

Research Article

A New Proof of the Pythagorean Theorem and Its Application to Element Decompositions in Topological Algebras

Fred Greensite

*Department of Radiological Sciences, University of California Irvine Medical Center,
Orange, CA 92868, USA*

Correspondence should be addressed to Fred Greensite, fredg@uci.edu

Received 28 February 2012; Accepted 3 June 2012

Academic Editor: Attila Gilányi

Copyright © 2012 Fred Greensite. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a new proof of the Pythagorean theorem which suggests a particular decomposition of the elements of a topological algebra in terms of an “inverse norm” (addressing unital algebraic structure rather than simply vector space structure). One consequence is the unification of Euclidean norm, Minkowski norm, geometric mean, and determinant, as expressions of this entity in the context of different algebras.

1. Introduction

Apart from being unital topological $*$ -algebras, matrix algebras, special Jordan algebras and Cayley-Dickson algebras would seem to have little else in common. For example, the matrix algebra $\mathbb{R}^{n \times n}$ is associative but noncommutative, the special Jordan algebra derived from it is nonassociative but commutative (and so are the spin factor Jordan algebras), and Cayley-Dickson algebras are both nonassociative and noncommutative (apart from the three lowest dimensional instances). However, these latter sets of algebras share an interesting feature. Each is associated with a function $f(s)$ that vanishes on the nonunits and provides a decomposition of every unit as

$$s = f(s)\nabla^* f(s^{-1}), \quad (1.1)$$

with

$$f\left(\nabla^* f\left(s^{-1}\right)\right) = 1, \quad (1.2)$$

where $\nabla^* f(s^{-1})$ indicates that the gradient ∇f is evaluated at s^{-1} following which the involution $*$ is applied. For example,

- (i) on an n -dimensional Cayley-Dickson algebra, $f(s)$ is the quadratic mean of the components multiplied by \sqrt{n} , that is, $f(s)$ is the Euclidean norm,
- (ii) on the algebra \mathbb{R}^n with component-wise addition and multiplication, $f(s)$ is the geometric mean of the absolute values of the components multiplied by \sqrt{n} ,
- (iii) on the matrix algebra $\mathbb{R}^{n \times n}$, $f(s)$ is the (principal) n th root of the determinant multiplied by \sqrt{n} ,
- (iv) on the spin factor Jordan algebras, $f(s)$ is the Minkowski norm.

Looked at another way, the Euclidean norm, the geometric mean, the n th root of the determinant, and the Minkowski norm are all expressions of the same thing in the context of different algebras. With respect to topological $*$ -algebras, this “thing” supercedes the Euclidean norm and determinant since neither is meaningful in all settings for which the solution to (1.1) and (1.2) is meaningful.

There is another aspect relevant to the Cayley-Dickson algebras. In addition to (1.1) and (1.2), there is a function f on the elements of the algebra such that for any unit s in the algebra,

$$s = f(s)\nabla f(s), \quad (1.3)$$

with

$$f(\nabla f(s)) = 1. \quad (1.4)$$

This equation set makes no reference to multiplicative structure; that is, it is a general property of the underlying vector space. Indeed, $f(s)$ is again the Euclidean norm. In fact, (1.3) and (1.4) can be derived from the (Hilbert formulation) axioms of plane Euclidean geometry without use of the Pythagorean theorem—and as such can be used as the centerpiece of a new proof of that theorem.

So, we first prove the Pythagorean Theorem by deriving (1.3) and (1.4), we then use the latter equations to develop (1.1) and (1.2), following which we demonstrate the assertions of the first paragraph of this Introduction. Ultimately, existence of the decomposition (1.1), (1.2) is forwarded as a kind of surrogate for the Pythagorean theorem in the context of topological algebras.

It will also be seen that there is a hierarchy related to the basic equations, evidenced by progressively more structure accompanying the solution function on particular algebras. The equations (1.1) and (1.2) and have a clear analogy with the form of (1.3) and (1.4). Both cases present a decomposition of the units of an algebra as the product of a particular function's value at that point multiplied by a unity-scaled orientation point dependent on the function's gradient. The equations are nonlinear, in general. However, along with the prescription that

$f(s)$ vanishes on the nonunits, the above particular matrix, Cayley-Dickson, and spin factor Jordan algebras, also happen to satisfy the additional property that there is a real constant α such that for any unit s in the algebra,

$$f(s)f(s^{-1}) = \alpha. \quad (1.5)$$

Replacing s with s^{-1} in (1.1), the above implies a decomposition of the inverse of a point as

$$s^{-1} = \alpha \frac{\nabla^* f(s)}{f(s)}. \quad (1.6)$$

Furthermore, multiplying both sides of (1.6) by $f(s)s$ indicates that the above is a linear equation for f . One can be even more restrictive and consider the set of algebras on which a function exists satisfying all of the above where in addition (1.5) is strengthened to $f(s_1)f(s_2) = \alpha f(s_1s_2)$ for any two units s_1, s_2 . In this case, the units form a group and $f(s)/\alpha$ is a homomorphism on this group. The octonions occupy a special place as an algebra satisfying this prescription that is not a matrix subalgebra.

2. Euclidean Decomposition

Theorem 2.1 (Pythagoras). *In a space satisfying the axioms of plane Euclidean geometry, the square of the hypotenuse of a right triangle is equal to the sum of the squares of its two other sides.*

The theorem hypothesis is assumed to indicate the Hilbert formulation of plane Euclidean geometry [1]. One will refer to the point set in question as the “Euclidean plane.” Points will be denoted by lower case Roman letters and real numbers by lower case Greek letters.

Proof. We begin by providing an outline of the proof.

