Points and Vectors in $\mathbb{R}^n$

(2/9/19)

Alex Nita

1 Points and Vectors

1.1 Definitions

Definition 1.1 We denote n**-dimensional Euclidean space** by

$$\begin{aligned} \mathbb{R}^n &= \overbrace{\mathbb{R} \times \cdots \times \mathbb{R}}^{n \text{ times}} \\ &= \{(x_1, \dots, x_n) \mid x_i \text{ is a real number}\} \end{aligned} \tag{1.1}$$

We will typically use capital letters P for the elements of $\mathbb{R}^n$, the n**-tuples** $(x_1, \dots, x_n)$, when thought of as **points**, in the sense of positional location,

$$P = (x_1, \dots, x_n) \tag{1.2}$$

We may also think of the elements of $\mathbb{R}^n$ as **vectors**, however, for example when working with things like velocities/forces/etc. ('vector quantities', typically residing in phase space rather than configuration space, in physics lingo). There is a certain notation which is found in both physics and math books, that emphasises the *vector* part of the elements of $\mathbb{R}^n$, and it is the *boldface* or *arrow*, and *angle-bracket* notations for the n-tuples:

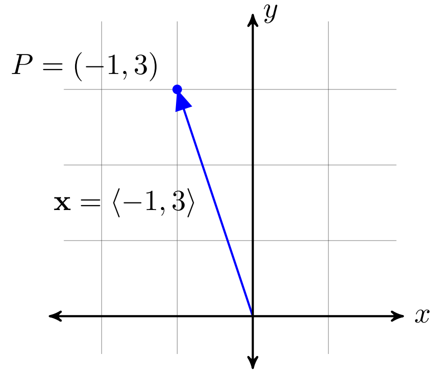
$$\mathbf{x} \quad \text{or} \quad \vec{x} = \langle x_1, \dots, x_n \rangle \tag{1.3}$$

Add to this another notation for the n-tuples, needed for our own *linear-algebraic* purposes in this course: the *column vector* notation,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{1.4}$$

Example 1.2 *Our two main examples are the Euclidean plane $\mathbb{R}^2$ and Euclidean three-dimensional space $\mathbb{R}^3$. We usually denote x_1 by x, x_2 by y, and x_3 by z, so that $P = (x, y)$ or (x, y, z), as the case may be.* ■

We visualize points and vectors differently. For example, in $\mathbb{R}^2$ the point $P = (-1, 3)$ is pictured as a dot, and the vector $\mathbf{x} = \langle -1, 3 \rangle$ is pictured as an arrow from the origin to P:



1.2 Algebraic Properties of Vectors

The essential difference between points and vectors, mathematically, is that points don't possess any algebraic properties, whereas vectors do. The key algebraic properties of vectors are *addition*, *scalar multiplication*, and the *dot* and *cross products*:

- We can **add** two vectors.
- We can **scale** any vector (multiply it by a real number).
- Rules (1) and (2) are subject to certain **rules** (**associativity, commutativity and distributivity rules**). It's not so much that the rules make vectors 'nice' to work with, in the sense of a happy coincidence. The **rules *define* the 'niceness' into the structure**. The rules determine how all this works, and in the process give vectors a certain symbolic elegance with which to decorate their geometric import.
- We can **'multiply'** two vectors in various different ways. The most well-known and useful product is the **dot product** or **scalar product** of two vectors, which results in a real number. This is a product we'll study in greater depth in this course.
- The **cross product** of two vectors results, not in a real number, but in another vector. This product, so promising at first, runs into the difficulty of not generalizing well to higher dimensions, working only in $\mathbb{R}^2$ and $\mathbb{R}^3$. It took some work, but Grassmann figured out how to get around this, by switching to the *wedge product* (see next bullet point).
- There are yet other products, such as the **tensor product**, the **wedge product** and the **Clifford product**, to name a few. They are studied in **geometric algebra** and ***multi*****linear algebra**. If we have time towards the end of the semester we'll talk about some of these, too.

We'll begin with the definition of addition.

1.2.1 Vector Addition

Definition 1.3 We define **addition** of vectors in $\mathbb{R}^n$ componentwise,

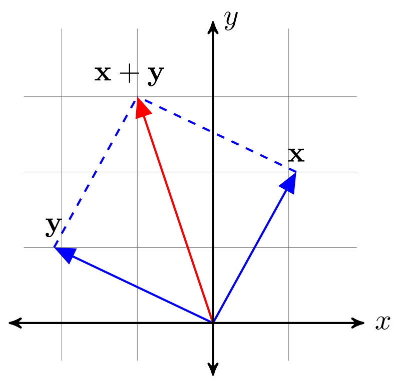
$$\begin{aligned} \mathbf{x}+\mathbf{y} &= \langle x_1, x_2, \ldots, x_n\rangle + \langle y_1, y_2, \ldots, y_n\rangle \\ &= \langle (x_1+y_1),\ (x_2+y_2), \ldots,\ (x_n+y_n)\rangle \end{aligned} \tag{1.5}$$

This is more easily seen in terms of columns:

$$\mathbf{x}+\mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{pmatrix} \tag{1.6}$$

■

Example 1.4 *Let us see what this means in* $\mathbb{R}^2$. *Take, say,* $\mathbf{x} = \langle 1, 2\rangle$ *and* $\mathbf{y} = \langle -2, 1\rangle$. *Then* $\mathbf{x}+\mathbf{y} = \langle 1-2, 2+1\rangle = \langle -1, 3\rangle$.



