathbb{R}^2$ is another polygon, with vertices $\mathbf{q}_1, \dots, \mathbf{q}_n$, ordered anticlockwise. Then we call the mapping $\mathbf{f} : \overline{P} \rightarrow \mathbb{R}^2$ defined by

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^n \phi_i(\mathbf{x}) \mathbf{q}_i, \quad \mathbf{x} \in \overline{P}, \quad (5)$$

a *barycentric mapping*.

Equation (1) implies that for $\mathbf{x} \in \overline{P}$, the point $\mathbf{f}(\mathbf{x})$ is an affine combination of the points $\mathbf{q}_i$. If the ϕ_i are non-negative then this combination is also convex and if in addition Q is convex then $\mathbf{f}(\overline{P}) \subseteq \overline{Q}$. Moreover, if the ϕ_i are positive in P , then $\mathbf{f}(P) \subseteq Q$.

The linear precision property (2) implies that if $Q = P$ then $\mathbf{f}$ is the identity mapping, and by the continuity of the ϕ_i , $\mathbf{f}$ is continuous and so if Q is a perturbation of P then $\mathbf{f}$ is a perturbation of the identity mapping. By (4), if $\mathbf{x}$ is the boundary point in (3) then

$$\mathbf{f}(\mathbf{x}) = (1 - \mu) \mathbf{q}_\ell + \mu \mathbf{q}_{\ell+1}. \quad (6)$$

Thus $\mathbf{f}$ maps ∂P to ∂Q in a piecewise linear fashion, mapping vertices and edges of ∂P to corresponding vertices and edges of ∂Q .

The question we will address is whether $\mathbf{f}$ is injective. If both P and Q are convex and if the GBCs are Wachspress coordinates then $\mathbf{f}$ is C^∞ on $\overline{P}$ (technically, $\mathbf{f}$ is C^∞ in an open set containing $\overline{P}$), and it was shown in [9] that the Jacobian of $\mathbf{f}$ is positive on $\overline{P}$. Then, by a theorem of [16], the positivity of $J(\mathbf{f})$ on $\overline{P}$ implies that $\mathbf{f}$ is injective.

We will use a similar approach to prove Theorem 1. But first let us observe that there is a simple formula for $J(\mathbf{f})$ at a vertex of P for *any* barycentric mapping $\mathbf{f}$ as long as $\mathbf{f}$ is C^1 at the vertex. Recall that the Jacobian of $\mathbf{f}$ is the determinant

$$J(\mathbf{f}) := |\mathbf{f}_x, \mathbf{f}_y| = \begin{vmatrix} f_x & f_y \\ g_x & g_y \end{vmatrix},$$

where

$$h_x := \frac{\partial h}{\partial x}, \quad h_y := \frac{\partial h}{\partial y},$$

and $\mathbf{x} = (x, y)$ and $\mathbf{f}(\mathbf{x}) = (f(\mathbf{x}), g(\mathbf{x}))$. For points $\mathbf{p} = (p^1, p^2)$, $\mathbf{q} = (q^1, q^2)$, $\mathbf{r} = (r^1, r^2)$ in $\mathbb{R}^2$ let

$$A(\mathbf{p}, \mathbf{q}, \mathbf{r}) := \frac{1}{2} \begin{vmatrix} 1 & 1 & 1 \\ p^1 & q^1 & r^1 \\ p^2 & q^2 & r^2 \end{vmatrix}, \quad (7)$$

the signed area of the triangle $[\mathbf{p}, \mathbf{q}, \mathbf{r}]$.

We will often use the two-dimensional cross product,

$$\mathbf{a} \times \mathbf{b} := \begin{vmatrix} a^1 & b^1 \\ a^2 & b^2 \end{vmatrix} = a^1 b^2 - a^2 b^1$$

for vectors $\mathbf{a} = (a^1, a^2)$ and $\mathbf{b} = (b^1, b^2)$. So, for example,

$$J(\mathbf{f}) = \mathbf{f}_x \times \mathbf{f}_y, \quad A(\mathbf{p}, \mathbf{q}, \mathbf{r}) = \frac{1}{2}(\mathbf{q} - \mathbf{p}) \times (\mathbf{r} - \mathbf{p}).$$

Lemma 1. *If $\mathbf{f} : \overline{P} \rightarrow \mathbb{R}^2$ is any barycentric mapping that is C^1 at $\mathbf{p}_i$ then*

$$J(\mathbf{f})(\mathbf{p}_i) = \frac{A(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{q}_{i+1})}{A(\mathbf{p}_{i-1}, \mathbf{p}_i, \mathbf{p}_{i+1})}.$$

Proof. At a point where $\mathbf{f}$ is C^1 , it has a directional derivative $D_{\mathbf{v}}\mathbf{f}$ in any vector direction $\mathbf{v} = (v^1, v^2)$ given by

$$D_{\mathbf{v}}\mathbf{f} := \lim_{\epsilon \rightarrow 0} \frac{\mathbf{f}(\mathbf{x} + \epsilon \mathbf{v}) - \mathbf{f}(\mathbf{x})}{\epsilon} = v^1 \mathbf{f}_x + v^2 \mathbf{f}_y.$$