A vector space structure is defined on the Euclidean plane E after identifying one of the vertices of the hypotenuse of the given right triangle as an origin o . We define the Euclidean norm implicitly as the function $f : E \rightarrow \mathbb{R}$ giving the length of a line segment from the origin to any given point in the plane. Since the Hilbert formulation includes continuity axioms, we can employ the usual notions relating to limits and thereby define directional derivatives. The crux of the proof, Lemma 2.3, is the demonstration that the parallel axiom implies the existence and continuity of the directional derivatives at points other than the origin, and the largest directional derivative at a point s associated with a unit length direction line segment has unit value and is such that its direction line segment is collinear with line segment \overline{os} . The novelty lies in the necessity that this be accomplished in the absence of an explicit formula for the Euclidean norm. A Cartesian axis system is now generated from the two sides of the given triangle forming the right angle, following which an isomorphism from E to vector space \mathbb{R}^2 is easily demonstrated. Using this isomorphism, Lemma 2.3 is seen to imply the existence of the gradient $\nabla f(t) \in E$ for t not at the origin, with its characteristic property regarding generation of a directional derivative from a particular specified direction line segment, and such that the origin, t , and $\nabla f(t)$ are collinear. It is then a simple matter to show $t = f(t)\nabla f(t)$ and $f(\nabla f(t)) = 1$. For $t = (\tau_1, \tau_2)$, a solution to the latter partial differential equation is supplied by $f(\tau_1, \tau_2) = \sqrt{\tau_1^2 + \tau_2^2}$. This solution is unique because $t = f(t)\nabla f(t)$ implies that t and $\nabla f(t)$

are collinear with the origin, so that the equation can be written as an ordinary differential equation—which is easily shown to have a unique solution. This explicit representation of the Euclidean norm implies the Pythagorean theorem, thus concluding the proof.

We now fill in details the above argument.

Given a right triangle $\Delta\{o, a, s\}$ with line segment \overline{os} as the hypotenuse, we define a function $f(t)$ that gives the length of the line segment \overline{ot} for any point t in the plane, with the convention that $f(o) = 0$. The continuity axioms that are part of the Hilbert formulation support the equivalent of the least upper bound axiom for the real number system, and we have the usual continuity properties related to \mathbb{R} , whose elements identify lengths of line segments in particular. Thus, the usual notion of limit can be defined and is assumed.

Definition 2.2. A direction line segment with respect to a particular point is any line segment one of whose endpoints is the given point. If the following limit exists, the *directional derivative* of f at t with respect to direction line segment \overline{tw} is

$$D_{\overline{tw}}f(t) \equiv \lim_{\epsilon \rightarrow 0} \frac{f(t \oplus \epsilon \overline{tw}) - f(t)}{\epsilon}, \quad (2.1)$$

where $t \oplus \epsilon \overline{tw}$ formally denotes a point on the line containing \overline{tw} such that the line segment defined by this point and t has length given by the product of $|\epsilon|$ and the length of \overline{tw} , and t lies between this point and w if and only if $\epsilon < 0$.

Lemma 2.3. For any point w different from s , the directional derivative of f at s specified by \overline{sw} exists and is continuous at s . Furthermore, the largest directional derivative at s associated with a direction line segment \overline{sw} of a particular fixed length is such that o, s, w are collinear with s between o and w , and if the direction line segment has unit length, then the largest directional derivative has unit value.

Proof of Lemma 2.3. Let the length of \overline{sw} be λ . If w lies on the line containing \overline{os} with s between o and w , it is clear that the directional derivative exists and has a value given by λ , since the expression inside the limit in (2.1) has this value for all $\epsilon \neq 0$. In the same way, we can demonstrate that the directional derivative exists when w is on the line containing \overline{os} but s is not between o and w —in which case the directional derivative has value $-\lambda$.

Thus, suppose \overline{sw} is not on the line containing \overline{os} . Figure 1 can be used to keep the following constructions in context. Consider the orthogonal projection of the point $s \oplus \epsilon \overline{sw}$ to the line containing \overline{os} , that is, the point p on the line containing \overline{os} such that $\angle\{o, p, s \oplus \epsilon \overline{sw}\}$ is a right angle (the expression $\angle\{o, p, s \oplus \epsilon \overline{sw}\}$ denotes the angle within the triangle $\Delta\{o, p, s \oplus \epsilon \overline{sw}\}$ that is formed by the two line segments \overline{op} and $\overline{p(s \oplus \epsilon \overline{sw})}$). Consider the circle centered at o of radius $f(s \oplus \epsilon \overline{sw})$. We claim that p is in the region enclosed by the circle, that is, $f(p) < f(s \oplus \epsilon \overline{sw})$. Indeed, suppose this is false. Consider the circle with center at o of radius $f(p)$, and let the tangent line to this circle at the point p be denoted T_p . Being a tangent, all of the points y of T_p other than p will be such that

$$f(y) > f(p) \geq f(s \oplus \epsilon \overline{sw}). \quad (2.2)$$

Note that T_p intersects the line containing \overline{os} at a right angle (a tangent to a circle at a particular point is perpendicular to the circle radius at that point). But there is only one line

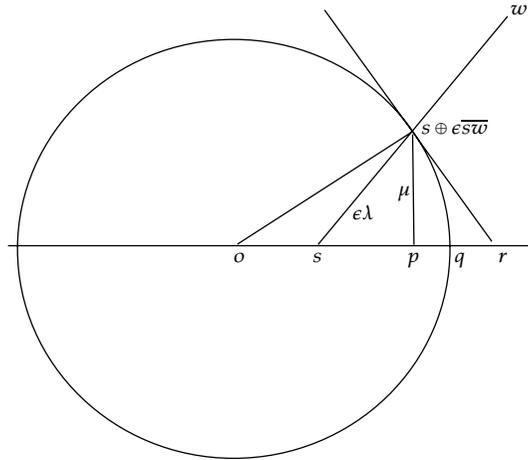


Figure 1: Diagram relating to proof that the function giving the length of a line segment with one endpoint at o has continuous directional derivatives at points other than o , Lemma 2.3. Three facts used in the proof are as follows: (1) $\Delta\{p, o, s \oplus \epsilon \overline{s w}\}$ is similar to $\Delta\{p, s \oplus \epsilon \overline{s w}, r\}$ because $\angle\{o, s \oplus \epsilon \overline{s w}, r\}$ is a right angle, (2) for all $\epsilon \neq 0$, the triangles $\Delta\{p, s, s \oplus \epsilon \overline{s w}\}$ are similar, (3) q is between p and r .

through p that meets the line containing \overline{os} at a right angle, and that is the line containing $\overline{p(s \oplus \epsilon \overline{s w})}$, as previously defined. So T_p must contain $\overline{p(s \oplus \epsilon \overline{s w})}$, that is, $s \oplus \epsilon \overline{s w}$ is on T_p but is not the point p . The first inequality of (2.2) would then imply that $y = s \oplus \epsilon \overline{s w} \in T_p$ is such that $f(s \oplus \epsilon \overline{s w}) > f(p)$. This contradicts the second inequality of (2.2). Hence, we have established our claim that $f(p) < f(s \oplus \epsilon \overline{s w})$.