Thus we see that to reach $\mathbf{x}+\mathbf{y}$, *we may first go to* $\mathbf{x}$, *then go in the direction of* $\mathbf{y}$ *to get to* $\mathbf{x}+\mathbf{y}$, *or else we may to to* $\mathbf{y}$ *first and then go in the direction of* $\mathbf{x}$. *This shows geometrically the algebraic rule called* commutativity, $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$. ■

Let us now observe some *direct consequence* of our definition, which evidently is just *componentwise* application of addition in $\mathbb{R}$. All the rules for real numbers apply here, where we have n-tuples of real numbers, since what we are doing now is just carrying out n applications of the same operation *at the same time*.

Remark 1.5 Don't be phased by the obviousness of these statements, because their obviousness isn't the point: their obviousness is emphasized here for later purposes. We'll take these properties, so obvious in $\mathbb{R}^n$, and use them as the *defining characteristics* of a new *generalization*, the *abstract concept* of a *vector space* V, in which things will have the same algebraic behavior as vectors in $\mathbb{R}^n$. This will cast a wider net, and capture also *real polynomials* $\mathbb{R}[x]$, as well as many *function spaces*, for example the space of *continuous functions* $C(\mathbb{R})$, or *smooth functions* $C^\infty(\mathbb{R})$, etc., in their vector space guise. The *space of* $m \times n$ *matrices*, $M_{m,n}(\mathbb{R})$, too, will turn out to satisfy these rules, making *it* a vector space. ■

Theorem 1.6 (Properties of Vector Addition) *For all vectors* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$*, we have*

(1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ *(commutativity of addition)*

(2) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ *(associativity of addition)*

(3) *The* **zero** *vector in* $\mathbb{R}^n$ $\boxed{\mathbf{0} = \langle 0, \ldots, 0\rangle}$ *is characterized by* $\mathbf{0} + \mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{x}$ *for all* $\mathbf{x} \in \mathbb{R}^n$*.*

(4) *Every* $\mathbf{x} \in \mathbb{R}^n$ *has a* **negative** *or* **additive inverse** $\boxed{-\mathbf{x} = \langle -x_1, \ldots, -x_n\rangle}$ *which is characterized by* $(-\mathbf{x}) + \mathbf{x} = \mathbf{x} + (-\mathbf{x}) = \mathbf{0}$

Proof: We'll prove (1), and leave the others as (slightly tedious) exercises. Let $\mathbf{x} = \langle x_1, \ldots, x_n\rangle, \mathbf{y} = \langle y_1, \ldots, y_n\rangle \in \mathbb{R}^n$ and follow along:

$$\begin{aligned} \mathbf{x} + \mathbf{y} &= \langle x_1, \ldots, x_n\rangle + \langle y_1, \ldots, y_n\rangle && (1.7) \\ &= \langle x_1 + y_1, \ldots, x_n + y_n\rangle && \text{(our new definition of } +\text{)} \\ &= \langle y_1 + x_1, \ldots, y_n + x_n\rangle && \text{(commutativity is true in } \mathbb{R}\text{, componentwise)} \\ &= \mathbf{y} + \mathbf{x} \end{aligned}$$

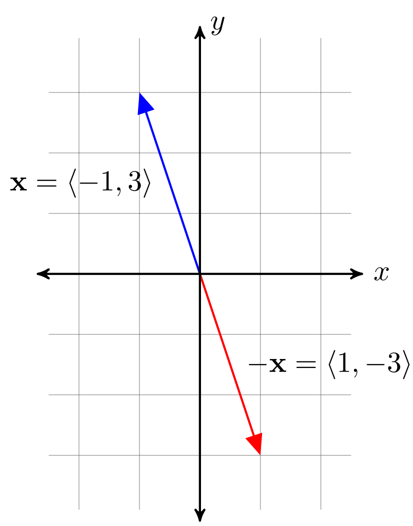
which proves (1). ■

Now follows a typical math move. *Define* subtraction to be the addition of negatives.

