TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI

Defining Orthogonality Leads to Defining Angles

Dr. Ralph deLaubenfels

TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI TSI

Teacher-Scholar Institute

Columbus, Ohio

2013

DEFINING ORTHOGONALITY LEADS TO DEFINING ANGLES

Ralph deLaubenfels

ABSTRACT. A reasonable definition of orthogonality in a normed vector space leads, via orthogonal projections, to two natural definitions of angles between pairs of vectors, one definition with cosine, one with cotangent. Having these definitions coincide is equivalent to the vector space having both the orthogonality and the norm defined by the same inner product $< \cdot >$. We then have the familiar definition of angle and orthogonal projection: the angle between $\vec{a}$ and $\vec{b}$ defined by

$$\cos^{-1}\left(\frac{< \vec{a}, \vec{b} >}{\|\vec{a}\|\|\vec{b}\|}\right) (*)$$

and the orthogonal projection of $\vec{b}$ onto $\vec{a}$ given by

$$P(\vec{b}, \vec{a}) = \left(\frac{< \vec{b}, \vec{a} >}{\|\vec{a}\|^2}\right) \vec{a}.$$

It is shown that this construction of angle from orthogonality provides a very simple proof of the agreement of (*) with any reasonable definition of angle in the plane. This emphasis on orthogonality provides natural motivation for the inner product immediately upon its introduction. One consequence of this approach is a short, natural and geometric (Pythagorean theorem) proof of the Cauchy-Schwarz inequality.

I. INTRODUCTION. Orthogonality rivals calculus both in applicability and conceptual signficance in mathematics, statistics, and physics. In normed vector spaces, orthogonality is usually realized via an *inner product*, which then can be used to define the angle between any pair of vectors with (*) of the Abstract; as a special case, two vectors are orthogonal when their inner product is zero.

But in introducing the inner product, or, more generally, any notion of angle between vectors, it is easier and more natural and motivated, to begin with orthogonality, even though it is merely one particular angle ($\frac{\pi}{2}$). We begin by listing desirable properties, O1–6, of orthogonality. These properties are sufficient to define orthogonal projection, which in turn is sufficient, via right triangles or unit circle, to define angle between a pair of vectors. In fact, angle may be defined with cosine or cotangent. The less familiar cotangent definition has better properties, in general. We spend much time exploring potential bad behaviour, via examples, of these definitions of angle, and present admittedly incomplete and tentative good behaviour. Section II culminates (Theorem 2.17) with a characterization of orthogonality relations that arise in the usual way from an inner product that also determines the norm: it is equivalent to the cosine and cotangent definitions just mentioned agreeing, and is also equivalent to the Pythagorean theorem being valid.

Section III is purely pedagogical, suggesting that the inner, or "dot," product be introduced by starting with orthogonality and the Pythagorean theorem.

II. ORTHOGONALITY TO ANGLES. Assume throughout this paper that V is a normed real vector space. Denote $O(\vec{a}, \vec{b})$ as shorthand for "$\vec{a}$ is orthogonal to $\vec{b}$" $(\vec{a}, \vec{b} \in V)$. Consider the following desirable properties of O.

O1. $O(\vec{a}, \vec{b})$ implies $O(\vec{b}, \vec{a})$.

O2. $O(\vec{a}, \vec{b})$ implies $O(\vec{a}, s\vec{b})$, for any real s.

O3. $O(\vec{a}, \vec{b}), O(\vec{a}, \vec{c})$ implies $O(\vec{a}, (\vec{b} + \vec{c}))$.

O4. $O(\vec{a}, \vec{a})$ if and only if $\vec{a} = \vec{0}$.

O5. For any $\vec{a}$ and two-dimensional subspace W containing $\vec{a}$, there exists nontrivial $\vec{c} \in W$ such that $O(\vec{a}, \vec{c})$.

O6. $O(\vec{a}, \vec{b})$ implies

$$\|\vec{a}\| \leq \|\vec{a} + \vec{b}\|.$$

Note that O1 is stating that the relation O is symmetric, O2 and O3 that, for any $\vec{a}$, the set of all vectors orthogonal to $\vec{a}$ is a vector space, while O5 is guaranteeing an adequate supply of orthogonal vectors. O6 relates orthogonality to norms.

Definition 2.1. If V satisfies O1–4 and $\vec{a}, \vec{b} \in V$, then the *(orthogonal) projection* of $\vec{b}$ onto $\vec{a}$, denoted $P(\vec{b}, \vec{a})$, is $t\vec{a}$, where $t = t(\vec{b}, \vec{a})$ satisfies $O(\vec{a}, (\vec{b} - t\vec{a}))$. The *component* of $\vec{b}$ in the direction $\vec{a}$, denoted $C(\vec{b}, \vec{a})$, is $t\|\vec{a}\|$.

Note that O2–4 imply that $t(\vec{b}, \vec{a})$, if it exists, is unique.

$$P(\vec{b}, \vec{a}) = t(\vec{b}, \vec{a})\, \vec{a}$$

$$C(\vec{b}, \vec{a}) = t(\vec{b}, \vec{a})\, \|\vec{a}\|$$

Proposition 2.2. Suppose V satisfies O1–4.

(a) O5 is equivalent to $P(\vec{b}, \vec{a})$ existing, for all $\vec{a}, \vec{b}$ in V.

(b) Under O5, O6 is equivalent to

$$\|P(\vec{b}, \vec{a})\| \leq \|\vec{b}\|, \;\; \forall \vec{a}, \vec{b}.$$

Proof: (a) Suppose O5 holds and $\vec{a}, \vec{b} \in V$. If $\vec{a}$ and $\vec{b}$ are collinear, then $P(\vec{b}, \vec{a}) = \vec{0}$.

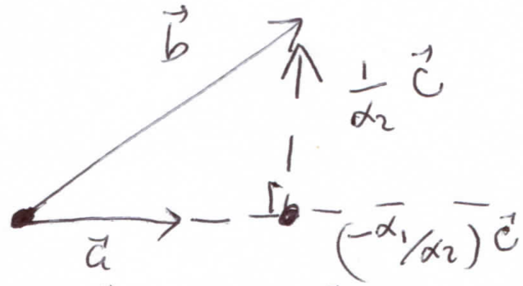
If $\vec{a}$ and $\vec{b}$ are not collinear, let $W \equiv \text{span}(\vec{a}, \vec{b})$. There exists $\vec{c} \in W$ perpendicular to $\vec{a}$. There exist scalars α_1, α_2 so that

$$\vec{c} = \alpha_1 \vec{a} + \alpha_2 \vec{b},$$

which implies

$$\frac{1}{\alpha_2}\vec{c} - \frac{\alpha_1}{\alpha_2}\vec{a} = \vec{b},$$

so that $-\frac{\alpha_1}{\alpha_2}\vec{a} = P(\vec{b}, \vec{a})$, since $\mathrm{O}(\vec{a}, (\vec{b} - (-\frac{\alpha_1}{\alpha_2}\vec{a}))) = O(\vec{a}, \frac{1}{\alpha_2}\vec{c}) = 0$, since $\mathrm{O}(\vec{a}, \vec{c}) = 0$.



Conversely, suppose $P(\vec{b}, \vec{a})$ exists for all $\vec{a}, \vec{b} \in V$ and suppose $\vec{a} \in V$ and W is a two-dimensional subspace containing $\vec{a}$. There exists $\vec{b} \in W$ so that $W = \text{span}(\vec{a}, \vec{b})$; choosing $\vec{c} \equiv (\vec{b} - P(\vec{b}, \vec{a}))$ satisfies O5.