Then at $\mathbf{p}_i$,

$$D_{\mathbf{p}_i - \mathbf{p}_{i-1}}\mathbf{f} = (\mathbf{p}_i^1 - \mathbf{p}_{i-1}^1)\mathbf{f}_x + (\mathbf{p}_i^2 - \mathbf{p}_{i-1}^2)\mathbf{f}_y, \quad (8)$$

$$D_{\mathbf{p}_{i+1} - \mathbf{p}_i}\mathbf{f} = (\mathbf{p}_{i+1}^1 - \mathbf{p}_i^1)\mathbf{f}_x + (\mathbf{p}_{i+1}^2 - \mathbf{p}_i^2)\mathbf{f}_y, \quad (9)$$

and by the piecewise linearity of $\mathbf{f}$ on ∂P given by (3) and (6),

$$\begin{aligned} D_{\mathbf{p}_{i+1} - \mathbf{p}_i}\mathbf{f} &= \lim_{\mu \rightarrow 0+} \frac{\mathbf{f}(\mathbf{p}_i + \mu(\mathbf{p}_{i+1} - \mathbf{p}_i)) - \mathbf{f}(\mathbf{p}_i)}{\mu} \\ &= \lim_{\mu \rightarrow 0+} \frac{(\mathbf{q}_i + \mu(\mathbf{q}_{i+1} - \mathbf{q}_i)) - \mathbf{q}_i}{\mu} = \mathbf{q}_{i+1} - \mathbf{q}_i, \end{aligned}$$

and similarly,

$$D_{\mathbf{p}_i - \mathbf{p}_{i-1}}\mathbf{f} = \mathbf{q}_i - \mathbf{q}_{i-1}.$$

So from (8–9) we deduce that

$$\begin{aligned} (\mathbf{q}_i - \mathbf{q}_{i-1}) \times (\mathbf{q}_{i+1} - \mathbf{q}_i) &= (D_{\mathbf{p}_i - \mathbf{p}_{i-1}}\mathbf{f}) \times (D_{\mathbf{p}_{i+1} - \mathbf{p}_i}\mathbf{f}) \\ &= ((\mathbf{p}_i^1 - \mathbf{p}_{i-1}^1)(\mathbf{p}_{i+1}^2 - \mathbf{p}_i^2) - (\mathbf{p}_i^2 - \mathbf{p}_{i-1}^2)(\mathbf{p}_{i+1}^1 - \mathbf{p}_i^1))(\mathbf{f}_x \times \mathbf{f}_y), \end{aligned}$$

and so

$$J(\mathbf{f}) = \frac{(\mathbf{q}_i - \mathbf{q}_{i-1}) \times (\mathbf{q}_{i+1} - \mathbf{q}_i)}{(\mathbf{p}_i - \mathbf{p}_{i-1}) \times (\mathbf{p}_{i+1} - \mathbf{p}_i)}.$$

□

Suppose now that $\mathbf{f} : \bar{P} \rightarrow \mathbb{R}^2$ is a barycentric mapping that is C^1 on $\bar{P}$. If both P and Q are convex then at each vertex $\mathbf{p}_i$, both $A(\mathbf{p}_{i-1}, \mathbf{p}_i, \mathbf{p}_{i+1})$ and $A(\mathbf{q}_{i-1}, \mathbf{q}_i, \mathbf{q}_{i+1})$ are positive and the lemma implies that $J(\mathbf{f})$ is positive at $\mathbf{p}_i$. This agrees with the Wachspress case studied in [9]. On the other hand, suppose that P is non-convex and Q convex. Then P must have a non-convex vertex, i.e. a vertex $\mathbf{p}_i$ with $A(\mathbf{p}_{i-1}, \mathbf{p}_i, \mathbf{p}_{i+1}) < 0$. Then $J(\mathbf{f})$ is negative at $\mathbf{p}_i$ while it is positive at the convex vertices of P . We conclude that $J(\mathbf{f})$ must change sign in ∂P and $\mathbf{f}$ cannot be injective. The same conclusion can be drawn in the case that P is convex and Q is concave. A similar observation concerning polynomial mappings from the unit square to a concave quadrilateral was made in [12].

With these facts in mind, it may seem surprising that the mean value mapping from a concave quadrilateral to a convex quadrilateral can be injective. But this does not contradict Lemma 1 because mean value coordinates are *not* C^1 at the polygon vertices. This lack of smoothness can thus be viewed as an *advantage* of mean value coordinates, compared with, for example, Wachspress coordinates. This same lack of smoothness must also be inherent in harmonic coordinates. A referee pointed out to us that a simple argument in [1, Figure 4] shows that if $\mathbf{p}_i$ is a non-convex vertex and the coordinates are non-negative in P (which is the case for harmonic coordinates) then ϕ_i cannot be C^1 at $\mathbf{p}_i$.

Suppose now that $\mathbf{f} : \bar{P} \rightarrow \mathbb{R}^2$ is a barycentric mapping and that $\mathbf{x} \in \bar{P}$ is a point at which f is C^1 . A useful identity was derived in [9] that relates $J(\mathbf{f})(\mathbf{x})$ to the signed areas of triangles formed by the vertices $\mathbf{q}_i$ of Q . For differentiable bivariate functions a, b, c , let $D(a, b, c)$ denote the 3×3 determinant

$$D(a, b, c) := \begin{vmatrix} a & b & c \\ a_x & b_x & c_x \\ a_y & b_y & c_y \end{vmatrix}. \quad (10)$$

Then [9, Lemma 1],

$$J(\mathbf{f})(\mathbf{x}) = 2 \sum_{1 \leq i < j < k \leq n} D(\phi_i, \phi_j, \phi_k)(\mathbf{x}) A(\mathbf{q}_i, \mathbf{q}_j, \mathbf{q}_k). \quad (11)$$

From this it follows that if Q is convex and if

$$D(\phi_i, \phi_j, \phi_k)(\mathbf{x}) \geq 0, \quad 1 \leq i < j < k \leq n, \quad (12)$$

with strict inequality for at least one choice of i, j, k , then $J(\mathbf{f})(\mathbf{x}) > 0$.