Definition 1.7 The **subtraction** of vectors in $\mathbb{R}^n$ is now elegantly defined to be addition of negatives:

$$\mathbf{x} - \mathbf{y} := \mathbf{x} + (-\mathbf{y})$$ ■

Geometrically, the negative $-\mathbf{x}$ of a vector is the reflection of $\mathbf{x}$ through the origin:



Remark 1.8 For those of you of a more mathematical bent, we note that, the way we've defined things, $\mathbb{R}^n$ ***is an abelian group under addition,*** $+$. Actually, if you go down the math rabbit hole far enough you begin to combine algebra with calculus/analysis and topology/geometry in order to create all manner of conceptual chimeras. But the biggest one of these is the general ***Lie group***, which is a group with extra topological and smooth structure. You need look no further than $(\mathbb{R}^n, +)$ to see one, but the group of all invertible matrices, the ***general linear group*** $\mathrm{GL}(n, \mathbb{R})$, provides us with another, and in fact with a whole slew of them: all of its subgroups. Hopefully we'll get to some of this highly interesting material towards the end of the course! ■

1.2.2 Scalar Multiplication

Definition 1.9 We define **scalar multiplication** of vectors in $\mathbb{R}^n$, by which we mean multiplication of a vector $\mathbf{x} \in \mathbb{R}^n$ by a real number $a \in \mathbb{R}$, componentwise:

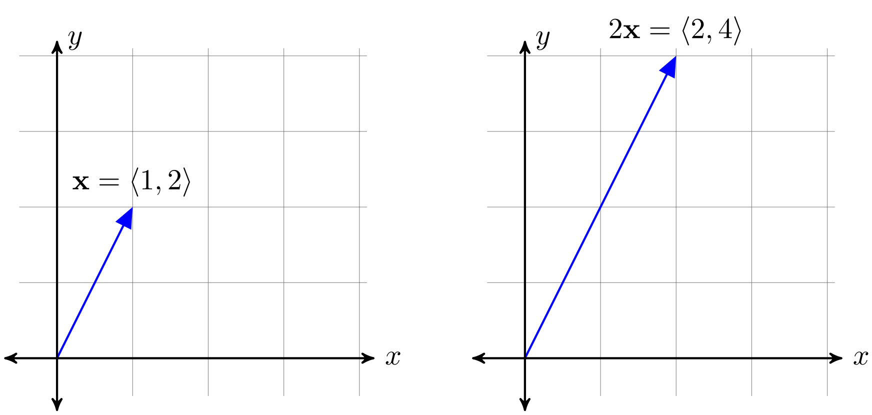
$$a\mathbf{x} = a\langle x_1, \ldots, x_n \rangle := \langle ax_1, \ldots, ax_n \rangle \quad \text{or} \quad a\mathbf{x} = a\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$$ ■

The geometric content of scalar multiplication may be seen in the following example:

Example 1.10 *Take, say, the vector* $\mathbf{x} = \langle 1, 2 \rangle$ *and the real number* $a = 2$. *Then, algebraically,*

$$a\mathbf{x} = 2\langle 1, 2 \rangle = \langle 2 \cdot 1, \ 2 \cdot 2 \rangle = \langle 2, 4 \rangle$$

which, geometrically means this:



Thus geometrically scalar multiplication has the effect of scaling the length *of the vector* $\mathbf{x}$. ■

Immediate properties of scalar multiplication are the following obvious ones, which I leave to you as an exercise to prove.

Theorem 1.11 (Properties of Scalar Multiplication) *For all* $a, b \in \mathbb{R}$ *and all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, *we have*

(1) $a(b\mathbf{x}) = (ab)\mathbf{x}$ *(associativity of scalar mult.)*

(2) $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$ *(distributivity over* $\mathbb{R}$*-addition)*

(3) $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$ *(distributivity over* $\mathbb{R}^n$*-addition)*

(4) $1\mathbf{x} = \mathbf{x}$. ■

For reasons that must remain obscure for now (having to do with the notion of *group actions*), we include condition (4) in the immediate foreground (in the theorem above), and keep the next condition separate, not so much because it is a *logical consequence* of the previous theorem, since indeed we can prove it directly from the definition, but because *we must keep an eye on our future goal: the abstract vector space. In that case, we won't have* $0\mathbf{x}$ *defined directly, but will derive its meaning from the above conditions, which will be taken for the definition of scalar multiplication.*

Corollary 1.12 *For all* $\mathbf{x} \in \mathbb{R}^n$, *we have* $0\mathbf{x} = \mathbf{0}$.