(b) Suppose O6 holds. Then, for any $\vec{a}, \vec{b} \in V$,

$$\|\vec{b}\| = \|P(\vec{b}, \vec{a}) + (\vec{b} - P(\vec{b}, \vec{a}))\| \geq \|P(\vec{b}, \vec{a})\|,$$

by orthogonality.

Conversely, suppose $\|P(\vec{b}, \vec{a})\| \leq \|\vec{b}\|$, for all $\vec{a}, \vec{b} \in V$. If $\vec{a}$ and $\vec{b}$ are orthogonal, then $\vec{a} = P((\vec{a} + \vec{b}), \vec{a})$, thus

$$\|\vec{a}\| \leq \|\vec{a} + \vec{b}\|,$$

as desired. $\square$

Definitions 2.3. Suppose $< \cdot >$ is an inner product on V. The orthogonality relationship on V is *determined by* $< \cdot >$ if

$$O(\vec{a}, \vec{b}) \iff < \vec{a}, \vec{b} >= 0.$$

The norm on V is *determined by* $< \cdot >$ if

$$\|\vec{a}\|^2 =< \vec{a}, \vec{a} >, \quad \forall \vec{a} \in V.$$

Remark 2.4. If the orthogonality is determined by the inner product $< \cdot >$, then O1–O5 are automatically satisfied, with t from Proposition 2.2(a) and Definition 2.1 given by

$$t(\vec{b}, \vec{a}) = \frac{< \vec{a}, \vec{b} >}{< \vec{a}, \vec{a} >}.$$

If the norm is also determined by $< \cdot >$, then O6 also follows, from the Pythagorean theorem. See Theorem 2.17 for a partial converse; see also Open Question 2.18.

Counterexample 2.5. Most orthogonality relations satisfying O1–5 do not come from an inner product. For a class of counterexamples, let's focus on $V \equiv \mathbf{R}^2$, with $(1, 0)$ orthogonal to $(0, 1)$.

Let ϕ be any injective map from $(-\frac{\pi}{2}, \frac{\pi}{2}]$ to itself such that $\phi(0) = \frac{\pi}{2}$ and, for all $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2}], \phi(\phi(\theta)) = \theta$ and $\phi(\theta) \neq \theta$. Define an orthogonality relationship O_ϕ by

$$O_\phi((\alpha \cos\theta_1, \beta \sin\theta_1), (\beta \cos\theta_2, \sin\theta_2))$$

if and only if

$$\alpha\beta = 0 \quad \text{or} \quad \phi(\theta_1) = \theta_2.$$

Only very particular functions ϕ will produce O_ϕ that is determined by an inner product. Since $(1,0)$ and $(0,1)$ are orthogonal, any inner product $< \cdot >$ will have the form

$$<(a,b),(c,d)>= ac<(1,0),(1,0)> +bd<(0,1),(0,1)>,$$

so that, letting

$$\omega \equiv \frac{<(1,0),(1,0)>}{<(0,1),(0,1)>},$$

(a,b) is orthogonal to (c,d) if and only if

$$bd = -\omega ac.$$

Letting $(a,b) \equiv (\cos\theta_1, \sin\theta_1)$, $(c,d) \equiv (\cos\theta_2, \sin\theta_2)$, orthogonality is equivalent to

$$\sin\theta_1 \sin\theta_2 = -\omega\cos\theta_1\cos\theta_2 \text{ or } \tan\theta_2 = -\omega\cot\theta_1.$$

Solving for θ_2, this implies that the orthogonality relationship O_ϕ is determined by an inner product if and only if

$$\phi(\theta) = \begin{cases} \tan^{-1}(-\omega\cot\theta) & \theta \neq 0 \\ \frac{\pi}{2} & \theta = 0 \end{cases}$$

for some positive ω.

The following illustrates how much O6 intertwines orthogonality and norm.

Proposition 2.6. If the norm of V is determined by an inner product and O1–6 holds, then this inner product also determines the orthogonality of V.

Proof: Denote by $< \cdot >$ the inner product determining the norm of V.

Suppose $\vec{a}$ and $\vec{b}$ are orthogonal. Then by O6

$$\|\vec{a}\|^2 + s^2\|\vec{b}\|^2 + 2s<\vec{a},\vec{b}>= \|\vec{a}+s\vec{b}\|^2 \geq \|\vec{a}\|^2 \text{ for all real } s;$$

a simple minimization argument shows that this implies that $<\vec{a},\vec{b}>=0$.

We must also show the converse, that $<\vec{a},\vec{b}>=0$ implies $\vec{a}$ and $\vec{b}$ are orthogonal.

For $t = t(\vec{a},\vec{b})$ as in Definition 2.1 (see Proposition 2.2), since $(\vec{b}-t\vec{a}) = (\vec{b}-P(\vec{b},\vec{a})$ is orthogonal to $\vec{a}$, we now know that

$$0 =<\vec{b}-P(\vec{b},\vec{a}),\vec{a}>=<\vec{b},\vec{a}> -t(\vec{b},\vec{a})<\vec{a},\vec{a}>,$$

so that

$$t(\vec{b},\vec{a}) = \frac{<\vec{b},\vec{a}>}{<\vec{a},\vec{a}>} \quad \text{or} \quad P(\vec{b},\vec{a}) = \frac{<\vec{b},\vec{a}>}{<\vec{a},\vec{a}>}\vec{a}.$$

In particular, if $<\vec{a},\vec{b}>=0$, then $P(\vec{b},\vec{a})=\vec{0}$, so that $\vec{a}$ is orthogonal to $\vec{b}$. □

The picture after Definition 2.1 and the definitions of cosine and cotangent lead to the following definitions of the angle from $\vec{b}$ to $\vec{a}$.

Definitions 2.7. For $\vec{a}, \vec{b}$ nontrivial,

$$\theta_1(\vec{b}, \vec{a}) \equiv \cos^{-1}\left(\frac{C(\vec{b}, \vec{a})}{\|\vec{b}\|}\right) \quad \text{(when defined).}$$

$$\theta_2(\vec{b}, \vec{a}) \equiv \cot^{-1}\left(\frac{C(\vec{b}, \vec{a})}{\|(\vec{b} - P(\vec{b}, \vec{a}))\|}\right).$$

Note that O6 is equivalent to $\theta_1(\vec{b}, \vec{a})$ being defined for all $\vec{a}, \vec{b}$ (see Definition 2.1 and Proposition 2.2).

Remark 2.8. If V has both the norm and orthogonality determined by the innner product $< \cdot >$, then (see Remark 2.4)

$$\theta_1(\vec{b}, \vec{a}) = \cos^{-1}\left(\frac{< \vec{a}, \vec{b} >}{\|\vec{a}\|\|\vec{b}\|}\right) = \theta_1(\vec{a}, \vec{b})$$

and

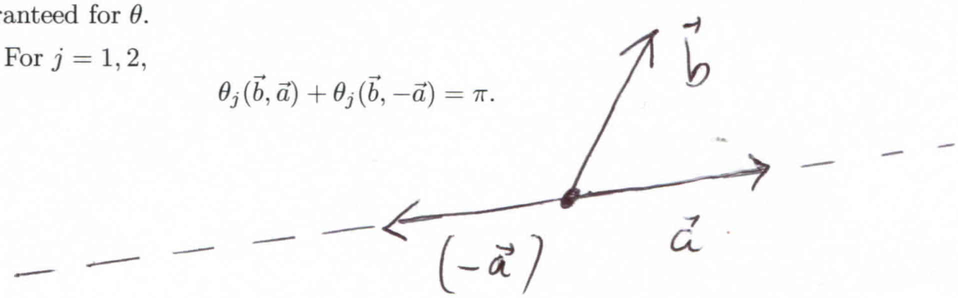
$$\theta_2(\vec{b}, \vec{a}) = \cot^{-1}\left(\frac{< \vec{a}, \vec{b} >}{\|\vec{a}\|\|\vec{b} - \frac{<\vec{a},\vec{b}>}{\|\vec{a}\|^2}\vec{a}\|}\right),$$

which, by the Pythagorean theorem, equals $\theta_1(\vec{b}, \vec{a})$ (see proof of Theorem 2.17(b) $\rightarrow$ (d)).