3 Mean value coordinates

Consider now the mean value (MV) coordinates [5, 13]. These can be defined as follows; see Figure 3. For $\mathbf{x} \in P$,

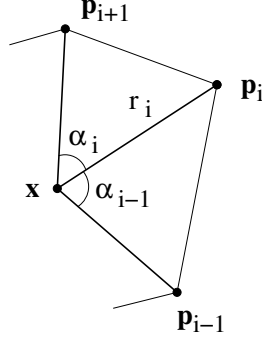


Figure 3: Notation for mean value coordinates.

$$\phi_i(\mathbf{x}) := w_i(\mathbf{x})/W(\mathbf{x}), \quad (13)$$

where

$$w_i := \frac{t_{i-1} + t_i}{r_i}, \quad W := \sum_{i=1}^n w_i, \quad (14)$$

and where r_i is the Euclidean distance $r_i := \|\mathbf{p}_i - \mathbf{x}\|$, and t_i is the half-angle tangent $t_i := \tan(\alpha_i/2)$. The angle $\alpha_i \in (-\pi, \pi)$ is the *signed* angle at $\mathbf{x}$ in the triangle $[\mathbf{x}, \mathbf{p}_i, \mathbf{p}_{i+1}]$, i.e.,

$$\text{sgn}(\alpha_i) := \text{sgn}(\mathbf{e}_i \times \mathbf{e}_{i+1}),$$

where $\mathbf{e}_i$ is the unit vector $\mathbf{e}_i := (\mathbf{p}_i - \mathbf{x})/r_i$.

It was shown in [13] that the MV coordinates $\phi_1, \dots, \phi_n$ are GBCs, i.e., they satisfy the linear precision properties (1) and (2) in P , and they have a unique continuous extension to ∂P , satisfying the Lagrange property (4) on ∂P .

Due to the anticlockwise ordering of the $\mathbf{p}_i$, $\alpha_i > 0$ in P in the case that P is convex, but for general P , the sign of α_i can be any value in $\{-1, 0, 1\}$.

The half-angle tangent t_i comes from the integral definition of mean value coordinates for positive α_i of [5], where it appears in the form

$$t_i = \frac{1 - c_i}{s_i},$$

where $s_i := \sin(\alpha_i)$ and $c_i := \cos(\alpha_i)$. We can therefore compute t_i without needing to evaluate trigonometric functions by the formula

$$t_i = \frac{1 - \mathbf{e}_i \cdot \mathbf{e}_{i+1}}{\mathbf{e}_i \times \mathbf{e}_{i+1}}, \quad (15)$$

when the α_i are positive. This formula is also correct for $\alpha_i < 0$, but requires defining the special case $t_i := 0$ if $\alpha_i = 0$. This can be avoided by using the alternative formula

$$t_i = \frac{s_i}{1 + c_i} = \frac{\mathbf{e}_i \times \mathbf{e}_{i+1}}{1 + \mathbf{e}_i \cdot \mathbf{e}_{i+1}}. \quad (16)$$

We note that an alternative way of expressing ϕ_i even when $\mathbf{x}$ is a boundary point was proposed recently in [11], where the issue of how to compute ϕ_i from the point of view of numerical stability has been studied in depth.

4 Gradients of mean value coordinates

In order to apply the condition (12) to MV coordinates, we need a suitable formula for their gradients. Applying the quotient rule to (13) gives

$$\nabla \phi_i = \frac{\nabla w_i}{W} - \frac{w_i \nabla W}{W^2},$$

and so it is sufficient to find a formula for ∇w_i . A formula for ∇w_i was derived in [6, Sec. 4.1], but it expresses ∇w_i as a linear combination of four vectors. We now derive a more compact formula for ∇w_i as a linear combination of just two vectors.

Here and later we will use the well known half-angle tangent identities

$$s_j = \frac{2t_j}{1 + t_j^2}, \quad c_j = \frac{1 - t_j^2}{1 + t_j^2}. \quad (17)$$

For any vector $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$, we define the vector

$$\mathbf{a}^\perp := (-a_2, a_1),$$

which is the rotation of $\mathbf{a}$ through a positive angle of $\pi/2$. We will show:

Theorem 2.

$$\nabla w_i = \frac{1}{2r_i} (u_{i-1} \mathbf{e}_{i-1}^\perp - u_i \mathbf{e}_{i+1}^\perp), \quad (18)$$

where

$$u_j := \left(\frac{1}{r_j} + \frac{1}{r_{j+1}} \right) (1 + t_j^2).$$

Notice that $u_j > 0$ and so the theorem shows that the vector ∇w_i is a positive linear combination of the two vectors $\mathbf{e}_{i-1}^\perp$ and $-\mathbf{e}_{i+1}^\perp$.