Proof: Suppose $\mathbf{x} \in \mathbb{R}^n$. If we scale by $0 \in \mathbb{R}$, we get, straight from our definition,

$$0\mathbf{x} = 0\langle x_1, \ldots, x_n\rangle = \langle 0x_1, \ldots, 0x_n\rangle = \langle 0, \ldots, 0\rangle = \mathbf{0}$$

since $0x_i = 0$ for all i. ■

Remark 1.13 Note, for future purposes, that we could have proved this differently, using only the conditions of the previous theorems: Since the zero vector satisfies $\mathbf{0} + \mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$, and since scalar multiplication is distributive, we get

$$\mathbf{0} + 0\mathbf{x} = 0\mathbf{x} = (0 + 0)\mathbf{x} = 0\mathbf{x} + 0\mathbf{x}$$

so that, subtracting $0\mathbf{x}$ from both sides gives $\mathbf{0} = 0\mathbf{x}$. This makes no use of *the particular way* we defined scalar multiplication, it just makes use of its *abstract properties* (and those of addition). ■

1.2.3 Standard Coordinate Vectors

We can use addition and scalar multiplicaton of vectors to *decompose* a given vector $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$ into its **components** x_i, as follows:

$$\begin{aligned} \mathbf{x} &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ x_n \end{pmatrix} \\ &= x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \end{aligned} \tag{1.8}$$

This calls for a definition:

Definition 1.14 Define the **standard (Cartesian coordinate) basis vectors**

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

That is, $\mathbf{e}_i = \langle 0, \ldots, 1, \ldots, 0 \rangle$ with the 1 in the ith slot and 0's everywhere else. ■

With this in hand, we can nicely rewrite equation (1.8) as

$$\mathbf{x} = \langle x_1, \ldots, x_n \rangle = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \cdots + x_n \mathbf{e}_n \tag{1.9}$$

Example 1.15 *When $n = 2$, we have special notation, which is more commonly found in physics texts:*

$$\begin{aligned} \mathbf{e}_1 &= \mathbf{i} = \langle 1, 0 \rangle \\ \mathbf{e}_2 &= \mathbf{j} = \langle 0, 1 \rangle \end{aligned} \tag{1.10}$$

When $n = 3$, we also write:

$$\begin{aligned} \mathbf{e}_1 &= \mathbf{i} = \langle 1, 0, 0 \rangle \\ \mathbf{e}_2 &= \mathbf{j} = \langle 0, 1, 0 \rangle \\ \mathbf{e}_3 &= \mathbf{k} = \langle 0, 0, 1 \rangle \end{aligned} \tag{1.11}$$

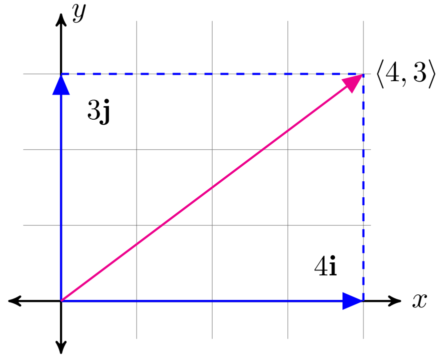
For example,

$$\langle 1, 3, -2 \rangle = 1\mathbf{i} + 3\mathbf{j} - 2\mathbf{k} \tag{1.12}$$

is decomposed into its components. As another example,

$$\langle 4, 3 \rangle = 4\mathbf{i} + 3\mathbf{j} \tag{1.13}$$

and this can be pictured as follows:



■

1.3 The Dot and Cross Products

1.3.1 The Dot Product

Definition 1.16 The **dot product** of two vectors $\mathbf{x} = \langle x_1, x_2, \ldots, x_n \rangle$ and $\mathbf{y} = \langle y_1, y_2, \ldots, y_n \rangle$ in $\mathbb{R}^n$ is defined by

$$\boxed{\mathbf{x} \cdot \mathbf{y} \;=\; \langle x_1, x_2, \ldots, x_n \rangle \cdot \langle y_1, y_2, \ldots, y_n \rangle \;=\; x_1 y_1 + x_2 y_2 + \cdots + x_n y_n} \tag{1.14}$$

or, more concisely, using summation notation,

$$\boxed{\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i}$$

■

Remark 1.17 We may dot in various ways, in practice, such as dotting the row of a matrix A with a column of a matrix B. Thus, we allow

$$\mathbf{x} \cdot \mathbf{y} \;=\; \langle x_1, x_2, \ldots, x_n \rangle \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i$$

■

Example 1.18 *Consider the vectors* $\mathbf{x} = \langle 1, 2, -5 \rangle$ *and* $\mathbf{y} = \langle -3, 2, 4 \rangle$ *in* $\mathbb{R}^3$. *Their dot product is*

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= \langle 1, 2, -5 \rangle \cdot \langle -3, 2, 4 \rangle \\ &= 1 \cdot (-3) + 2 \cdot 2 + (-5) \cdot 4 \\ &= -19 \end{aligned}$$

■

Definition 1.19 The **length** (or **magnitude** or **norm**) of a vector $\mathbf{x}$ in $\mathbb{R}^n$ will be defined as the square root of the dot product of $\mathbf{x}$ with itself:

$$\boxed{\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}} \tag{1.15}$$

Thus,

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

and therefore

$$\|\mathbf{x}\|^2 = \sum_{i=1}^{n} x_i^2 = \mathbf{x} \cdot \mathbf{x}$$ ■

Example 1.20 *The magnitude of the vector* $\mathbf{x} = \langle 4, 1 \rangle \in \mathbb{R}^2$ *is*

$$\|\langle 4, 1 \rangle\| = \sqrt{4^2 + 1^2} = \sqrt{17}$$ ■

Proposition 1.21 (Algebraic Properties of the Dot Product) *Let* $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ *be vectors in* $\mathbb{R}^n$ *and let* c *be a real number. Then,*