Since we will soon see that θ may be very poorly behaved, it seems appropriate to mention one good property guaranteed for θ.

Proposition 2.9. For $j = 1, 2$,

$$\theta_j(\vec{b}, \vec{a}) + \theta_j(\vec{b}, -\vec{a}) = \pi.$$



Proof: Since

$$t(\vec{b}, \vec{a})\vec{a} = P(\vec{b}, \vec{a}) = P(\vec{b}, -\vec{a}) = t(\vec{b}, -\vec{a})(-\vec{a}),$$
$$t(\vec{b}, -\vec{a}) = -t(\vec{b}, \vec{a}).$$

This implies that

$$\cos\theta_1(\vec{b}, -\vec{a}) = \frac{t(\vec{b}, -\vec{a})\| - \vec{a}\|}{\|\vec{b}\|} = \frac{-t(\vec{b}, \vec{a})\|\vec{a}\|}{\|\vec{b}\|} = -\cos\theta_1(\vec{b}, \vec{a}),$$

which implies that

$$\theta_1(\vec{b}, \vec{a}) + \theta_1(\vec{b}, -\vec{a}) = \pi.$$

An almost identical argument shows the same result for θ_2. □

In Example 2.10 and More Examples 2.13(a), we will see how badly behaved $\theta_j, j = 1, 2$, can easily be.

Example 2.10. Let V be $\mathbf{R}^2$ with orthogonality defined in the usual way: $\vec{a} \equiv (a_1, a_2)$ is orthogonal to $\vec{b} \equiv (b_1, b_2)$ iff

$$0 = \vec{a} \cdot \vec{b} \equiv (a_1 b_1 + a_2 b_2).$$

For nontrivial real x, y, define

$$\vec{a} \equiv (x, 0), \vec{b} \equiv (x, y).$$

Proposition. (1) $\theta_1(\vec{a}, \vec{b}) = \theta_1(\vec{b}, \vec{a}) \iff$

$$\|(x,y)\|^2 = (x^2 + y^2)\|(1,0)\|^2.$$

(2) $\theta_2(\vec{a}, \vec{b}) = \theta_2(\vec{b}, \vec{a}) \iff$

$$\frac{\|(x,y)\|}{\|(y,-x)\|} = \frac{\|(1,0)\|}{\|(0,1)\|}.$$

(3) $\theta_1(\vec{b}, \vec{a}) = \theta_2(\vec{b}, \vec{a}) \iff$

$$\|(x,y)\|^2 = x^2\|(1,0)\|^2 + y^2\|(0,1)\|^2.$$

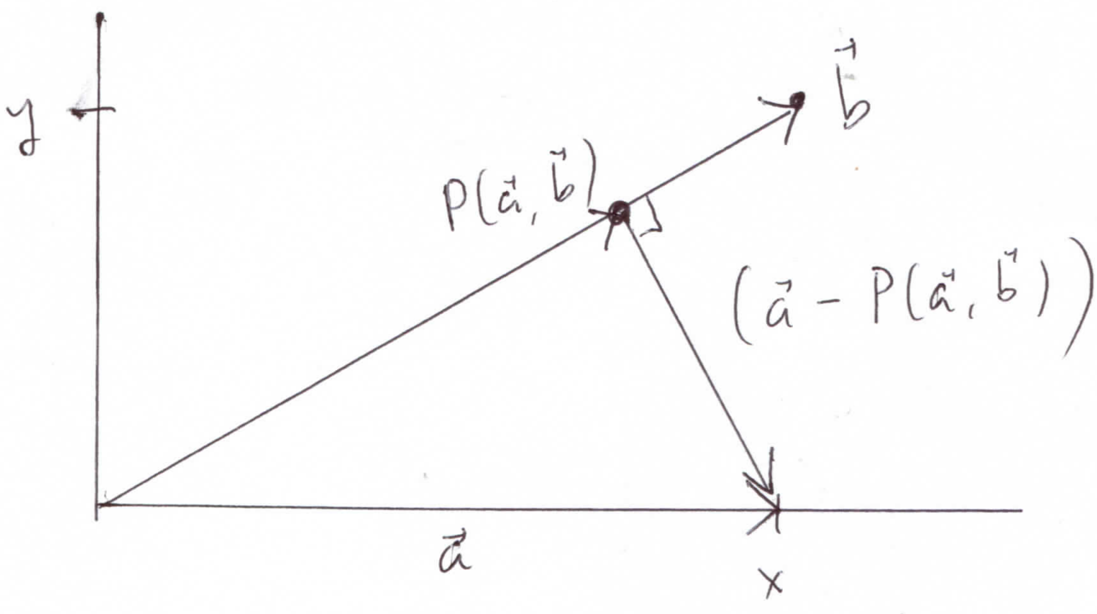
Note in particular that (1) implies (2) and (3), since (1) implies that $\|(0,1)\| = \|(1,0)\|$. Assertion (1) also implies that the same inner product is determining both the norm and the orthogonality relationship.

Proof of Proposition: Clearly $P(\vec{b}, \vec{a}) = \vec{a}$, while a calculation shows that

$$P(\vec{a}, \vec{b}) = \frac{x^2}{x^2 + y^2}(x, y),$$

so that

$$\vec{a} - P(\vec{a}, \vec{b}) = \frac{xy}{x^2 + y^2}(y, -x).$$



Thus

$$\cos\theta_1(\vec{b}, \vec{a}) = \frac{\|(x,0)\|}{\|(x,y)\|} = \frac{x\|(1,0)\|}{\|(x,y)\|}, \quad \text{while} \quad \cos\theta_1(\vec{a}, \vec{b}) = \frac{\|P(\vec{a}, \vec{b})\|}{\|\vec{a}\|} = \frac{x}{x^2 + y^2}\frac{\|(x,y)\|}{\|(1,0)\|},$$

so that (1) follows. Also

$$\cot\theta_2(\vec{b}, \vec{a}) = \frac{\|(x,0)\|}{\|(0,y)\|} = \frac{x}{y}\frac{\|(1,0)\|}{\|(0,1)\|}, \quad \text{while} \quad \cot\theta_2(\vec{a}, \vec{b}) = \frac{\|P(\vec{a}, \vec{b})\|}{\|\vec{a} - P(\vec{a}, \vec{b})\|} = \frac{x}{y}\frac{\|(x,y)\|}{\|(y,-x)\|},$$

giving us (2).

For (3), denote, for $j = 1, 2$, θ_j for $\theta_j(\vec{b}, \vec{a})$. Then

$$\theta_1 = \theta_2 \iff \frac{\|\vec{a}\|^2}{\|\vec{b} - \vec{a}\|^2} = \cot^2\theta_2 = \cot^2\theta_1 = \frac{\cos^2\theta_1}{1 - \cos^2\theta_1} = \frac{\frac{\|\vec{a}\|^2}{\|\vec{b}\|^2}}{1 - \frac{\|\vec{a}\|^2}{\|\vec{b}\|^2}} = \frac{\|\vec{a}\|^2}{\|\vec{b}\|^2 - \|\vec{a}\|^2}$$

$$\iff \frac{\|(x,0)\|^2}{\|(0,y)\|^2} = \frac{\|(x,0)\|^2}{\|(x,y)\|^2 - \|(x,0)\|^2} \iff \|(x,y)\|^2 = \|(x,0)\|^2 + \|(0,y)\|^2 = x^2\|(1,0)\|^2 + y^2\|(0,1)\|^2.$$

$\square$

Note that the norm in (3) of the Proposition above is determined by the inner product

$$< (x_1, y_1), (x_2, y_2) > \equiv x_1 x_2 \|(1,0)\|^2 + y_1 y_2 \|(0,1)\|^2.$$

Proposition 2.6 now implies that, in order that O6 hold, $\|(1,0)\|$ must equal $\|(0,1)\|$.