Proof. Following [6, Sec. 4.1], we can compute ∇w_i by writing w_i as

$$w_i = \frac{t_{i-1}}{r_i} + \frac{t_i}{r_i},$$

and applying the quotient rule to each part. It was shown in [6, Sec. 4.1] that

$$\nabla t_j = \frac{t_j}{s_j} \left(\frac{\mathbf{e}_j^\perp}{r_j} - \frac{\mathbf{e}_{j+1}^\perp}{r_{j+1}} \right),$$

which was then combined with the fact that $\nabla r_i = -\mathbf{e}_i$ to show that

$$\nabla \left(\frac{t_{i-1}}{r_i} \right) = \frac{t_{i-1}}{r_i s_{i-1}} \left(\frac{\mathbf{e}_{i-1}^\perp}{r_{i-1}} - \frac{\mathbf{e}_i^\perp}{r_i} \right) + \frac{t_{i-1}}{r_i^2} \mathbf{e}_i, \quad (19)$$

$$\nabla \left(\frac{t_i}{r_i} \right) = \frac{t_i}{r_i s_i} \left(\frac{\mathbf{e}_i^\perp}{r_i} - \frac{\mathbf{e}_{i+1}^\perp}{r_{i+1}} \right) + \frac{t_i}{r_i^2} \mathbf{e}_i. \quad (20)$$

Now let us observe that $\mathbf{e}_i$ and $\mathbf{e}_i^\perp$ form an orthonormal system of vectors in $\mathbb{R}^2$, and so

$$\mathbf{e}_{i-1}^\perp = \cos \left(\frac{\pi}{2} - \alpha_{i-1} \right) \mathbf{e}_i + \sin \left(\frac{\pi}{2} - \alpha_{i-1} \right) \mathbf{e}_i^\perp = s_{i-1} \mathbf{e}_i + c_{i-1} \mathbf{e}_i^\perp.$$

Solving this for $\mathbf{e}_i$ gives

$$\mathbf{e}_i = \frac{1}{s_{i-1}} \mathbf{e}_{i-1}^\perp - \frac{c_{i-1}}{s_{i-1}} \mathbf{e}_i^\perp.$$

Then substituting this into (19), we deduce

$$\nabla \left(\frac{t_{i-1}}{r_i} \right) = \frac{t_{i-1}}{r_i s_{i-1}} \left(\frac{1}{r_{i-1}} + \frac{1}{r_i} \right) \mathbf{e}_{i-1}^\perp - \frac{t_{i-1}}{r_i^2 s_{i-1}} (1 + c_{i-1}) \mathbf{e}_i^\perp,$$

and using the facts

$$\frac{t_j}{s_j} = \frac{1 + t_j^2}{2}, \quad \frac{t_j(1 + c_j)}{s_j} = 1,$$

we find

$$\nabla \left(\frac{t_{i-1}}{r_i} \right) = \frac{u_{i-1}}{2r_i} \mathbf{e}_{i-1}^\perp - \frac{\mathbf{e}_i^\perp}{r_i^2}. \quad (21)$$

Similarly,

$$\mathbf{e}_{i+1}^\perp = \cos\left(\frac{\pi}{2} + \alpha_i\right)\mathbf{e}_i + \sin\left(\frac{\pi}{2} + \alpha_i\right)\mathbf{e}_i^\perp = -s_i\mathbf{e}_i + c_i\mathbf{e}_i^\perp,$$

and so

$$\mathbf{e}_i = \frac{c_i}{s_i}\mathbf{e}_i^\perp - \frac{1}{s_i}\mathbf{e}_{i+1}^\perp,$$

and substituting this into (20) leads to

$$\nabla\left(\frac{t_i}{r_i}\right) = \frac{t_i}{r_i^2 s_i}(1 + c_i)\mathbf{e}_i^\perp - \frac{t_i}{r_i s_i}\left(\frac{1}{r_i} + \frac{1}{r_{i+1}}\right)\mathbf{e}_{i+1}^\perp,$$

and therefore,

$$\nabla\left(\frac{t_i}{r_i}\right) = \frac{\mathbf{e}_i^\perp}{r_i^2} - \frac{u_i}{2r_i}\mathbf{e}_{i+1}^\perp. \quad (22)$$

Taking the sum of (21) and (22), the term $\mathbf{e}_i^\perp/r_i^2$ cancels out and we are left with the formula (18). $\square$

We initially derived (18) using the integral formula for w_i and differentiating the integrand. However, this method of proof is rather long, and we saw, with the benefit of hindsight, that (18) follows more easily from the calculations of [6].

5 Quadrilateral MV mappings

We will assume *for the rest of the paper* that P is a quadrilateral, i.e., $n = 4$, and that $\phi_1, \phi_2, \phi_3, \phi_4$ are the MV coordinates with respect to P . Then the determinant condition (12) reduces to

$$D_i(\mathbf{x}) := D(\phi_{i-1}, \phi_i, \phi_{i+1})(\mathbf{x}) \geq 0, \quad i = 1, 2, 3, 4, \quad (23)$$

with strict inequality for at least one i . In this section we show that this condition holds for all $\mathbf{x} \in P$.

We start with an observation about the signs of the t_i . By the definition of the α_i , $\cos(\alpha_i/2) > 0$ while $\sin(\alpha_i/2)$ and t_i have the same sign as α_i . Thus t_i may be negative but the sum $t_{i-1} + t_i$ is positive.

Lemma 2. $t_{i-1} + t_i > 0$ for all $i = 1, 2, 3, 4$.

From this lemma it follows that also $w_i > 0$ and $\phi_i > 0$, a property that was observed in [14].