(1) $\mathbf{x} \cdot \mathbf{x} = |\mathbf{x}|^2$

(2) $|c\mathbf{x}| = |c||\mathbf{x}|$

(3) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$ *(commutativity)*

(4) $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$ *(distributivity over addition)*

(5) $(c\mathbf{x}) \cdot \mathbf{y} = c(\mathbf{x} \cdot \mathbf{y}) = \mathbf{x} \cdot (c\mathbf{y})$ *(associativity and commutativity of scalar multiplication and dot multiplication)*

(6) $\mathbf{0} \cdot \mathbf{x} = 0$

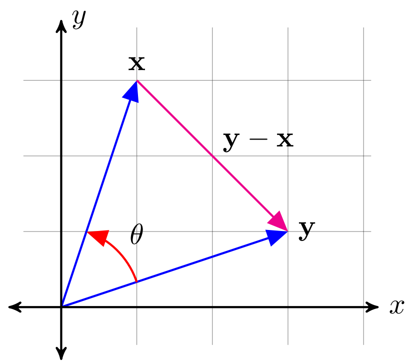
Proof: A worthy exercise. ■

In two and three dimensions, the dot product has a very *geometric* interpretation:

Proposition 1.22 *Let* $\mathbf{x} = \langle x_1, y_1, z_1 \rangle$ *and let* $\mathbf{y} = \langle x_2, y_2, z_2 \rangle$ *be two vectors in* $\mathbb{R}^3$*, and let* θ *be the angle between them. Then,*

$$\boxed{\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos\theta} \tag{1.16}$$

Proof: Consider the triangle formed by the vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{y} - \mathbf{x}$.



By the Law of Cosines

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$$

Using the fact that $|\mathbf{x}|^2 = \mathbf{x} \cdot \mathbf{x}$, we can rewrite the above as

$$(\mathbf{y} - \mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}) = \mathbf{x} \cdot \mathbf{x} + \mathbf{y} \cdot \mathbf{y} - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$$

Distributing on the left and simplifying, we get

$$\cancel{\mathbf{y} \cdot \mathbf{y}} - 2\mathbf{x} \cdot \mathbf{y} + \cancel{\mathbf{x} \cdot \mathbf{x}} = \cancel{\mathbf{x} \cdot \mathbf{x}} + \cancel{\mathbf{y} \cdot \mathbf{y}} - 2\|\mathbf{x}\|\|\mathbf{y}\|$$

That is,

$$\cancel{-2}\mathbf{x} \cdot \mathbf{y} = \cancel{-2}\|\mathbf{x}\|\|\mathbf{y}\| \qquad \blacksquare$$

Remark 1.23 We observe that if we were clever, we could generalize this to $\mathbb{R}^n$ by simply *defining* the angle θ between $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ to be

$$\theta := \cos^{-1}\left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\right)$$

which is, of course, what the clever people have already done. $\blacksquare$

1.3.2 Cross Product

Note that this section requires knowledge of determinants, which is to be found below, in the section on matrices.

The cross product is a *vector* product, meaning that multiplying two vectors this way results in another vector (this is why the dot product is sometimes called the *scalar* product, to distinguish it from this vector product). The cross product is only defined in 3 dimensions, i.e. only on $\mathbb{R}^3$.[1]

Let $\mathbf{u} = \langle a, b, c\rangle$ and $\mathbf{v} = \langle d, e, f\rangle$ be two vectors in $\mathbb{R}^3$. We define their **cross product** to be the vector gotten by computing the following determinant:

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= \langle a, b, c\rangle \times \langle d, e, f\rangle = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a & b & c \\ d & e & f \end{vmatrix} \\ &= \mathbf{i}\begin{vmatrix} b & c \\ e & f \end{vmatrix} - \mathbf{j}\begin{vmatrix} a & c \\ d & f \end{vmatrix} + \mathbf{k}\begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{aligned} \tag{1.17}$$

Carrying out this computation to its awful end, we get

$$\mathbf{u} \times \mathbf{v} = (bf - ce)\mathbf{i} - (af - cd)\mathbf{j} + (ae - bd)\mathbf{k} = \langle bf - ce,\ cd - af,\ ae - bd\rangle$$

But it is easier to remember the equation (1.17) in terms of determinants and just perform the rest of the computation by hand in particular cases.