Let $P_{\overline{os}}$ be the line perpendicular to \overline{os} at o . Since s is not the point o , the line containing $\overline{s w}$ intersects $P_{\overline{os}}$ in at most one point. In what follows, we assume $|\epsilon|$ is small enough so that no point of $\overline{s(s \oplus \epsilon \overline{s w})}$ is on $P_{\overline{os}}$. Again consider the circle centered at o with diameter $\overline{f(s \oplus \epsilon \overline{s w})}$. Its tangent at $s \oplus \epsilon \overline{s w}$ must intersect the line containing \overline{os} at the point r (since $s \oplus \epsilon \overline{s w}$ is not on $P_{\overline{os}}$, the tangent cannot be parallel to \overline{os}). Being on a tangent but not the point of tangency itself, r is necessarily external to the region enclosed by the circle ($f(r)$ is greater than the circle radius). This circle intersects the line containing \overline{os} at two points defining a diameter of the circle. We denote as q the one of these two points such that q is between p and r .

We are first required to show the existence of the limit in (2.1) (for $t = s$), which we can write as $\lim_{\epsilon \rightarrow 0} (f(q) - f(s))/\epsilon$, since q and $s \oplus \epsilon \overline{s w}$ both lie on the aforementioned circle. To do this, we will initially assume that $\epsilon > 0$ and show that $\lim_{\epsilon \rightarrow 0^+} (f(q) - f(s))/\epsilon$ exists, after which it will be clear that an entirely analogous argument establishes the same value for $\lim_{\epsilon \rightarrow 0^-} (f(q) - f(s))/\epsilon$.

We have

$$\frac{f(q) - f(s)}{\epsilon \lambda} = \frac{f(q) - f(p)}{\epsilon \lambda} + \frac{f(p) - f(s)}{\epsilon \lambda}, \tag{2.3}$$

$$\frac{f(q) - f(p)}{\epsilon \lambda} \leq \frac{f(r) - f(p)}{\epsilon \lambda}, \tag{2.4}$$

since q is between p and r . Note that $\epsilon\lambda$ is the length of $\overline{s(s \oplus \epsilon\overline{sw})}$, according to Definition 2.2. The triangle $\Delta\{p, o, s \oplus \epsilon\overline{sw}\}$ is similar to $\Delta\{p, s \oplus \epsilon\overline{sw}, r\}$ (ultimately, because the tangent line to the circle at $s \oplus \epsilon\overline{sw}$ implies that $\angle\{o, s \oplus \epsilon\overline{sw}, r\}$ is a right angle). Let μ be the length of $\overline{p(s \oplus \epsilon\overline{sw})}$. It follows that

$$\frac{f(r) - f(p)}{\mu} = \frac{\mu}{f(p)}. \quad (2.5)$$

If s and p are ever the same point for some value of ϵ , then $\overline{s(s \oplus \epsilon\overline{sw})}$ is perpendicular to \overline{os} , and will remain so for any other ϵ , so that s and p will always be the same point. In that case, $f(s) = f(p)$, and $\mu = \epsilon\lambda$. It then follows that the right-hand side of (2.5) tends to zero with ϵ (since this right-hand side is $(\epsilon\lambda)/f(s)$), so the right-hand side of (2.4) also tends to zero with ϵ , which means that the first term on the right-hand side of (2.3) tends to zero with ϵ . But since $f(s) = f(p)$, the second term on the right-hand side of (2.3) is zero. It then follows that $\lim_{\epsilon \rightarrow 0} (f(q) - f(s))/\epsilon = 0$ and, in particular, the required limit exists.

Thus, suppose s and p are different, and consider $\Delta\{p, s, s \oplus \epsilon\overline{sw}\}$. This defines a set of similar triangles for all values of $\epsilon \neq 0$, because the angle $\angle\{p, s, s \oplus \epsilon\overline{sw}\}$ does not change as ϵ varies and $\angle\{s, p, s \oplus \epsilon\overline{sw}\}$ remains a right angle. Consequently, $\mu/(\epsilon\lambda)$ is a nonzero constant for all $\epsilon > 0$ since this is the ratio of two particular sides of each triangle in this set of similar triangles. This means that $\lim_{\epsilon \rightarrow 0^+} \mu = 0$, since $\mu/(\epsilon\lambda)$ could not otherwise remain a constant because $\epsilon\lambda$ tends to zero with ϵ . On the other hand, because $|f(p) - f(s)|/(\epsilon\lambda)$ is also constant as ϵ varies (being a ratio of a different combination of sides of these same triangles), it must also follow that $\lim_{\epsilon \rightarrow 0^+} f(p) = f(s) \neq 0$ (which also means that o is not between p and s for small ϵ). So, it must be that the right-hand side of (2.5) tends to zero as $\epsilon > 0$ tends to zero. Consequently, the right-hand side of (2.4) also tends to zero as $\epsilon > 0$ tends to zero (because, as we have noted, $\mu/(\epsilon\lambda)$ is constant as $\epsilon > 0$ varies). Hence, the first term on the right-hand side of (2.3) also tends to zero as $\epsilon > 0$ becomes small. This means that the left-hand side of (2.3) has the same limit as $\epsilon > 0$ tends to zero as the second term on the right-hand side of (2.3) (assuming the limit exists). But, we have already noted that the term $|f(p) - f(s)|/(\epsilon\lambda)$ is a constant as ϵ varies, since it is determined by the ratio of particular sides of $\Delta\{p, s, s \oplus \epsilon\overline{sw}\}$, and (as we have noted) for any $\epsilon \neq 0$ all such triangles are similar. In fact (removing the absolute value sign in the numerator), we further claim that

$$\kappa \equiv \frac{f(p) - f(s)}{\epsilon\lambda} \quad (2.6)$$

is a constant for all small $\epsilon > 0$. To see this, recall that for small $\epsilon > 0$, o is not between s and p , and $\angle\{p, s, s \oplus \epsilon\overline{sw}\}$ is constant. Now, if p changes from being between o and s versus not being between o and s as $\epsilon > 0$ varies, the latter angle must change from being an acute angle to being an obtuse angle, or vice versa, which contradicts the fact that the angle is constant for all $\epsilon > 0$. For small $\epsilon > 0$, it follows that s is always between o and p , or p is always between o and s —which establishes our claim that (2.6) is constant. Therefore,

$$\lim_{\epsilon \rightarrow 0^+} \frac{f(q) - f(s)}{\epsilon} = \lambda\kappa. \quad (2.7)$$