Proof. If P is convex, then $\alpha_i > 0$ for all i and

$$t_i = \frac{\sin(\alpha_i/2)}{\cos(\alpha_i/2)} > 0.$$

It remains to consider the case that P is non-convex, as in Figure 4. If $\mathbf{x}$ is

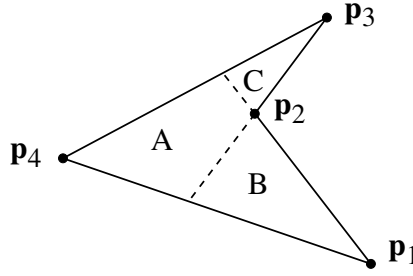


Figure 4: Non-convex quadrilateral.

in A (the kernel of P), then $\alpha_i > 0$ for all i and again $t_i > 0$. Suppose next that $\mathbf{x} \in B$. Then $\alpha_2 < 0$ and $\alpha_1, \alpha_3, \alpha_4 > 0$, and further $|\alpha_2| < |\alpha_1|$ and $|\alpha_2| < |\alpha_3|$. Therefore

$$\alpha_1 + \alpha_2 > 0, \quad \alpha_2 + \alpha_3 > 0,$$

implying

$$0 < \alpha_{i-1} + \alpha_i < 2\pi, \quad i = 1, 2, 3, 4,$$

and so

$$\sin\left(\frac{\alpha_{i-1} + \alpha_i}{2}\right) > 0, \quad i = 1, 2, 3, 4.$$

Hence

$$t_{i-1} + t_i = \frac{\sin\left(\frac{\alpha_{i-1} + \alpha_i}{2}\right)}{\cos\left(\frac{\alpha_{i-1}}{2}\right) \cos\left(\frac{\alpha_i}{2}\right)} > 0.$$

The argument for $\mathbf{x} \in C$ is similar. □

Another ingredient in the proof is the following trigonometric identity, relating t_1, t_2, t_3, t_4 .

Lemma 3.

$$t_4 = \frac{p}{q}, \quad \text{where} \quad \begin{aligned} p &:= t_1 + t_2 + t_3 - t_1 t_2 t_3, \\ q &:= t_1 t_2 + t_1 t_3 + t_2 t_3 - 1 > 0. \end{aligned}$$

This is a special case of a more general identity for half-angle tangents involving symmetric polynomials [2], but we need to take care that $q \neq 0$.

Proof. Let $s'_i := \sin(\alpha_i/2)$ and $c'_i := \cos(\alpha_i/2)$ for $i = 1, 2, 3, 4$. Since $-\pi < \alpha_i < \pi$, $c'_i > 0$.

Since $\sum_{i=1}^4 (\alpha_i/2) = \pi$, applying the angle-sum formulas for sine and cosine yields

$$\begin{aligned} s'_4 &= \sin\left(\pi - \frac{\alpha_4}{2}\right) = \sin\left(\frac{\alpha_1}{2} + \frac{\alpha_2}{2} + \frac{\alpha_3}{2}\right) \\ &= c'_2 c'_3 s'_1 + c'_1 c'_3 s'_2 + c'_1 c'_2 s'_3 - s'_1 s'_2 s'_3, \\ c'_4 &= -\cos\left(\pi - \frac{\alpha_4}{2}\right) = -\cos\left(\frac{\alpha_1}{2} + \frac{\alpha_2}{2} + \frac{\alpha_3}{2}\right) \\ &= c'_3 s'_1 s'_2 + c'_2 s'_1 s'_3 + c'_1 s'_2 s'_3 - c'_1 c'_2 c'_3. \end{aligned}$$

Since $t_4 = s'_4/c'_4$, we obtain $t_4 = p/q$ by dividing both s'_4 and c'_4 by $c'_1 c'_2 c'_3$, which is positive. Moreover,

$$p = \frac{s'_4}{c'_1 c'_2 c'_3} \quad \text{and} \quad q = \frac{c'_4}{c'_1 c'_2 c'_3},$$

and so p has the same sign as s'_4 and q is positive. □

We now derive an explicit formula for the determinants in (23).

Theorem 3.

$$D_i = \frac{(t_{i-2} + t_{i-1})u_i + (t_i + t_{i+1})u_{i-1}}{2W^2 r_{i-1} r_i r_{i+1}}, \quad 1 \leq i \leq 4.$$

By Lemma 2, Theorem 3 implies that $D_i(\mathbf{x}) > 0$ for all $\mathbf{x} \in P$ and $i = 1, 2, 3, 4$.

Proof. Observe first that by [9, Lemma 2],

$$D_i = \frac{1}{W^3} D(w_{i-1}, w_i, w_{i+1}),$$

and we can write this latter determinant as

$$D(w_{i-1}, w_i, w_{i+1}) = \begin{vmatrix} w_{i-1} & w_i & w_{i+1} \\ \nabla w_{i-1} & \nabla w_i & \nabla w_{i+1} \end{vmatrix}.$$

From (14) and Theorem 2,

$$D(w_{i-1}, w_i, w_{i+1}) = \frac{1}{4r_{i-1}r_i r_{i+1}} E_i,$$

where

$$E_i := \begin{vmatrix} t_{i-2} + t_{i-1} & t_{i-1} + t_i & t_i + t_{i+1} \\ \mathbf{a}_{i-1} & \mathbf{a}_i & \mathbf{a}_{i+1} \end{vmatrix}, \quad (24)$$

and

$$\mathbf{a}_j := u_{j-1} \mathbf{e}_{j-1}^\perp - u_j \mathbf{e}_{j+1}^\perp,$$

and it remains to show that

$$E_i = 2W((t_{i-2} + t_{i-1})u_i + (t_i + t_{i+1})u_{i-1}), \quad 1 \leq i \leq 4.$$