Example 1.24 *Let* $\mathbf{u} = \langle 1, 2, -2\rangle$ *and* $\mathbf{v} = \langle -8, 5, 4\rangle$*. Then,*

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= \langle 1, 2, -2\rangle \times \langle -8, 5, 4\rangle \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & -2 \\ -8 & 5 & 4 \end{vmatrix} \\ &= \mathbf{i}\begin{vmatrix} 2 & -2 \\ 5 & 4 \end{vmatrix} - \mathbf{k}\begin{vmatrix} 1 & -2 \\ -8 & 4 \end{vmatrix} + \mathbf{j}\begin{vmatrix} 1 & 2 \\ -8 & 5 \end{vmatrix} \\ &= (8 + 10)\mathbf{i} - (4 - 16)\mathbf{j} + (5 + 16)\mathbf{k} \\ &= \boxed{18\mathbf{i} + 12\mathbf{j} + 21\mathbf{k}} \quad \textit{or} \quad \boxed{\langle 18, 12, 21\rangle} \end{aligned}$$

Example 1.25 *Let us compute* $\mathbf{i} \times \mathbf{j}$*:*

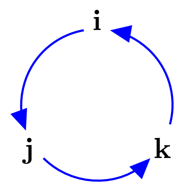
$$\mathbf{i} \times \mathbf{j} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \mathbf{i}\begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} - \mathbf{j}\begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 0\mathbf{i} - 0\mathbf{j} + 1\mathbf{k} = \mathbf{k}$$

[1] It generalizes to other dimensions only once we switch to the **wedge product** (or **exterior product**) in multilinear algebra. Moving past the algebra we are led to a *differential* type of wedge product in the apparatus of **differential forms**. For further reading on this, see Knapp [2] and Gallier [1] for the algebra, and Munkres [3] for the calculus side.

By similar calculations, which we leave to you, we also have the relations $\mathbf{j} \times \mathbf{k} = \mathbf{i}$ *and* $\mathbf{k} \times \mathbf{i} = \mathbf{j}$*:*

$$\boxed{\begin{aligned} \mathbf{i} \times \mathbf{j} &= \mathbf{k} \\ \mathbf{k} \times \mathbf{i} &= \mathbf{j} \\ \mathbf{j} \times \mathbf{k} &= \mathbf{i} \end{aligned}} \tag{1.18}$$

Note that the vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ *are cyclically permuted:*



(1.19)

in slots of the equation $\_ \times \_ = \_$

■

Proposition 1.26 (Algebraic Properties of the Cross Product) *Let* $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ *be vectors in* $\mathbb{R}^3$ *and let* c *be a real number. Then,*

(1) $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$ *(anti-commutativity)*

(2) $\mathbf{x} \times (\mathbf{y} + \mathbf{z}) = \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z}$, $(\mathbf{x} + \mathbf{y}) \times \mathbf{z} = \mathbf{x} \times \mathbf{z} + \mathbf{y} \times \mathbf{z}$ *(distributivity over addition)*

(3) $(c\mathbf{x}) \times \mathbf{y} = c(\mathbf{x} \times \mathbf{y}) = \mathbf{x} \times (c\mathbf{y})$ *(associativity of scalar multiplication and cross multiplication)*

Proof: Let us prove (1), and leave the rest as easy exercises. Let $\mathbf{x} = \langle a, b, c\rangle$ and $\mathbf{y} = \langle d, e, f\rangle$. Then,

$$\mathbf{y} \times \mathbf{x} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ d & e & f \\ a & b & c \end{vmatrix} = \mathbf{i} \begin{vmatrix} e & f \\ b & c \end{vmatrix} - \mathbf{j} \begin{vmatrix} d & f \\ a & c \end{vmatrix} + \begin{vmatrix} d & e \\ a & b \end{vmatrix} = (ce - bf)\mathbf{i} - (cd - af)\mathbf{j} + (bd - ae)\mathbf{k}$$

However, according to the calculation (1.17) above,

$$\mathbf{x} \times \mathbf{y} = (bf - ce)\mathbf{i} - (af - cd)\mathbf{j} + (ae - bd)\mathbf{k} = -\mathbf{y} \times \mathbf{x}$$ ■

Proposition 1.27 *For any vectors* $\mathbf{u}$, $\mathbf{v}$ *in* $\mathbb{R}^3$ *we have the identity*

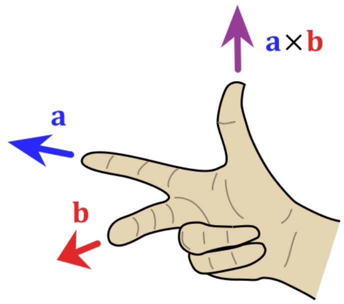
$$\boxed{\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\| \sin\theta} \tag{1.20}$$

and consequently

$$\boxed{\mathbf{u} \times \mathbf{v} = \big(\|\mathbf{u}\|\|\mathbf{v}\| \sin\theta\big)\mathbf{n}} \tag{1.21}$$

where θ *is the angle between the vectors and* $\mathbf{n} = \frac{\mathbf{u} \times \mathbf{v}}{\|\mathbf{u} \times \mathbf{v}\|}$ *is the unit vector in the direction of* $\mathbf{u} \times \mathbf{v}$*, determined by the* **right-hand rule***, illustrated in the following diagram,*[2]

[2]Which I got from the Wikipedia page on the cross product, http://en.wikipedia.org/wiki/Cross_product.