An analogous argument establishes $\lim_{\epsilon \rightarrow 0^+} (f(q) - f(s))/\epsilon = \lambda\kappa$, because $\Delta\{p, s, (s \oplus \epsilon\overline{sw})\}$ are similar triangles for *all* nonzero values of ϵ . This establishes the limit in (2.1).

Next, we establish the continuity of each directional derivative. Consider any sequence of line segments $\{\overline{s_i w_i}\}$ such that $\{s, w, w_i, s_i\}$ form a parallelogram for all i , and the limit of the length of the line segments $\{\overline{s_i s}\}$ is zero (which implies that the limit of the length of the line segments $\{\overline{w_i w}\}$ is also zero). To establish the continuity of a directional derivative at s it is required to show that the left-hand-side of the following equation is zero:

$$\lim_{i \rightarrow \infty} |D_{\overline{s w}} f(s) - D_{\overline{s_i w_i}} f(s)| = \lim_{i \rightarrow \infty} |\kappa - \kappa_i| \lambda. \quad (2.8)$$

For each i , κ_i is the ratio of particular sides of any member of a particular set of similar triangles, as is also the case for κ as above. Furthermore, as i increases, the ratio of the lengths of the sides of the triangle relevant to κ_i converges to the same ratio as that of the corresponding triangles relevant to κ , since s_i converges to s and w_i converges to w . Therefore, $|D_{\overline{s w}} f(s) - D_{\overline{s_i w_i}} f(s)|$ necessarily tends to zero, implying continuity of the directional derivative at s .

We next show that the largest directional derivative at s associated with a direction line segment \overline{sw} of a particular fixed length is such that o, s, w are collinear with s between o and w . We have from (2.7) (and the equation in the following sentence) that the directional derivative is $\lambda\kappa$. Thus, we only need to show that $|f(p) - f(s)| \leq \epsilon\lambda$, with equality if and only if \overline{sw} is on the line containing \overline{os} , because once that is established it is easy to show that if o, s, w are collinear, then κ will be negative if s is not between o and w , and will be positive otherwise.

Now, $\lambda\epsilon$ is the length of the hypotenuse of right triangle $\Delta\{s, p, s \oplus \epsilon\overline{sw}\}$, and $|f(p) - f(s)|$ is the length of the side that is on the line containing \overline{os} . So, we only need to show that a nonhypotenuse side of a right triangle has a shorter length than the hypotenuse. Thus, consider any right triangle $\Delta\{a, b, c\}$ with $\angle\{a, b, c\}$ being the right angle (vertex a here has nothing to do with vertex a of our earlier given right triangle $\Delta o, a, s$). Consider a circle centered at a having radius \overline{ac} . This circle intersects the line containing \overline{ab} at a point b' . Suppose the length of $\overline{ab'}$ is less than the length of \overline{ab} , meaning that the length of the hypotenuse is smaller than the length of one of the other sides. Then b is external to the circle. So, consider a second circle centered at a but now with radius given by the length of \overline{ab} . This circle intersects the line containing \overline{ab} at the point b . The two circles are concentric, with the circle containing b lying wholly external to the circle containing c . Now, the second circle (containing b) has a tangent at b making a right angle with the line containing \overline{ab} (tangents are perpendicular to the radius at the point of tangency). But the line containing \overline{bc} is also perpendicular to the line containing \overline{ab} (since $\angle\{a, b, c\}$ is already given as a right angle). So the point c (which lies on our first circle of radius given by the length of \overline{ac}) must also be a point on the tangent line at the point b of our second circle. This is impossible since, given two concentric circles, a tangent to a point on the circle of greater radius cannot intersect the circle of smaller radius. This contradicts the assumption that the hypotenuse is smaller than the length of one of the sides. Furthermore, the hypotenuse cannot equal the length of one of the other sides of the right triangle, because in that case a circle centered at a with radius equal to the length of the hypotenuse would intersect the right triangle at the two points b and c , which would require that a tangent line to the circle at b (making a right angle with the line containing \overline{ab}) would have another point of intersection with the circle (i.e., at c , again since the line containing \overline{bc} is also perpendicular to the line containing \overline{ab} and there can be

only one such perpendicular). Thus, unless $s, p, s \oplus \epsilon \overline{sw}$ are collinear, $|f(p) - f(s)| < \epsilon \lambda$, and it is immediately verified that $|f(p) - f(s)| = \epsilon \lambda$ if the points are collinear. It is then evident that the value of the largest directional derivative is λ . Thus, if $\lambda = 1$, the largest directional derivative has unit value. Hence, Lemma 2.3 is proved. \square

Definition 2.2 actually suggests two operations, and these will be referred to in Lemma 2.5 as “the Definition 2.2 associated operations.” That is, we define the (more general) expression $c \oplus \gamma \overline{uv}$ to represent $c \oplus \gamma v \oplus (-\gamma)u$ subject to the following.