We may assume without loss of generality that $i = 1$ and the task is to show that

$$E_1 = 2W((t_3 + t_4)u_1 + (t_1 + t_2)u_4).$$

From (24),

$$E_1 = (t_3 + t_4)D_{12} + (t_4 + t_1)D_{24} + (t_1 + t_2)D_{41},$$

where $D_{i,j} := \mathbf{a}_i \times \mathbf{a}_j$, and since $\mathbf{a}^\perp \times \mathbf{b}^\perp = \mathbf{a} \times \mathbf{b}$ for any vectors $\mathbf{a}, \mathbf{b}$ in $\mathbb{R}^2$,

$$\begin{aligned} D_{12} &= (u_4 \mathbf{e}_4 - u_1 \mathbf{e}_2) \times (u_1 \mathbf{e}_1 - u_2 \mathbf{e}_3) = u_1 u_4 s_4 + u_2 u_4 s_3 + u_1^2 s_1 + u_1 u_2 s_2, \\ D_{24} &= (u_1 \mathbf{e}_1 - u_2 \mathbf{e}_3) \times (u_3 \mathbf{e}_3 - u_4 \mathbf{e}_1) = (u_1 u_3 - u_2 u_4)(\mathbf{e}_1 \times \mathbf{e}_3), \\ D_{41} &= (u_3 \mathbf{e}_3 - u_4 \mathbf{e}_1) \times (u_4 \mathbf{e}_4 - u_1 \mathbf{e}_2) = u_3 u_4 s_3 + u_1 u_3 s_2 + u_4^2 s_4 + u_1 u_4 s_1. \end{aligned}$$

Next observe that

$$\frac{1}{2} \sum_{j=1}^4 s_j u_j = \sum_{j=1}^4 \left(\frac{1}{r_j} + \frac{1}{r_{j+1}} \right) t_j = \sum_{j=1}^4 \frac{1}{r_j} (t_j + t_{j-1}) = \sum_{j=1}^4 w_j = W. \quad (25)$$

Therefore,

$$\begin{aligned} D_{12} &= 2W u_1 - (u_1 u_3 - u_2 u_4) s_3, \\ D_{41} &= 2W u_4 + (u_1 u_3 - u_2 u_4) s_2, \end{aligned}$$

and so

$$E_1 = 2W((t_3 + t_4)u_1 + (t_1 + t_2)u_4) + (u_1u_3 - u_2u_4)K,$$

where

$$K := -(t_3 + t_4)s_3 + (t_4 + t_1)(\mathbf{e}_1 \times \mathbf{e}_3) + (t_1 + t_2)s_2,$$

and it remains to show that $K = 0$.

By (17),

$$\mathbf{e}_1 \times \mathbf{e}_3 = \sin(\alpha_1 + \alpha_2) = s_1c_2 + c_1s_2 = 2\frac{(t_1 + t_2)(1 - t_1t_2)}{(1 + t_1^2)(1 + t_2^2)},$$

and the only occurrences of t_4 in K are in the sums $t_3 + t_4$ and $t_4 + t_1$. We can eliminate t_4 from these sums using Lemma 3:

$$\begin{aligned} t_3 + t_4 &= \frac{t_3q + p}{q} = \frac{(1 + t_3^2)(t_1 + t_2)}{q}, \\ t_4 + t_1 &= \frac{p + t_1q}{q} = \frac{(1 + t_1^2)(t_2 + t_3)}{q}. \end{aligned}$$

Thus it follows that

$$K = \frac{2(t_1 + t_2)}{(1 + t_2^2)q} \left(-t_3(1 + t_2^2) + (t_2 + t_3)(1 - t_1t_2) + t_2q \right),$$

which by the definition of q turns out to be zero. $\square$

6 Determinants on the boundary

By [13, Remark 4.5], $\phi_1, \phi_2, \phi_3, \phi_4$ are C^∞ at all points of ∂P except the vertices. The following lemma shows that the determinant condition (23) holds at such points.

Lemma 4. *For any $\mathbf{y}$ in the relative interior of the edge $[\mathbf{p}_1, \mathbf{p}_2]$,*

$$D_i(\mathbf{y}) = \begin{cases} \frac{w_4(\mathbf{y})}{2\|\mathbf{p}_2 - \mathbf{p}_1\|} & \text{if } i = 1, \\ \frac{w_3(\mathbf{y})}{2\|\mathbf{p}_2 - \mathbf{p}_1\|} & \text{if } i = 2, \\ 0 & \text{if } i = 3, 4. \end{cases}$$

Proof. We use the fact that

$$D_i(\mathbf{y}) = \lim_{\mathbf{x} \rightarrow \mathbf{y}, \mathbf{x} \in P} D_i(\mathbf{x}).$$

As $\mathbf{x} \rightarrow \mathbf{y}$ for $\mathbf{x} \in P$, $t_1 \rightarrow \infty$ while t_2, t_3, t_4 and r_1, r_2, r_3, r_4 have finite limits.

Consider the case $i = 1$, so that for $\mathbf{x} \in P$,

$$D_1 = \frac{(t_3 + t_4)u_1 + (t_1 + t_2)u_4}{2W^2 r_4 r_1 r_2}.$$

Its numerator is

$$(t_3 + t_4) \left(\frac{1}{r_1} + \frac{1}{r_2} \right) t_1^2 + O(t_1), \quad \text{as } t_1 \rightarrow \infty.$$

Recalling (25),

$$W = \left(\frac{1}{r_1} + \frac{1}{r_2} \right) t_1 + O(1), \quad \text{as } t_1 \rightarrow \infty,$$

and so the denominator of D_1 is

$$2r_4 r_1 r_2 \left(\frac{1}{r_1} + \frac{1}{r_2} \right)^2 t_1^2 + O(t_1), \quad \text{as } t_1 \rightarrow \infty.$$

Therefore,

$$\lim_{\mathbf{x} \rightarrow \mathbf{y}, \mathbf{x} \in P} D_1 = \frac{t_3 + t_4}{2r_4(r_1 + r_2)} = \frac{w_4}{2\|\mathbf{p}_2 - \mathbf{p}_1\|}.$$