Here are some particular choices of norm.

(a) First take the norm defined by the inner product

$$< \vec{a}, \vec{b} > \equiv 2a_1 b_1 + a_2 b_2 : \qquad \|(x,y)\|^2 \equiv 2x^2 + y^2.$$

Then

$$\cot\theta_2(\vec{a}, \vec{b}) = \frac{x}{y}\sqrt{\frac{(2x^2 + y^2)}{(x^2 + 2y^2)}}, \quad \text{while} \quad \cot\theta_2(\vec{b}, \vec{a}) = \frac{\sqrt{2}x}{y};$$

not equal.

$$\cos\theta_1(\vec{a}, \vec{b}) = \frac{x}{x^2 + y^2}\sqrt{x^2 + \frac{1}{2}y^2}, \quad \text{while} \quad \cos\theta_1(\vec{b}, \vec{a}) = \sqrt{\frac{2x^2}{2x^2 + y^2}};$$

also not equal.

$$\theta_2(\vec{b}, \vec{a}) = \theta_1(\vec{b}, \vec{a}), \quad \text{but} \quad \theta_2(\vec{a}, \vec{b}) \neq \theta_1(\vec{a}, \vec{b}).$$

Thus even in a space with norm and orthogonality determined by inner products we see not-so-good behaviour of θ_1 and θ_2, because the inner product determining the norm is different than the inner product determining the orthogonality.

Note that, by Proposition 2.6, O6 must fail, for some $\vec{a}, \vec{b}$ in $\mathbf{R}^2$. An example is

$$\vec{a} \equiv 2(1,1), \ \vec{b} \equiv (-1,1);$$

then $\vec{a}$ is orthogonal to $\vec{b}$, but

$$\|\vec{a}\|^2 = 12, \|\vec{a} + \vec{b}\|^2 = \|(1,3)\|^2 = 11.$$

(b) Take the ℓ^∞ norm

$$\|(x,y)\| \equiv \max|x|, |y|.$$

Then

$$\theta_2(\vec{a}, \vec{b}) = \cot^{-1}\left(\frac{x}{y}\right) = \theta_2(\vec{b}, \vec{a}),$$

while, for $|y| \leq x$,

$$\theta_1(\vec{a}, \vec{b}) = \cos^{-1}\left(\frac{x^2}{x^2 + y^2}\right), \quad \theta_1(\vec{b}, \vec{a}) = 0.$$

The value of zero for $\theta_1(\vec{b}, \vec{a})$ is particularly counterintuitive: an angle of zero between $\vec{a}$ and $\vec{b}$, with neither vector a multiple of the other. Do we call $\vec{a}$ and $\vec{b}$ parallel because the angle between them is zero?

(c) Now take the ℓ^1 norm

$$\|(x,y)\| \equiv |x| + |y|.$$

Then again

$$\theta_2(\vec{a}, \vec{b}) = \cot^{-1}\left(\frac{x}{y}\right) = \theta_2(\vec{b}, \vec{a}),$$

while, for $|y| \leq x$,

$$\theta_1(\vec{a}, \vec{b}) = \cos^{-1}\left(\frac{x(x+y)}{x^2+y^2}\right), \quad \theta_1(\vec{b}, \vec{a}) = \cos^{-1}\left(\frac{x}{x+y}\right).$$

For a very familiar special case, take $x = y > 0$, so that

$$\theta_1((1,1),(1,0)) = \theta_1(\vec{b}, \vec{a}) = \cos^{-1}\left(\frac{\|(0,1)\|}{\|(1,1)\|}\right) = \cos^{-1}(\frac{1}{2}) = \frac{\pi}{3},$$

while

$$\theta_1((1,0),(1,1)) = \theta_1(\vec{a}, \vec{b}) = \cos^{-1}\left(\frac{\|(\frac{1}{2},\frac{1}{2})\|}{\|(1,0)\|}\right) = \cos^{-1}(\frac{1}{1}) = 0;$$

besides again producing an angle of zero between two vectors that are not collinear, this illustrates that θ_1 is not symmetric; that is, $\theta_1(\vec{b}, \vec{a})$ does not equal $\theta_1(\vec{a}, \vec{b})$. We have also seen above that θ_2 is not symmetric.

θ_1 also fails to be additive: again with $x = y > 0$,

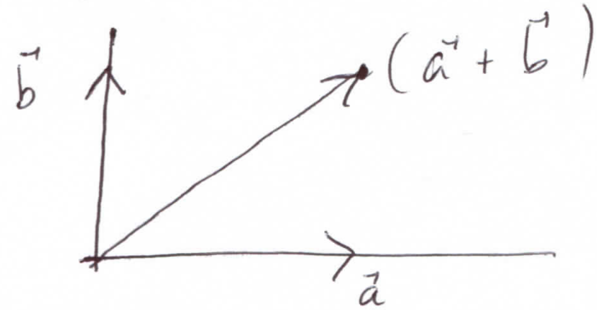
$$\theta_1(\vec{a}, \vec{b}) + \theta_1(\vec{b}, \vec{b} - \vec{a}) = \theta_1((1,0),(1,1)) + \theta_1((1,1),(0,1)) = \frac{\pi}{4} + \cos^{-1}\left(\frac{\|(1,0)\|}{\|(1,1)\|}\right) = \frac{\pi}{4} + \frac{\pi}{3} \neq \frac{\pi}{2}$$
$$= \theta_1((1,0),(0,1)) = \theta_1(\vec{a}, \vec{b} - \vec{a}).$$

It is interesting that θ_2 is additive in a sense very similar to that just mentioned.

Proposition 2.11. Suppose V is a normed vector space with an orthogonality relation O satisfying O1–5. Then

$$\theta_2(\vec{a} + \vec{b}, \vec{a}) + \theta_2(\vec{a} + \vec{b}, \vec{b}) = \frac{\pi}{2},$$

whenever $\vec{a}$ is orthogonal to $\vec{b}$.



Proof: For $\vec{a}$ orthogonal to $\vec{b}$,

$$\cot\left(\theta_2(\vec{a} + \vec{b}, \vec{a})\right) \equiv \frac{\|P(\vec{a} + \vec{b}, \vec{a})\|}{\|\vec{a} + \vec{b} - P(\vec{a} + \vec{b}, \vec{a})\|} = \frac{\|\vec{a}\|}{\|\vec{b}\|};$$

similarly,

$$\cot\left(\theta_2(\vec{a} + \vec{b}, \vec{b})\right) = \frac{\|\vec{b}\|}{\|\vec{a}\|}.$$

Since

$$\cot\left(\theta_2(\vec{a} + \vec{b}, \vec{a})\right) \cot\left(\theta_2(\vec{a} + \vec{b}, \vec{b})\right) = 1,$$

$\left(\theta_2(\vec{a} + \vec{b}, \vec{a})\right) + \left(\theta_2(\vec{a} + \vec{b}, \vec{b})\right) = \frac{\pi}{2}$ (using the trig identity $\cot(\phi + \psi) = \frac{\cot\phi \cot\psi - 1}{\cot\phi + \cot\psi}$). $\square$

We shall see (Theorem 2.17) that having θ_1 be symmetric and additive is equivalent to V having an inner product that determines both the orthogonality and the norm.