- (I) The multiplication of a scalar with a point, γv , is defined to be the point on the line containing \overline{ov} such that $f(\gamma v) = |\gamma|f(v)$ and such that o is between this new point and v if and only if γ is negative.
- (II) The “sum” of two points, $h_1 \oplus h_2$, is defined as follows.
 - (i) If o, h_1, h_2 are not collinear, $h_1 \oplus h_2$ is defined to be the point z such that the vertices $\{o, h_1, h_2, z\}$ form a parallelogram.
 - (ii) If either h_1 or h_2 is the point o , then $h_1 \oplus h_2$ is the point that is not o , or is o if h_1 and h_2 are both o .
 - (iii) If h_1 and h_2 are the same point, then $h_1 \oplus h_2$ is the point on the line containing $\overline{oh_1}$ such that $f(h_1 \oplus h_2) = 2f(h_1)$ and o is not between h_1 and $h_1 \oplus h_2$.
 - (iv) If o, h_1, h_2 are distinct and collinear,
 - (1) when o is not between h_1, h_2 then $h_1 \oplus h_2$ is the point on the line containing $\overline{h_1 h_2}$ such that $f(h_1 \oplus h_2) = f(h_1) + f(h_2)$ and o is not between h_1 and $h_1 \oplus h_2$,
 - (2) when o is between h_1, h_2 ,
 - (a) if $f(h_2) > f(h_1)$ then $h_1 \oplus h_2$ is the point on the line containing $\overline{h_1 h_2}$ such that $f(h_1 \oplus h_2) = f(h_2) - f(h_1)$ and o is between h_1 and $h_1 \oplus h_2$,
 - (b) if $f(h_2) < f(h_1)$ then $h_1 \oplus h_2$ is the point on the line containing $\overline{h_1 h_2}$ such that $f(h_1 \oplus h_2) = f(h_1) - f(h_2)$ and o is between h_2 and $h_1 \oplus h_2$,
 - (c) if $f(h_1) = f(h_2)$ then $h_1 \oplus h_2$ is o .

With this understanding, it is clear that $t \oplus \epsilon \overline{tw}$ as defined in Definition 2.2 is the same thing as $t \oplus \epsilon w \oplus (-\epsilon)t$. Of course, this suggests operations on a vector space.

Remark 2.4. The central role of the derivative of the norm function as featured in Definition 2.2 is not without precedent. The derivative of the norm also plays an important role in semi-inner product spaces [2, 3] (and premanifolds [4]), where the condition that the space be continuous (or, alternatively, uniformly continuous) can be shown to be equivalent to the condition that the norm is Gateaux differentiable (or, alternatively, uniformly Frechet differentiable). Naturally, once the isomorphism between E and \mathbb{R}^2 is established, our definition is seen to be analogous to the standard one.

Lemma 2.5. *There is an isomorphism between the vector space \mathbb{R}^2 with component-wise addition of elements and component-wise multiplication of elements by scalars, and the Euclidean plane with the Definition 2.2 associated operations.*

Proof of Lemma 2.5. We first identify a particular Cartesian axis system on the plane. One Cartesian system is already present, consisting of the lines containing the line segments forming the right angle of the right triangle given at the outset of this proof (i.e., \overline{oa} and \overline{as}).

However, since the length function $f(w)$ is referenced to o , we use the latter axis system to set up a different Cartesian system at o . Using the parallel axiom, consider the line through o that is parallel to the line containing \overline{as} , and furthermore (again using the parallel axiom) consider a point b on this new parallel line such that the vertices $\{o, a, s, b\}$ form a rectangle. The lines containing \overline{oa} and \overline{ob} are our Cartesian system (the “ \overline{oa} -axis” and the “ \overline{ob} -axis”).

We identify o with the ordered pair $(0,0)$. Any point t in the plane different from o is associated with a unique ordered pair (τ_a, τ_b) implied by the line segment \overline{ot} . That is, we take the orthogonal projection of t to the line containing \overline{oa} and take $|\tau_a|$ to be the length of the line segment formed by o and this projection of t . τ_a is negative or positive depending on whether or not o is between a and the projection of t to the line containing \overline{oa} . τ_b is defined analogously with respect to b and the \overline{ob} -axis. Conversely, every ordered pair of real numbers is associated with a point in the plane. That is, for (χ_a, χ_b) we find the point x_a on the \overline{oa} -axis with $f(x_a) = |\chi_a|$ such that o is between x_a and a if the sign of χ_a is negative, and o is not between x_a and a if the sign of χ_a is positive—and similarly for a point x_b on the \overline{ob} -axis relating to χ_b . Then the point x associated with (χ_a, χ_b) is the point such that $\{o, x_a, x, x_b\}$ is a rectangle. Its existence and uniqueness is guaranteed by the parallel axiom. It is further obvious that the first construction associates t to (τ_a, τ_b) if and only if the second construction associates (τ_a, τ_b) to t .

The above is therefore a one-to-one mapping between the points of the Euclidean plane and the points of \mathbb{R}^2 (ordered pairs of real numbers), explicitly employing the parallel and betweenness axioms. \mathbb{R}^2 becomes a vector space once we specify that ordered pairs (i.e., vectors) representing points in the plane can be added together component-wise and multiplied by scalars component-wise. The isomorphism between vector space \mathbb{R}^2 and E with the Definition 2.2 associated operations is then easily verified. Thus, Lemma 2.5 is established. \square

It follows that the isomorphism in the above lemma leads to an expression for the directional derivative in the vector space \mathbb{R}^2 that gives the same result as it did in the original Euclidean plane. In particular, basic arguments from multivariable calculus establish the existence of a total derivative, the gradient $\nabla f(s)$, as the ordered pair of directional derivatives of f at s with direction line segments defined by unit length line segments parallel to the \overline{oa} -axis and \overline{ob} -axis such that the directional derivatives are the inner product of $\nabla f(s)$ with the ordered pair in \mathbb{R}^2 corresponding to a particular direction line segment. For example, this is seen from

$$\begin{aligned} f(a_1, b_1) - f(a_0, b_0) &= [f(a_1, b_1) - f(a_0, b_1)] + [f(a_0, b_1) - f(a_0, b_0)] \\ &= \frac{\partial f(\xi_a)}{\partial a}(a_1 - a_0) + \frac{\partial f(\xi_b)}{\partial b}(b_1 - b_0), \end{aligned} \quad (2.9)$$

with ξ_a, ξ_b given by the mean value theorem. Applying the definition of directional derivative to both sides above one sees that the directional derivative is given by the usual inner product of $\nabla f(s) = (\partial f(s)/\partial a, \partial f(s)/\partial b)$ with the direction line segment. This is accomplished without use of the Euclidean norm or prior use of the inner product operation. Being an ordered pair, $\nabla f(s)$ is a point in the plane, and we can refer to $f(\nabla f(s))$. Furthermore, we have already established that the largest directional derivative of f at s associated with a line segment \overline{sw} of length λ is such that o, s, w are collinear with s between o and w . A standard argument establishes that $o, \nabla f(s)$, and w are then collinear, and o is not between s and

w (because the gradient is proportional to the direction line segment associated with the greatest directional derivative and, according to Lemma 2.3, this direction line segment \overline{sw} lies on the line containing \overline{os} with o not between s and w). So, we have $s = \beta \nabla f(s)$, for $\beta > 0$. Furthermore, a standard multivariable calculus argument establishes that the magnitude of $\nabla f(s)$ (the length of $\overline{o\nabla f(s)}$, i.e., $f(\nabla f(s))$) is the value of the largest directional derivative associated with a direction line segment of length unity—which according to Lemma 2.3 is unity. Thus, $f(\nabla f(s)) = 1$ (the magnitude of $\nabla f(s)$ is unity), so that $s = f(s)\nabla f(s)$ since $f(s)$ is the length of the hypotenuse \overline{os} .