When $i = 2$,

$$D_2 = \frac{(t_4 + t_1)u_2 + (t_2 + t_3)u_1}{2W^2 r_1 r_2 r_3},$$

and its numerator is

$$(t_2 + t_3) \left(\frac{1}{r_1} + \frac{1}{r_2} \right) t_1^2 + O(t_1),$$

and its denominator is

$$2r_1 r_2 r_3 \left(\frac{1}{r_1} + \frac{1}{r_2} \right)^2 t_1^2 + O(t_1),$$

and so

$$\lim_{\mathbf{x} \rightarrow \mathbf{y}, \mathbf{x} \in P} D_2 = \frac{t_2 + t_3}{2r_3(r_1 + r_2)} = \frac{w_3}{2\|\mathbf{p}_2 - \mathbf{p}_1\|}.$$

For the cases $i = 3, 4$, the numerator of D_i is $O(t_1)$ and its denominator grows like t_1^2 , and so

$$\lim_{\mathbf{x} \rightarrow \mathbf{y}, \mathbf{x} \in P} D_i = 0.$$

□

7 Global injectivity

We will now complete the proof of Theorem 1. By Theorem 3 and Lemma 4, the Jacobian of $\mathbf{f}$ is positive in P and at all points of ∂P except the four vertices. So by the Inverse Function Theorem (IFT) [21, Theorem 9.24], $\mathbf{f}$ is locally injective at all such points. It remains to show that this implies the (global) injectivity of $\mathbf{f}$ as a mapping from $\overline{P}$ to $\overline{Q}$. The method of proof uses the path-connectedness of P and the compactness of $\overline{P}$, similar to that of [16]. Let us say that $\mathbf{x} \in \overline{P}$ is *multiple* if there exists $\mathbf{y} \in \overline{P}$, $\mathbf{y} \neq \mathbf{x}$, such that $\mathbf{f}(\mathbf{y}) = \mathbf{f}(\mathbf{x})$. Since $\mathbf{f}$ maps ∂P to ∂Q injectively and since $\mathbf{f}$ maps P to Q , ∂P contains no multiple points, and the task is to show that P contains no multiple points either.

Proof of Theorem 1. Suppose, for the sake of contradiction, that there exists some $\mathbf{x} \in P$ that is multiple. Then choose any point $\mathbf{x}_B \in \partial P$ that is not a vertex of P and such that the interior of the line segment $[\mathbf{x}, \mathbf{x}_B]$ lies entirely in P , and let $\gamma(t)$ be the parameterization

$$\gamma(t) := (1 - t)\mathbf{x} + t\mathbf{x}_B, \quad 0 \leq t \leq 1.$$

Let

$$T := \{t \in [0, 1] : \gamma(t) \text{ is multiple}\},$$

and let $t_* := \sup T$.

We will next show that $\gamma(t_*)$ is multiple. If $t_* = 0$ then $\gamma(t_*) = \mathbf{x}$, which is multiple. So suppose that $t_* > 0$. Then there exists a sequence (t_n) in T converging to t_* . Since γ is continuous, $\gamma(t_n) \rightarrow \gamma(t_*)$. Since $\gamma(t_n)$ is multiple, there is some $\mathbf{x}_n \in \overline{P}$, $\mathbf{x}_n \neq \gamma(t_n)$, such that $\mathbf{f}(\mathbf{x}_n) = \mathbf{f}(\gamma(t_n))$. Since $\overline{P}$ is compact, the sequence $(\mathbf{x}_n)$ has a convergent subsequence in $\overline{P}$, and so we may assume that both $\gamma(t_n) \rightarrow \gamma(t_*)$ and $\mathbf{x}_n \rightarrow \mathbf{x}_*$ for some $\mathbf{x}_* \in \overline{P}$.

Since $J(\mathbf{f})(\gamma(t_*)) > 0$, by IFT there are open sets U and V in $\mathbb{R}^2$ such that $\gamma(t_*) \in U$, $\mathbf{f}(\gamma(t_*)) \in V$, $\mathbf{f}$ is injective on U and $\mathbf{f}(U) = V$. Then for large enough n , $\gamma(t_n) \in U$ and so $\mathbf{x}_n \notin U$ and so $\mathbf{x}_* \notin U$, and we deduce that $\mathbf{x}_* \neq \gamma(t_*)$. On the other hand, since $\mathbf{f}$ is continuous in $\overline{P}$, $\mathbf{f}(\gamma(t_n)) \rightarrow \mathbf{f}(\gamma(t_*))$ and $\mathbf{f}(\mathbf{x}_n) \rightarrow \mathbf{f}(\mathbf{x}_*)$, and therefore, $\mathbf{f}(\mathbf{x}_*) = \mathbf{f}(\gamma(t_*))$. Thus we have shown that $\gamma(t_*)$ is multiple as claimed.

We now consider the two cases: (i) $t_* = 1$ and (ii) $t_* < 1$.

Case (i) If $t_* = 1$ then $\gamma(t_*) = \mathbf{x}_B$. This implies that $\mathbf{x}_B$ is multiple which is a contradiction.