The following is essentially a consequence of (3) of the Proposition in Example 2.10.

Proposition 2.12. If the orthogonality relation O is determined by an inner product and

$$\theta_1(\vec{b}, \vec{a}) = \theta_1(\vec{a}, \vec{b})$$

for all $\vec{a}, \vec{b}$ in V, then there exists an inner product that determines both the norm and the orthogonality of V.

Proof: Denote by $\|\cdot\|_V$ the norm in V, and by $< \cdot >$ the inner product that determines the orthogonality in V. To show that the norm of V is determined by an inner product, it is sufficient to show it satisfies the parallelogram law. So fix $\vec{a}, \vec{b} \in V$. Define

$$\vec{v} \equiv \frac{\vec{a}}{\sqrt{< \vec{a}, \vec{a} >}}, \quad \vec{w} \equiv \frac{(\vec{b} - P(\vec{b}, \vec{a}))}{\sqrt{< (\vec{b} - P(\vec{b}, \vec{a})), (\vec{b} - P(\vec{b}, \vec{a})) >}}.$$

Then $< \vec{v}, \vec{v} > = < \vec{w}, \vec{w} > = 1$ and $< \vec{v}, \vec{w} > = 0$ (equivalently, $\vec{v}$ is orthogonal to $\vec{w}$ in V). Define $U : \mathbf{R}^2 \to \text{span}(\vec{a}, \vec{b})$ by

$$U((x, y)) \equiv (x\vec{v} + y\vec{w}) \quad (x, y \in \mathbf{R}).$$

Have $\mathbf{R}^2$ inherit the norm and orthogonality of V:

$$\|(x, y)\| \equiv \|U((x, y))\|_V, \quad (x_1, y_1) \text{ orthogonal to } (x_2, y_2) \iff < U((x_1, y_1)), U((x_2, y_2)) > = 0.$$

A calculation shows that

$$< U((x_1, y_1)), U((x_2, y_2)) > = x_1x_2 + y_1y_2 \equiv (x_1, y_1) \cdot (x_2, y_2),$$

the usual inner product in $\mathbf{R}^2$, thus we have the usual orthogonality in $\mathbf{R}^2$, as in Example 2.10.

$\theta_1((x, y), (x, 0)) = \theta_1(U(x, y), U(x, 0)) = \theta_1(U(x, 0), U(x, y)) = \theta_1((x, 0), (x, y))$ for all nontrivial real x, y, by hypothesis. By (1) of the Proposition in Example 2.10, for any real x, y,

$$\|x\vec{v} + y\vec{w}\|_V^2 = \|U((x, y))\|^2 = \|(x, y)\|^2 = (x^2 + y^2)\|(1, 0)\|^2 = (x^2 + y^2)\|U((1, 0))\|^2 = (x^2 + y^2)\|\vec{v}\|_V^2$$
$$= < (x\vec{v} + y\vec{w}), (x\vec{v} + y\vec{w}) >_2,$$

where the inner product $< \cdot >_2$ is defined by

$$< (x_1\vec{v} + y_1\vec{w}), (x_2\vec{v} + y_2\vec{w}) >_2 \equiv x_1x_2\|\vec{v}\|^2 + y_1y_2 \quad (x_1, x_2, y_1, y_2 \in \mathbf{R}).$$

In particular, this shows that $\text{span}(\vec{a}, \vec{b})$ has its norm determined by an inner product, hence satisfies the parallelogram law:

$$\|\vec{a} + \vec{b}\|^2 + \|\vec{a} - \vec{b}\|^2 = 2\left(\|\vec{a}\|^2 + \|\vec{b}\|^2\right).$$

Since $\vec{a}, \vec{b}$ are arbitrary vectors in V, this shows that the norm of V is determined by an inner product. The remainder of the proposition follows from Proposition 2.6, after we observe that O6 is satisfied since $\theta_1(\vec{a}, \vec{b})$ is defined for all $\vec{a}, \vec{b}$ in V. □

In Examples 2.13, (a) gives more examples of badly behaved angles, while (b) demonstrates bad behavior of orthogonality.

More Examples 2.13. (a) Take the usual orthogonality in $\mathbf{R}^2$, as in Example 2.10, and define

$$\vec{a} \equiv \frac{5}{3}(1, 1), \quad \vec{b} \equiv (2, 1).$$

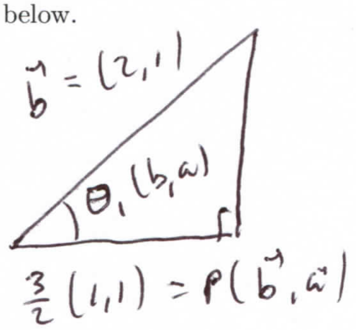
Then

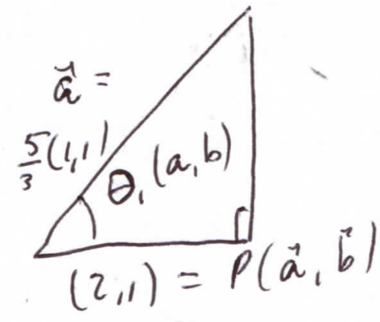
$$P(\vec{b}, \vec{a}) = \frac{3}{2}(1, 1), \ \vec{b} - P(\vec{b}, \vec{a}) = \frac{1}{2}(1, -1), \ \ P(\vec{a}, \vec{b}) = \vec{b}, \ \vec{a} - P(\vec{a}, \vec{b}) = \frac{1}{3}(-1, 2),$$

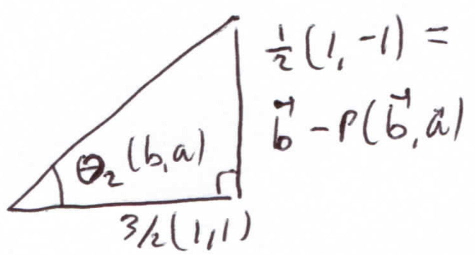
thus

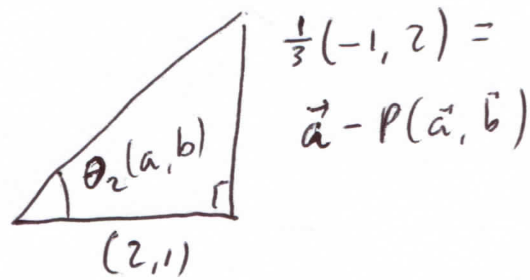
$$\cos\left(\theta_1(\vec{a}, \vec{b})\right) = \frac{\|(2, 1)\|}{\frac{5}{3}\|(1, 1)\|}, \ \cos\left(\theta_1(\vec{b}, \vec{a})\right) = \frac{\frac{3}{2}\|(1, 1)\|}{\|(2, 1)\|}, \ \cot\left(\theta_2(\vec{a}, \vec{b})\right) = \frac{\|(2, 1)\|}{\frac{1}{3}\|(-1, 2)\|}, \ \cot\left(\theta_2(\vec{b}, \vec{a})\right) = \frac{\frac{3}{2}\|(1, 1)\|}{\frac{1}{2}\|(1, -1)\|},$$

as drawn below.