Proof: Let $\mathbf{u} = \langle a, b, c\rangle$ and $\mathbf{v} = \langle d, e, f\rangle$. By the calculation (1.17) we have that $\mathbf{u} \times \mathbf{v} = \langle bf - ce,\ cd - af,\ ae - bd\rangle$, and consequently

$$\begin{aligned}
\|\mathbf{u}\|^2\|\mathbf{v}\|^2 \sin^2\theta &= \|\mathbf{u}\|^2\|\mathbf{v}\|^2(1-\cos^2\theta) \\
&= \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - \|\mathbf{u}\|^2\|\mathbf{v}\|^2\cos^2\theta \\
&= \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u}\cdot\mathbf{v})^2 \\
&= (a^2+b^2+c^2)(d^2+e^2+f^2) - (ad+be+cf)^2 \\
&= (a^2d^2+b^2e^2+c^2f^2+b^2f^2+c^2e^2+c^2d^2+a^2f^2+a^2e^2+b^2d^2) \\
&\qquad -(a^2d^2+b^2e^2+c^2f^2+2bcef+2acdf+2abde) \\
&= \left(b^2f^2-2bcef+c^2e^2\right)+\left(c^2d^2-2acdf+a^2f^2\right)+\left(a^2e^2-2abde+b^2d^2\right) \\
&\qquad +(a^2d^2+b^2e^2+c^2f^2)-(a^2d^2+b^2e^2+c^2f^2) \\
&= (bf-ce)^2+(cd-af)^2+(ae-bd)^2 \\
&= \langle bf-ce,\ cd-af,\ ae-bd\rangle\cdot\langle bf-ce,\ cd-af,\ ae-bd\rangle \\
&= \|\mathbf{u}\times\mathbf{v}\|^2
\end{aligned}$$

Taking the square root gives the result. ■

1.3.3 Interaction Between the Dot and Cross Products

Proposition 1.28 *Let* $\mathbf{a} = \langle a_1, a_2, a_3 \rangle,\ \mathbf{b} = \langle b_1, b_2, b_3 \rangle,\ \mathbf{c} = \langle c_1, c_2, c_3 \rangle$ *be vectors in* $\mathbb{R}^3$*. Then,*

(1) $\begin{array}{l} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \\ (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{c} \cdot \mathbf{b})\mathbf{a} \end{array}$ *(vector triple product)*

(2) $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) =$
$\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})$ *(scalar triple product)*

(3) $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} + (\mathbf{c} \times \mathbf{a}) \times \mathbf{b} + (\mathbf{b} \times \mathbf{c}) \times \mathbf{a} = \mathbf{0}$ *(Jacobi identity)*

Proof: (1) This is a direct computation:

$$\begin{aligned}
\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} & -\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} & \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \end{vmatrix} \\
&= \Big\langle a_2 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix}, \\
&\qquad a_3 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_1 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}, \\
&\qquad\quad -a_1 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} \Big\rangle \\
&= \big\langle a_2(b_1c_2 - b_2c_1) + a_3(b_1c_3 - b_3c_1), \\
&\qquad a_3(b_2c_3 - b_3c_2) - a_1(b_1c_2 - b_2c_1), \\
&\qquad\quad -a_1(b_1c_3 - b_3c_1) - a_2(b_2c_3 - b_3c_2)\big\rangle \\
&= \big\langle a_2b_1c_2 - a_2b_2c_1 + a_3b_1c_3 - a_3b_3c_1, \\
&\qquad a_3b_2c_3 - a_3b_3c_2 - a_1b_1c_2 + a_1b_2c_1, \\
&\qquad\quad -a_1b_1c_3 + a_1b_3c_1 - a_2b_2c_3 + a_2b_3c_2\big\rangle \\
&\quad + \langle a_1b_1c_1,\ a_2b_2c_2,\ a_3b_3c_3 \rangle - \langle a_1b_1c_1,\ a_2b_2c_2,\ a_3b_3c_3 \rangle \\
&= \big\langle (a_2c_2 + a_3c_3)b_1 - (a_2b_2 + a_3b_3)c_1, \\
&\qquad (a_1c_1 + a_3c_3)b_2 - (a_1b_1 + a_3b_3)c_2, \\
&\qquad\quad (a_1c_1 + a_2c_2)b_3 - (a_1b_1 + a_2b_2)c_3\big\rangle \\
&\quad + \langle a_1b_1c_1,\ a_2b_2c_2,\ a_3b_3c_3 \rangle - \langle a_1b_1c_1,\ a_2b_2c_2,\ a_3b_3c_3 \rangle \\
&= \Big\langle (a_1c_1 + a_2c_2 + a_3c_3)b_1 - (a_1b_1 + a_2b_2 + a_3b_3)c_1, \\
&\qquad (a_1c_1 + a_2c_2 + a_3c_3)b_2 - (a_1b_1 + a_2b_2 + a_3b_3)c_2, \\
&\qquad\quad (a_1c_1 + a_2c_2 + a_3c_3)b_3 - (a_1b_1 + a_2b_2 + a_3b_3)c_3 \Big\rangle
\end{aligned}$$