In fact, it is easy to see that the equations of the last sentence of the prior paragraph pertain not just s but to any point t in the plane different from o . They hold trivially if t is a point on the \overline{oa} -axis or \overline{ob} -axis. For any other point t , we can consider \overline{ot} to be the hypotenuse of a right triangle $\Delta\{o, a_t, t\}$, where a_t is the orthogonal projection of t to the \overline{oa} -axis, and then proceed in the same manner as we have already done for $\Delta\{o, a, s\}$ (noting that our axes are unchanged). Also, as stated at the outset, $f(o) = 0$, and lately we have the identification of o as the ordered pair $(0, 0)$. Including the latter, the equations in the last sentence of the prior paragraph constitute a partial differential equation. For t identified in our Cartesian system as (τ_a, τ_b) , it is easily verified by standard differentiation that one solution is $f(t) = \sqrt{\tau_a^2 + \tau_b^2}$. To show that this solution is unique, consider that $t = f(t)\nabla f(t)$ means that t and $\nabla f(t)$ and $(0, 0)$ are collinear. Thus, for the points x on the line containing ot , we have the ordinary differential equation $x = f(x)[df(x)/dx]$ with $f(df(x)/dx) = 1$ and $f(0, 0) = 0$. This has a unique solution, so that the already identified solution, $f(t) = \sqrt{\tau_a^2 + \tau_b^2}$, must be the only solution. Thus, we have derived the Euclidean norm, and hence proved the Pythagorean theorem.

Equations (1.3) and (1.4) could also be used in the definition of the Euclidean norm as follows.

Corollary 2.6. *A function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is the Euclidean norm if and only if it is continuous, vanishes at the origin, and at any other point it satisfies (1.3) and (1.4).*

Equations (1.3) and (1.4) indicate that, from a differentiable viewpoint, the Euclidean norm is a scaling-orientation function in the decomposition of a point as the product of a scalar with a unity-scaled orientation point derived from the function's gradient. We can consider this set of equations to represent "Euclidean decomposition."

But the Euclidean norm does not address the multiplicative structure of an algebra and so does not have an essential role in most algebras. Instead, we shall see that the role of $\nabla f(s)$ on the vector space \mathbb{R}^n is taken up by $\nabla^* f(s^{-1})$ on topological $*$ -algebras over \mathbb{R}^n , and we will consider (1.3) and (1.4) so modified to represent "Jacobian decomposition."

3. Jacobian Decomposition and Inverse Norm

The "defining" equations of the Euclidean norm, (1.3) and (1.4), make no reference to the multiplicative structure of an algebra. Nevertheless, the Euclidean norm has application to the Cayley-Dickson algebras (an unending sequence of real unital topological $*$ -algebras beginning with the only four real normed division algebras, \mathbb{R} , \mathbb{C} , the quaternions, and the octonions). Each of these algebras is characterized by a basis $\{e_0, \dots, e_m\}$, with $m = 2^k - 1$ for any nonnegative integer k . Multiplication is distributive and thus defined by a multiplication table relevant to $\{e_i\}$. In particular, $e_0^2 = 1$ and $e_i^2 = -1$ for $i \neq 0$. Any point $s = \alpha_0 e_0 + \dots + \alpha_m e_m$ (with each $\alpha_i \in \mathbb{R}$) has a conjugate $s^* \equiv \alpha_0 - \dots - \alpha_m e_m$, with $*$ evidently an involution.

In particular, $ss^* = \alpha_0^2 + \cdots + \alpha_m^2 = [f(s)]^2$, where $f(s)$ is the Euclidean norm. Thus, $s^{-1} = s^*/[f(s)]^2$. Hence, the Euclidean norm in this case helps express the inverse of a point. Being the Euclidean norm, $f(s)$ satisfies (1.3) and (1.4), and so also defines a Euclidean decomposition of the point. But given the above involution, it is easy to show that

$$\nabla f(s) = \nabla^* f(s^{-1}), \quad (3.1)$$

(where, as always, $\nabla^* f(s^{-1})$ represents evaluation of the gradient ∇f at the point s^{-1} followed by application of the involution). Substituting the above into (1.3), we obtain $s = f(s)\nabla^* f(s^{-1})$. Since (1.4) holds for all $s \neq 0$, and all such points have inverses, we must also have $f(\nabla^* f(s^{-1})) = 1$. Now we have a formulation for the Euclidean norm that makes reference to unital algebraic structure. On Cayley-Dickson algebras, it is equivalent to the Pythagorean theorem. When this latter decomposition exists on a topological $*$ -algebra but $f(s)$ is not the Euclidean norm, it can be considered to be an algebraic ghost of the Pythagorean theorem.

Definition 3.1. For a topological $*$ -algebra A defined on \mathbb{R}^n , a continuous function $f : \mathbb{C}^n \rightarrow \mathbb{C}$ is an *inverse norm* if it is zero on the nonunits of A , as a function restricted to the domain \mathbb{R}^n it is differentiable on the units of A , and for any unit,

$$s = f(s)\nabla^* f(s^{-1}), \quad (3.2)$$

with

$$f(\nabla^* f(s^{-1})) = 1. \quad (3.3)$$

The above equations mimic the equations for the Euclidean norm referred to in Corollary 2.6, but instead decompose a point as a function's value at the point multiplied by a unity-scaled orientation point dependent on the function's gradient at the *inverse* of the point (or, alternatively to the Euclidean norm's direct expression of a point, (3.2) and (3.3) express the *inverse* of a point—i.e., substituting s^{-1} for s in the latter equations). Thus, we use the term “inverse norm.”