In ℓ^1,

$$\theta_1(\vec{b},\vec{a}) = 0, \theta_1(\vec{a},\vec{b}) = \cos^{-1}\left(\frac{9}{10}\right), \theta_2(\vec{b},\vec{a}) = \cot^{-1}(3) = \theta_2(\vec{a},\vec{b}).$$

In ℓ^∞,

$$\theta_1(\vec{b},\vec{a}) = \cos^{-1}\left(\frac{3}{4}\right), \theta_1(\vec{a},\vec{b}) \text{ is undefined}, \theta_2(\vec{b},\vec{a}) = \cot^{-1}(3) = \theta_2(\vec{a},\vec{b}).$$

(b) When leaving the usual orthogonality of Example 2.10, it is surprisingly difficult to get all of O1–6 satisfied. It will simplify discussion to write

$$(x_1, y_1) \cdot (x_2, y_2) \equiv x_1 x_2 + y_1 y_2$$

for the usual inner product in $\mathbf{R}^2$, defining the usual orthogonality

$$\vec{x} \text{ orthogonal to } \vec{y} \iff \vec{x} \cdot \vec{y} = 0.$$

Consider $\mathbf{R}^2$ with the ℓ^p norm

$$\|(x,y)\|_p^p \equiv |x|^p + |y|^p,$$

for $1 \le p < \infty$.

In defining $\mathrm{O}((x_1,y_1),(x_2,y_2))$, to get O6 we need (x_2,y_2) parallel to the level curve

$$g(x,y) \equiv |x|^p + |y|^p \equiv \|(x,y)\|_p^p = \|(x_1,y_1)\|_p^p;$$

for $x_1, y_1 > 0$, since the gradient of $g(x,y) \equiv x^p + y^p$ is $\nabla g(x,y) = (px^{p-1}, py^{p-1})$, we need, since $(y_1^{p-1}, -x_1^{p-1}) \cdot \nabla g(x_1,y_1) = 0$,

$$(x_2, y_2) \text{ a multiple of } (y_1^{p-1}, -x_1^{p-1}).$$

However, for $p \ne 2$, this definition of $\mathrm{O}((x_1,y_1),(x_2,y_2))$ does not satisfy O1, in general.

Take, in particular, $p = 1$. The largest relation satisfying O6 (it also satisfies O2 and O4) has

$$O(\vec{a}, t(1,-1)) \ (t \text{ real }) \quad \text{for } \vec{a} \text{ in the closed first or third quadrant } xy \ge 0;$$

$$O(\vec{a}, t(1,1)) \ (t \text{ real }) \quad \text{for } \vec{a} \text{ in the closed second or fourth quadrant } xy \le 0.$$

Notice that this orthogonality relation does not satisfy O3; for example, if $\vec{a} \equiv (1,0), \vec{b} \equiv (2,2), \vec{c} \equiv (1,-1)$.

For the sake of counterintuition, consider, as in Example 2.10,

$$\vec{a} \equiv (1,0), \vec{b} \equiv (1,1).$$

Then $P(\vec{b},\vec{a})$ appears to equal $2\vec{a}$, since $\vec{b}-2\vec{a}=(-1,1)$

$$\theta_1(\vec{b},\vec{a})=\cos^{-1}\left(\frac{\|2\vec{a}\|}{\|\vec{b}\|}\right)=0,$$

while

$$\theta_2(\vec{b},\vec{a})=\cot^{-1}\left(\frac{\|2\vec{a}\|}{\|(-1,1)\|}\right)=\frac{\pi}{4}.$$

On the other hand, $\vec{b}$ is orthogonal to $\vec{a}$, so $P(\vec{b},\vec{a})$ should equal $\vec{0}$.

The largest relation satisfying O2–6 has

$$O(\vec{a},t(1,-1))\ (t \text{ real })\quad \text{for } \vec{a} \text{ in the open first or third quadrant } xy>0;$$
$$O(\vec{a},t(1,1))\ (t \text{ real })\quad \text{for } \vec{a} \text{ in the open second or fourth quadrant } xy<0.$$

In order that O satisfy both O1 and O6, it would consist, in its entirety, of

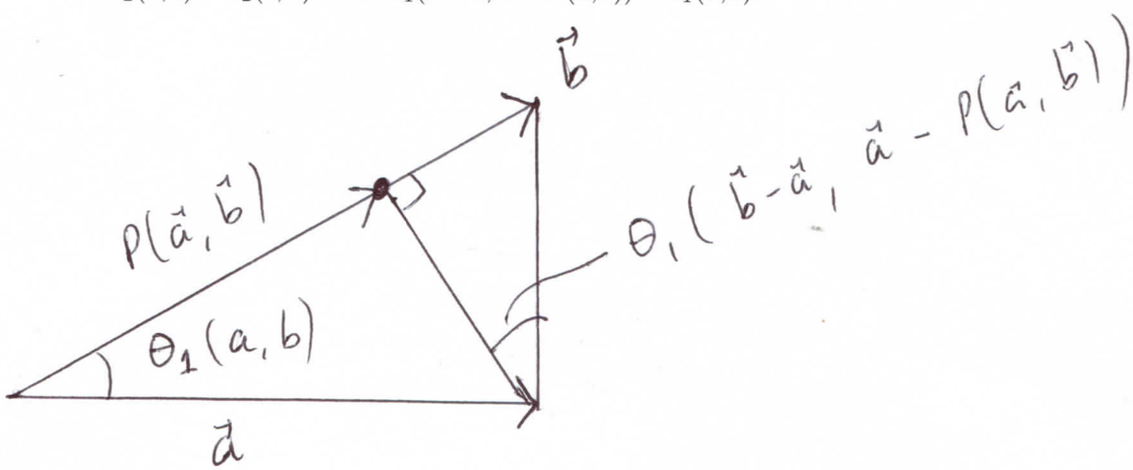
$$O(s(1,1),t(1,-1)),O(s(1,-1),t(1,1)),\ \ s,t \text{ real }.$$

It is not hard to see that O5 would then fail.

Before focusing on ideal behaviour, let's relate symmetry of θ_2 to geometry of θ_1.

Proposition 2.14. Suppose $\vec{a},\vec{b}$ are vectors such that $\vec{a}=P(\vec{b},\vec{a})$ and $\theta_1(\vec{b}-\vec{a},\vec{a}-P(\vec{a},\vec{b}))$ and $\theta_1(\vec{a},\vec{b})$ are defined. Then

$$\theta_2(\vec{a},\vec{b})=\theta_2(\vec{b},\vec{a}) \iff \theta_1(\vec{b}-\vec{a},\vec{a}-P(\vec{a},\vec{b}))=\theta_1(\vec{a},\vec{b}).$$



Proof:

$$\cot\left(\theta_2(\vec{a},\vec{b})\right)\equiv\frac{\|P(\vec{a},\vec{b})\|}{\|\vec{a}-P(\vec{a},\vec{b})\|}=\frac{\|\vec{a}\|}{\|\vec{b}-\vec{a}\|}\equiv\cot\left(\theta_2(\vec{b},\vec{a})\right)$$

if and only if

$$\cos\left(\theta_1(\vec{a},\vec{b})\right)\equiv\frac{\|P(\vec{a},\vec{b})\|}{\|\vec{a}\|}=\frac{\|\vec{a}-P(\vec{a},\vec{b})\|}{\|\vec{b}-\vec{a}\|}\equiv\cos\left(\theta_1(\vec{b}-\vec{a},\vec{a}-P(\vec{a},\vec{b})\right).$$

□

Definition 2.15. (V,O) *satisfies the Pythagorean theorem* if

$$\|\vec{a}+\vec{b}\|^2=\|\vec{a}\|^2+\|\vec{b}\|^2$$

whenever $\vec{a}$ and $\vec{b}$ are orthogonal.

Note that O6 follows from the Pythagorean theorem.

The following is surely well known, but its proof is included for completeness.

Lemma 2.16 If V satisfies the Pythagorean theorem, for some O satisfying O1–5, then the norm of V is determined by an inner product.