Case (ii) If $t_* < 1$ then $\gamma(t_*) \in P$ and so $\mathbf{x}_* \in P$ as well. Then $J(\mathbf{f})(\mathbf{x}_*) > 0$ and so again by IFT, there are open sets U' and V' in $\mathbb{R}^2$ such that $\mathbf{x}_* \in U'$, $\mathbf{f}(\mathbf{x}_*) \in V'$, $\mathbf{f}$ is injective on U' and $\mathbf{f}(U') = V'$. Let $d := \|\mathbf{x}_* - \gamma(t_*)\|$. We can find an open disk $D \subseteq U$ centred at $\gamma(t_*)$ with radius at most $d/2$ and an open disk $D' \subseteq U'$ centred at $\mathbf{x}_*$ with radius at most $d/2$. Further by IFT, the inverse $\mathbf{g}$ of $\mathbf{f}$ from V to U is continuous and the inverse $\mathbf{g}'$ of $\mathbf{f}$ from V' to U' is continuous. Therefore, $\mathbf{f}(D) = \mathbf{g}^{-1}(D)$ and $\mathbf{f}(D') = (\mathbf{g}')^{-1}(D')$ are open sets. Let $V'' := \mathbf{f}(D) \cap \mathbf{f}(D')$ and let $U_2 := \mathbf{g}(V'')$ and $U'_2 := \mathbf{g}'(V'')$. For small enough $\epsilon > 0$, $\gamma(t_* + \epsilon) \in U_2$ and then $\mathbf{f}(\gamma(t_* + \epsilon)) \in V''$ and $\mathbf{x}_\epsilon := \mathbf{g}'(\mathbf{f}(\gamma(t_* + \epsilon))) \in U'_2$. Then $\mathbf{x}_\epsilon \neq \gamma(t_* + \epsilon)$ and $\mathbf{f}(\mathbf{x}_\epsilon) = \mathbf{f}(\gamma(t_* + \epsilon))$ and so $\gamma(t_* + \epsilon)$ is multiple which contradicts the definition of t_* . $\square$

Declaration

The authors declare that they have no conflicting interests.

References

- [1] D. Anisimov, C. Deng, and K. Hormann, *Subdividing barycentric coordinates*, Comput. Aided Geom. Design **43** (2016), 172–185.
- [2] M. Bronstein, *Simplification of real elementary functions*, Proceedings of the ACM-SIGSAM international symposium on symbolic and algebraic computation, 1989, pp. 207–211.
- [3] G. Choquet, *Sur un type de transformation analytique généralisant la représentation conforme et défini au moyen de fonctions harmoniques*, Bull. Sci. Math. **69** (1945), 156–165.
- [4] C. Dyken and M. S. Floater, *Transfinite mean value interpolation*, Comput. Aided Geom. Design **26** (2009), 117–134.
- [5] M. S. Floater, *Mean value coordinates*, Comput. Aided Geom. Design **20** (2003), 19–27.
- [6] ———, *Wachspress and mean value coordinates*, Approximation Theory XIV: San Antonio 2013, Springer, 2014, pp. 81–102.

- [7] M. S. Floater, A. Gillette, and N. Sukumar, *Gradient bounds for Wachspress coordinates on polytopes*, SIAM J. Numer. Anal. **52** (2014), 515–532.
- [8] M. S. Floater, G. Kos, and M. Reimers, *Mean value coordinates in 3D*, Comput. Aided Geom. Design **22** (2005), 623–631.
- [9] M. S. Floater and J. Kosinka, *On the injectivity of Wachspress and mean value mappings between convex polygons*, Adv. Comput. Math. **32** (2010), 163–174.
- [10] M. S. Floater and M.-J. Lai, *Polygonal spline spaces and the numerical solution of the poisson equation*, SIAM J. Numer. Anal. **54** (2016), 797–824.
- [11] C. Fuda and K. Hormann, *A new stable method to compute mean value coordinates*, Comput. Aided Geom. Design **111** (2024), 102310, pp. 1–16.
- [12] J. Gravesen, A. Evgrafov, D.-M. Nguyen, and P. Nørtoft, *Planar parameterization in Isogeometric Analysis*, Mathematical Methods for Curves and Surfaces: 8th international conference, Springer, 2014, pp. 189–212.
- [13] K. Hormann and M. S. Floater, *Mean value coordinates for arbitrary planar polygons*, ACM Trans. on Graph. **25** (2006), 1424–1441.
- [14] K. Hormann and M. Tarini, *A quadrilateral rendering primitive*, Proceedings of the ACM SIGGRAPH/EUROGRAPHICS conference on graphics hardware, 2004, pp. 7–14.
- [15] T. Ju, S. Schaefer, and J. Warren, *Mean value coordinates for closed triangular meshes*, ACM Trans. on Graph. **24** (2005), 561–566.
- [16] H. Kestelman, *Mappings with non-vanishing Jacobian*, Amer. Math. Month. **78** (1971), 662–663.
- [17] H. Kneser, *Lösung der Aufgabe 41*, Jahresber. Deutsch. Math.-Verien. **35** (1926), 123–124.
- [18] Y. Lipman, J. Kopf, D. Cohen-Or, and D. Levin, *GPU-assisted positive mean value coordinates for mesh deformation*, Symposium on Geometry Processing 2007, Eurographics Association, 2007, pp. 117–123.

- [19] T. Radó, *Aufgabe 41*, Jahresber. Deutsch. Math.-Verien. **35** (1926), 49.
- [20] A. Rand, A. Gillette, and C. Bajaj, *Interpolation error estimates for mean value coordinates over convex polygons*, Adv. Comput. Math. **39** (2013), 327–347.
- [21] W. Rudin, *Principles of mathematical analysis, 3rd edition*, McGraw Hill, New York, 1976.
- [22] T. Schneider, K. Hormann, and M. S. Floater, *Bijective composite mean value mappings*, Symposium on Geometry Processing 2013, Eurographics Association, 2013, pp. 137–146.
- [23] J. Warren, *Barycentric coordinates for convex polytopes*, Adv. Comput. Math. **6** (1996), 97–108.
- [24] O. Weber, M. Ben-Chen, C. Gotsman, and K. Hormann, *A complex view of barycentric mappings*, Comput. Graph. Forum **30** (2011), 1533–1542.