Of course, from (3.1) we have already shown the following.

Theorem 3.2. *The Euclidean norm is an inverse norm on the Cayley-Dickson algebras.*

However, inverse norms have applicability well beyond the Cayley-Dickson algebras.

Theorem 3.3 (Jacobi). *For s a member of the algebra of real matrices $\mathbb{R}^{n \times n}$, let $f(s) \equiv \det(s)$. If s is a unit, then*

$$s = f(s)\nabla^* f(s^{-1}), \quad (3.4)$$

where $*$ indicates matrix transpose.

The above is a well-known immediate consequence of the Jacobi's formula in matrix calculus (the latter expresses gradient of the determinant in terms of the adjugate matrix).

Corollary 3.4. *For the algebra of real matrices $\mathbb{R}^{n \times n}$,*

$$f(s) \equiv \sqrt{n} (\det(s))^{1/n} \quad (3.5)$$

is an inverse norm.

Proof. $f(s)$ is evidently continuous everywhere, as well as differentiable on the units (the invertible matrices), and it vanishes on the nonunits. For any unit s on this algebra, we have $\det(s) \neq 0$, and

$$\nabla^* f(s) = \left[\frac{\partial f(s)}{\partial s_{ij}} \right]^* = \left(\frac{\sqrt{n}}{n} \right) (\det(s))^{-1+1/n} (\nabla(\det(s)))^*. \quad (3.6)$$

For a unit s , (3.4) implies $(\nabla \det(s))^* = s^{-1} \det(s)$. Substituting this into (3.6), and using (3.5), we obtain

$$\nabla^* f(s) = \frac{f(s)}{n} s^{-1} = \frac{s^{-1}}{f(s^{-1})}. \quad (3.7)$$

If we evaluate $\nabla^* f$ on the left-hand-side above at the point s^{-1} instead of evaluating it at s , we obtain (3.2). Applying f in (3.5) to both sides of (3.7) we obtain (3.3). \square

In analogy with Euclidean decomposition (1.3) and (1.4), we can consider the equations of Definition 3.1 to represent "Jacobian decomposition" (i.e., in view of Theorem 3.3).

For the algebra \mathbb{R}^n with component-wise addition and multiplication, it is also easy to show that an inverse norm is given by the product of \sqrt{n} with the geometric mean of the absolute values of the components of a point. That is, for a point $s = (s_1, \dots, s_n)$, set

$$f(s) \equiv \sqrt{n} \left(\prod_{i=1}^n |s_i| \right)^{1/n}. \quad (3.8)$$

Note that f is continuous and f vanishes on the nonunits. If s is a unit, then

$$\frac{\partial f(s)}{\partial s_j} = \frac{\sqrt{n}}{n} \left(\prod_{i=1}^n |s_i| \right)^{-1+1/n} \operatorname{sgn}(s_j) \prod_{i \neq j} |s_i| = \frac{\sqrt{n}}{n} \left(\prod_{i=1}^n |s_i| \right)^{1/n} \frac{1}{s_j} = \frac{f(s)}{n} \left(\frac{1}{s_j} \right). \quad (3.9)$$

It is then a simple task to verify that $f(s)$ satisfies the requirements of Definition 3.1 (with $*$ as the identity).

Now we turn to Jordan algebras.

Theorem 3.5. *The Minkowski Norm is an inverse norm on the spin factor jordan algebra.*

Proof. Thinking of \mathbb{R}^n in the format of $\mathbb{R} \oplus \mathbb{R}^{n-1}$, write its points as $s = w + \mathbf{z}$, with $w \in \mathbb{R}$ and $\mathbf{z} = (z_1, \dots, z_{n-1}) \in \mathbb{R}^{n-1}$. We introduce a multiplication operation such that

$$(\omega_a + \mathbf{z}_1)(\omega_b + \mathbf{z}_2) \equiv (\omega_a\omega_b + \mathbf{z}_1 \cdot \mathbf{z}_2) + (\omega_a\mathbf{z}_2 + \omega_b\mathbf{z}_1), \tag{3.10}$$

where “ \cdot ” is the usual inner product on \mathbb{R}^{n-1} . This multiplication defines a commutative but nonassociative algebra, the spin factor Jordan algebra [5]. The multiplicative identity element is evidently the point where $w = 1$ and $\mathbf{z} = 0$. An inverse element exists for points $w + \mathbf{z}$ such that $\mathbf{z} \cdot \mathbf{z} \neq w^2$. That is, $(w + \mathbf{z})(w + \mathbf{z})^{-1} = 1$ for

$$(w + \mathbf{z})^{-1} = \frac{-w + \mathbf{z}}{-w^2 + \mathbf{z} \cdot \mathbf{z}}, \tag{3.11}$$

when $\mathbf{z} \cdot \mathbf{z} \neq w^2$.

Now we define $f : \mathbb{C}^n \rightarrow \mathbb{C}$ such that for $s = (w, z_1, \dots, z_{n-1}) \in \mathbb{C}^n$,

$$f(s) = \sqrt{-w^2 + z_1^2 + \dots + z_{n-1}^2}, \tag{3.12}$$

where the above square root represents the principal value. On the domain comprised of the units of the spin factor Jordan algebra (the points $w + \mathbf{z}$, $w \in \mathbb{R}$, $\mathbf{z} \in \mathbb{R}^3$, such that $w^2 \neq \mathbf{z} \cdot \mathbf{z}$), we have

$$\nabla f(w + \mathbf{z}) = \frac{-w + \mathbf{z}}{f(w + \mathbf{z})} = (w + \mathbf{z})^{-1} f(w + \mathbf{z}) = \frac{(w + \mathbf{z})^{-1}}{f((w + \mathbf{z})^{-1})}, \tag{3.13}$$

where the three equalities follow from (3.11) and (3.12). Hence, on the units, we have $s^{-1} = f(s^{-1})\nabla f(s)$. Taking the involution $*$ to be the identity, the latter equation is equivalent to the first equation of Definition 3.1.