Proof: It is sufficient to prove the parallelogram law

$$\|\vec{a}+\vec{b}\|^2 + \|\vec{a}-\vec{b}\|^2 = 2\left(\|\vec{a}\|^2 + \|\vec{b}\|^2\right).$$

Denote by t_+ the unique number from O5 such that

$$(\vec{b} - t_+\vec{a}) \text{ is orthogonal to } \vec{a};$$

by t_- the unique number such that

$$(-\vec{b} - t_-\vec{a}) \text{ is orthogonal to } \vec{a};$$

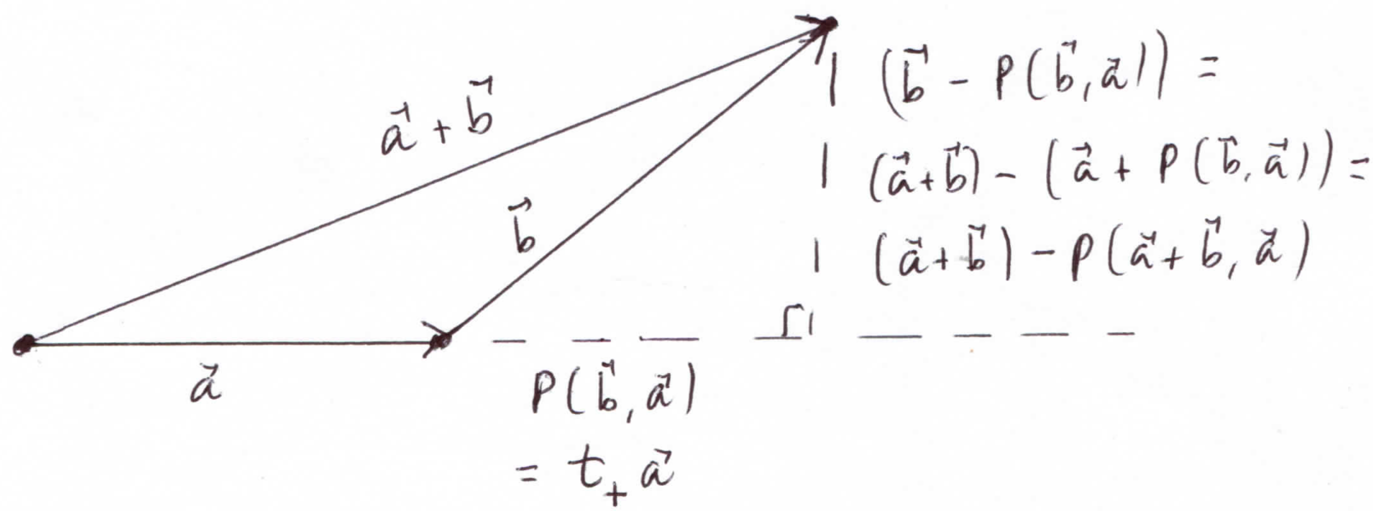
that is, $P(\vec{b},\vec{a}) = t_+\vec{a}, P(-\vec{b},\vec{a}) = t_-\vec{a}$.

In the picture below, notice that

$$(\vec{a}+\vec{b}) - (\vec{a} + P(\vec{b},\vec{a})) = (\vec{b} - P(\vec{b},\vec{a})),$$

thus

$$P((\vec{a}+\vec{b}),\vec{a}) = \vec{a} + P(\vec{b},\vec{a}).$$



By the Pythagorean theorem,

$$\|(\vec{a}+\vec{b})\|^2 = \|\vec{a} + P(\vec{b},\vec{a})\|^2 + \|\vec{b} - P(\vec{b},\vec{a})\|^2$$

and

$$\|\vec{b}\|^2 = \|P(\vec{b},\vec{a})\|^2 + \|\vec{b} - P(\vec{b},\vec{a})\|^2;$$

subtracting gives

$$\|(\vec{a}+\vec{b})\|^2 = \|\vec{b}\|^2 + \|\vec{a}+P(\vec{b},\vec{a})\|^2 - \|P(\vec{b},\vec{a})\|^2 = \|\vec{b}\|^2 + \|(1+t_+)\vec{a}\|^2 - \|t_+\vec{a}\|^2 = \|\vec{b}\|^2 + \left[(1+t_+)^2 - t_+^2\right]\|\vec{a}\|^2$$
$$= \|\vec{b}\|^2 + (1+2t_+)\|\vec{a}\|^2.$$

Identically,

$$\|(\vec{a}-\vec{b})\|^2 = \|\vec{b}\|^2 + (1+2t_-)\|\vec{a}\|^2,$$

so that

$$\|\vec{a}+\vec{b}\|^2 + \|\vec{a}-\vec{b}\|^2 = 2\left((1+t_+ + t_-)\|\vec{a}\|^2 + \|\vec{b}\|^2\right).(*)$$

By O2, $(-\vec{b} + t_+\vec{a}) = -(\vec{b} - t_+\vec{a})$ is orthogonal to $\vec{a}$. Thus $t_- = -t_+$, so that (*) implies the parallelogram law.

Since the norm of V satisfies the paralleogram law, it is determined by an inner product. $\square$

Theorem 2.17. Suppose (V, O) is as in O1–5. Then the following are equivalent.
(a) $\theta_1(\vec{a}+\vec{b},\vec{a})+\theta_1(\vec{a}+\vec{b},\vec{b})=\frac{\pi}{2}$, when $\vec{a}$ is orthogonal to $\vec{b}, \vec{a}, \vec{b}$ nontrivial.
(b) (V, O) satisfies the Pythagorean theorem.
(c) There exists an inner product $< \cdot,\cdot >$ that determines both the norm and the orthogonality relation on V.
(d) $\theta_1(\vec{b},\vec{a})=\theta_2(\vec{b},\vec{a})$ for all nontrivial $\vec{a}, \vec{b}$ in V.

If any of (a)–(d) hold, we then have, for any $\vec{a}, \vec{b}$ in V, the orthogonal projection of $\vec{b}$ onto $\vec{a}$ given by

$$P(\vec{b},\vec{a})=\left(\frac{<\vec{b},\vec{a}>}{\|\vec{a}\|^2}\right)\vec{a}.$$

Proof: For the equivalence of (a) and (b), assume $\vec{a}$ is orthogonal to $\vec{b}$ and $\vec{c}\equiv(\vec{a}+\vec{b})$. Let $\phi_1\equiv\theta_1(\vec{c},\vec{a}), \phi_2\equiv\theta_1(\vec{c},\vec{b})$.

Then, applying cosine to both sides, (a) is equivalent to

$$\cos(\phi_1)\cos(\phi_2)=\sin(\phi_1)\sin(\phi_2);$$

squaring both sides, this is equivalent to

$$\cos^2(\phi_1)\cos^2(\phi_2)=(1-\cos^2(\phi_1))(1-\cos^2(\phi_2))=1-\cos^2(\phi_1)-\cos^2(\phi_2)+\cos^2(\phi_1)\cos^2(\phi_2),$$

or

$$\cos^2(\phi_1)+\cos^2(\phi_2)=1;$$

by definition, this is equivalent to

$$1=\left(\frac{\|P(\vec{c},\vec{a})\|}{\|\vec{c}\|}\right)^2+\left(\frac{\|P(\vec{c},\vec{b})\|}{\|\vec{c}\|}\right)^2=\frac{\|\vec{a}\|^2}{\|\vec{c}\|^2}+\frac{\|\vec{b}\|^2}{\|\vec{c}\|^2},$$

which is equivalent to

$$\|\vec{c}\|^2=\|\vec{a}\|^2+\|\vec{b}\|^2,$$

the Pythagorean theorem.
(b) $\to$ (c). The existence of an inner product $< \cdot >$ that determines the norm of V follows from Lemma 2.16. (b) also clearly implies O6. Proposition 2.6 now gives us (c).
(c) $\to$ (b) follows from