Applying f to both sides of the first equality in (3.13) and using (3.12), we obtain $f(\nabla f(s)) = 1$ for any unit s . This is equivalent to the second equation of Definition 3.1. \square

On the other hand, the Jordan algebra obtained from the algebra of matrices $\mathbb{R}^{n \times n}$ has the product of two of its members A, B as given by $A \star B \equiv (AB + BA)/2$ where AB and BA indicate the usual matrix product. We then have $\mathbf{1} = A \star A^{-1} = AA^{-1}$, where $\mathbf{1}$ is the identity element in the algebra (in this case, $\mathbf{1} = \text{diag} \{1, 1, \dots, 1\}$). Therefore, A^{-1} is the usual matrix inverse. Consequently, the associated inverse norms are the same as those for the algebra of matrices $\mathbb{R}^{n \times n}$. Thus, Jacobian decomposition holds for the Jordan algebra obtained from the matrix algebra.

Supplying an inverse norm nominally requires solution of a nonlinear partial differential equation (3.2). However, if we apply further restrictions on the nature of $f(s)$, one can obtain a linear equation. In particular, for each algebra example considered up till now there is a constant $\alpha \in \mathbb{R}$ such that

$$f(s)f(s^{-1}) = \alpha, \tag{3.14}$$

so that (3.2) evaluated at s^{-1} instead of s implies (1.6) and

$$\alpha s \nabla^* f(s) = f(s) \mathbf{1}. \quad (3.15)$$

However, not all unital algebras have an inverse norm satisfying (3.15). First, since an inverse norm is zero on nonunits, and a nonunital algebra consists only of nonunits, the inverse norm on a nonunital algebra is identically zero. On the other hand, one might ask whether an inverse norm satisfying (3.15) exists on the unital hull [5] of a nonunital topological algebra.

Theorem 3.6. *The unital hull of a nonunital topological algebra does not have an inverse norm satisfying (3.15).*

Proof. The unital hull of a nonunital algebra A is defined by elements $\hat{s} \equiv (\sigma, s)$ for $\sigma \in \mathbb{R}$ and $s = (s_1, \dots, s_n) \in A$, with component-wise addition of elements and multiplication defined by

$$(\sigma, s) \star (\tau, t) \equiv (\sigma\tau, \sigma t + \tau s + st), \quad (3.16)$$

where st indicates the product between elements of A . The identity element is $\mathbf{1} = (1, 0)$. Equation (3.15) requires that the units satisfy

$$(\sigma, s) \star \left(\frac{\partial f(\hat{s})}{\partial \sigma}, \nabla f(\hat{s}) \right) = \frac{f(\hat{s})}{\alpha} \mathbf{1}, \quad (3.17)$$

where $\nabla \equiv (\partial/\partial s_1, \dots, \partial/\partial s_n)$. Using the multiplication rule, we can write this as

$$\left(\sigma \frac{\partial f(\hat{s})}{\partial \sigma}, \sigma \nabla f(\hat{s}) + s \frac{\partial f(\hat{s})}{\partial \sigma} + s \nabla f(\hat{s}) \right) = \left(\frac{f(\hat{s})}{\alpha}, 0 \right). \quad (3.18)$$

Thus, it is required that $\sigma [\partial f(\hat{s})/\partial \sigma] = f(\hat{s})/\alpha$, so that $f(\hat{s}) = \sigma^{1/\alpha} h(s)$, where $h(s)$ indicates some function independent of σ . In addition, the right-hand-side of (3.18) requires that the second component of the left-hand-side of (3.18) be zero. But with regard to variation in σ , the requirement that $f(\hat{s}) = \sigma^{1/\alpha} h(s)$ means that the first term of this second component is $O(\sigma^{1+1/\alpha})$, the second term is $O(\sigma^{-1+1/\alpha})$, and the third term is $O(\sigma^{1/\alpha})$. It is thus impossible for this second component to remain zero as σ varies unless f is identically zero. But in that case, the equations of Definition 3.1 cannot be satisfied for the units. Thus, an inverse norm satisfying (3.15) does not exist. \square

One can get even more restrictive and consider algebras for which not only is $f(s)f(s^{-1})$ constant on the units (i.e., (3.14) is satisfied) but in addition the units constitute a group on which a multiple of $f(s)$ is homomorphism. In fact, the inverse norm on the first four Cayley-Dickson algebras satisfies this prescription (the latter are the only real normed division algebras by Hurwitz's theorem, and the Euclidean norm is a homomorphism on their units). The inverse norm $f(s) = \sqrt{n}(\det(s))^{1/n}$ on the matrix algebra $\mathbb{R}^{n \times n}$ also satisfies this requirement (i.e., $f(s)/\sqrt{n}$ is a homomorphism on the group of units). Along these

lines, we observe that the Cayley-Dickson algebras and the set of subalgebras of the real matrix algebra $\mathbb{R}^{n \times n}$ overlap on the algebras of real numbers, the complex numbers, and quaternions—which happen to be the only associative real normed division algebras. The Cayley-Dickson algebra sequence contains the only other real normed division algebra—the (noncommutative/nonassociative) algebra of octonions. Thus, the octonions provide an example of an algebra with an inverse norm that is a homomorphism on the group of units, but without a representation as a matrix subalgebra. This point of nonoverlap is typical of the exceptionalism of the octonions [6].

References

- [1] D. Hilbert, *The Foundations of Geometry*, The Open Court, LaSalle, Ill, USA, 1950.
- [2] J. R. Giles, “Classes of semi-inner product spaces,” *Transactions of the American Mathematical Society*, vol. 129, no. 3, pp. 436–446, 1967.
- [3] A. G. Horváth, “Semi-indefinite inner product and generalized Minkowski spaces,” *Journal of Geometry and Physics*, vol. 60, no. 9, pp. 1190–1208, 2010.
- [4] A. G. Horváth, “Premanifolds,” *Note di Matematica*, vol. 31, no. 2, pp. 17–51, 2011.
- [5] K. McCrimmon, *A Taste of Jordan Algebras*, Springer, New York, NY, USA, 2004.
- [6] J. C. Baez, “The octonions,” *Bulletin of the American Mathematical Society*, vol. 39, no. 2, pp. 145–205, 2002.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