$$\|\vec{a}\|^2+2<\vec{a},\vec{b}>+\|\vec{b}\|^2=\|\vec{a}+\vec{b}\|^2.$$

(d) $\to$ (a) follows from Proposition 2.11.
(b) $\to$ (d). The Pythagorean theorem implies that

$$\|\vec{b}\|^2=\|P(\vec{b},\vec{a})\|^2+\|\vec{b}-P(\vec{b},\vec{a})\|^2,$$

thus

$$\left[\sin\left(\theta_1(\vec{b},\vec{a})\right)\right]^2=1-\left[\cos\left(\theta_1(\vec{b},\vec{a})\right)\right]^2=1-\frac{\|P(\vec{b},\vec{a})\|^2}{\|\vec{b}\|^2}=\frac{\|\vec{b}\|^2-\|P(\vec{b},\vec{a})\|^2}{\|\vec{b}\|^2}=\frac{\|\vec{b}-P(\vec{b},\vec{a})\|^2}{\|\vec{b}\|^2},$$

so that

$$\cot\left(\theta_1(\vec{b},\vec{a})\right)\equiv\frac{\cos\left(\theta_1(\vec{b},\vec{a})\right)}{\sin\left(\theta_1(\vec{b},\vec{a})\right)}=\frac{\left(\frac{C(\vec{b},\vec{a})}{\|\vec{b}\|}\right)}{\left(\frac{\|\vec{b}-P(\vec{b},\vec{a})\|}{\|\vec{b}\|}\right)}=\frac{C(\vec{b},\vec{a})}{\|\vec{b}-P(\vec{b},\vec{a})\|}\equiv\cot\left(\theta_2(\vec{b},\vec{a})\right),$$

as desired. $\square$

Note that (a)–(d) of Theorem 2.17 all imply O6; recall in particular that $\theta_1(\vec{b}, \vec{a})$ being defined for all $\vec{a}, \vec{b}$ is equivalent to O6.

Open Questions 2.18. Suppose O1–6 are satisfied.

(1) Does there exist an inner product that determines orthogonality in V? Recall (Counterexample 2.5) that O1–5 do not imply such an inner product.

(2) Must the norm in O6 be determined by an inner product?

(3) If orthogonality is determined by an inner product, must the norm in O6 be determined by an inner product? (Compare with Proposition 2.6).

If (2) is true, Proposition 2.6 implies (a)–(d) of Theorem 2.17.

III. A SMALL BIT OF PEDAGOGY ABOUT THE DOT PRODUCT. There appear to be two standard methods of introducing the dot (also known as inner) product, at least in $\mathbf{R}^3$.

One approach is geometric:

$$\vec{a} \cdot \vec{b} \equiv \|\vec{a}\| \|\vec{b}\| \cos\theta \quad (*),$$

where θ is the (smaller) angle between $\vec{a}$ and $\vec{b}$.

This definition makes sense only when angle makes sense, which is really only in $\mathbf{R}^2$. In $\mathbf{R}^3$, one could sketch the plane containing $\vec{a}$ and $\vec{b}$ and pretend it's $\mathbf{R}^2$. In $\mathbf{R}^n, n > 3$, this definition does not make sense. For the vector spaces of most interest, the infinite-dimensional ones, this definition definitely makes no sense; what is the angle between two functions on $[0,1]$?

This definition is also unmotivated; angle might be of interest and norm might be of interest, but why make this particular combination of angle and norm into a formal definition? When one relates the dot product to orthogonal projections and components, it might become of interest just because orthogonal projections and components are of interest, in physics and minimization problems. But the item of interest is *orthogonality*, that is, the particular angle of $\frac{\pi}{2}$; it seems unnatural to insert arbitrary angles into a definition that appears to be used only for the angle of $\frac{\pi}{2}$.

The second approach is algebraic:

$$(a_1, a_2, a_3) \cdot (b_1, b_2, b_3) \equiv (a_1 b_1 + a_2 b_2 + a_3 b_3) \quad (**).$$

Except for a desire to have some sort of definition of vector product, after seeing vector sum and scalar multiplication, this definition is *definitely* unmotivated; why this particular arrangement of the components of $\vec{a}$ and $\vec{b}$?

Of course the ultimate goal is the equality of (*) and (**); this is the equating of geometry (intuition) and algebra (precision) that has been arguably the dominant and most successful theme of the many amazing successes of western mathematics. But one likes to spend as little time as possible on completely unmotivated definitions and exposition.

The pedagogy suggested here begins by introducing, via drawings and hand gestures, orthogonality, and its motivations as mentioned above. With the Pythagorean theorem in mind, applied to a triangle with sides $\vec{a}, \vec{b}$ and $(\vec{a} + \vec{b})$, it takes only a moment to expand $\|\vec{a} + \vec{b}\|^2$ and see the appearance of $\vec{a} \cdot \vec{b}$, as in (**), as the extra stuff that appears in addition to $\|\vec{a}\|^2$ and $\|\vec{b}\|^2$. Taking as our fundamental axiom the Pythagorean theorem, the characterization of orthogonality as the dot product being zero is an immediate consequence. The orthogonal projection $P(\vec{b}, \vec{a})$ is defined geometrically as the multiple of $\vec{a}$, call it $t\vec{a}$, such that

$$(\vec{b} - t\vec{a}) \text{ is orthogonal to } \vec{a}.$$

(See Definition 2.1.)

The algebraic characterization of orthogonality and a small bit of algebra then quickly implies that

$$t = \frac{< \vec{b}, \vec{a} >}{< \vec{a}, \vec{a} >} = \frac{< \vec{b}, \vec{a} >}{\|\vec{a}\|^2}.$$

The Cauchy-Schwarz inequality follows now from the Pythagorean theorem, applied to the orthogonal sum

$$\vec{b} = \left(\vec{b} - P(\vec{b}, \vec{a})\right) + P(\vec{b}, \vec{a}),$$

so that

$$\|\vec{b}\|^2 = \| \left(\vec{b} - P(\vec{b}, \vec{a})\right) \|^2 + \|P(\vec{b}, \vec{a})\|^2 \geq \|P(\vec{b}, \vec{a})\|^2 = \|\frac{<\vec{b}, \vec{a}>}{\|\vec{a}\|^2}\vec{a}\|^2 = \frac{|<\vec{b}, \vec{a}>|}{\|\vec{a}\|^2},$$

with equality occurring if and only if $\vec{b} - P(\vec{b}, \vec{a}) = \vec{0}$, or $\vec{b} = P(\vec{b}, \vec{a})\vec{a}$, a multiple of $\vec{a}$.

It should be mentioned that traditional algebraic proofs of the Cauchy-Schwarz inequality, like the traditional introduction of dot product, suffer from an unmotivated strategy.

The characterization of angle in (*) in terms of the inner product follows from the drawing in Definition 2.1, along with either the unit circle or right triangle definition of cosine.

For more details, see "Ordinary Differential Equations, Linear Algebra, and Partial Differential Equations," Chapter III, subsection "ORTHOGONALITY," pages 21-23 and APPENDIX TWO, pages 117–121, under www.teacherscholarinstitute.com/FreeMathBooksCollege.html.

Some other pedagogy related to dot product chronologically: The equation for a line in $\mathbf{R}^n$ is naturally placed before the dot product, and immediately after defining collinear or parallel vectors. This provides more motivation for considering orthogonality, as the idea complementary to parallelness. The equation for a plane in $\mathbf{R}^3$ is naturally placed after equating orthogonality with a zero dot product and before considering orthogonal projections and angles; it's a nice illustration of the power of algebraic characterizations of orthogonality.