$$= \big\langle (\mathbf{a}\cdot\mathbf{c})b_1 - (\mathbf{a}\cdot\mathbf{b})c_1,\quad (\mathbf{a}\cdot\mathbf{c})b_2 - (\mathbf{a}\cdot\mathbf{b})c_2,\quad (\mathbf{a}\cdot\mathbf{c})b_3 - (\mathbf{a}\cdot\mathbf{b})c_3 \big\rangle$$

$$= (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$$

The other expression in (1) follows from this one and the anti-commutativity of the cross product: $-(\mathbf{a}\times\mathbf{b})\times\mathbf{c} = \mathbf{c}\times(\mathbf{a}\times\mathbf{b}) = (\mathbf{c}\cdot\mathbf{b})\mathbf{a} - (\mathbf{c}\cdot\mathbf{a})\mathbf{b}$ so that $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{c}\cdot\mathbf{b})\mathbf{a}$.

For (2), we simply compute, using some basic facts about determinants:

$$\begin{aligned}
(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c} &= \left\langle \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, -\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \right\rangle \cdot \langle c_1, c_2, c_3 \rangle \\
&= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} c_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} c_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} c_3 \\
&= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \\
&= \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (\mathbf{c}\times\mathbf{a})\cdot\mathbf{b} = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) \\
&= \begin{vmatrix} b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \end{vmatrix} = (\mathbf{b}\times\mathbf{c})\cdot\mathbf{a} = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) \\
&= \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})
\end{aligned}$$

since each of the other two determinants is obtained from the first by 2 row interchanges, which are equal to $(-1)^2$ times the first.

The Jacobi identity (3) follows from the vector triple product: $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} + (\mathbf{c}\times\mathbf{a})\times\mathbf{b} + (\mathbf{b}\times\mathbf{c})\times\mathbf{a} = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c} + (\mathbf{c}\cdot\mathbf{b})\mathbf{a} - (\mathbf{c}\cdot\mathbf{a})\mathbf{b} + (\mathbf{b}\cdot\mathbf{a})\mathbf{c} - (\mathbf{b}\cdot\mathbf{c})\mathbf{a} = \mathbf{0}$. ■

1.3.4 Geometric Properties of the Dot and Cross Products

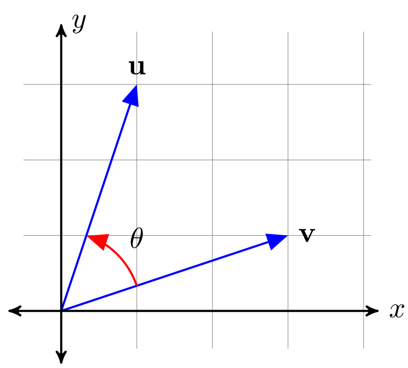
The formulas

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta \tag{1.22}$$
$$\|\mathbf{u}\times\mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\|\sin\theta \tag{1.23}$$

for all vectors **u** and **v** in $\mathbb{R}^3$ forming an acute angle θ, already contain significant geometric information. The purely algebraic definition of $\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3$, which a priori doesn't say anything about angles and orthogonality, turns out to in fact give precisely that information. Indeed, we can use purely algebraic information about **u** and **v**, namely their magnitudes and dot product, to gain the important geometric information about the acute angle they form:

$$\theta = \cos^{-1}\left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\right) \tag{1.24}$$

(provided, of course, that neither vector is the zero vector $\mathbf{0} = \langle 0,0,0\rangle$, else we'd be dividing by 0).



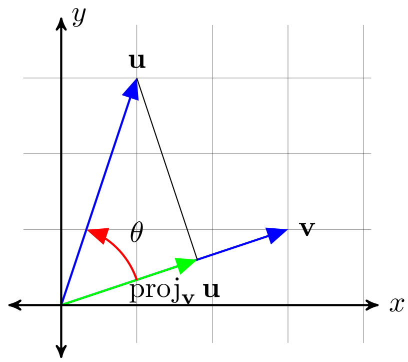
This fact motivates the following definition. We say two vectors **u** and **v** are **orthogonal** or **perpendicular** if $\mathbf{u}\cdot\mathbf{v} = 0$, and we denote this by

$$\boxed{\mathbf{u}\perp\mathbf{v}} \tag{1.25}$$

From equation (1.22) we immediately get that

$$\mathbf{u}\perp\mathbf{v} \iff \cos\theta = 0 \iff \theta = \frac{\pi}{2} \tag{1.26}$$

Next, consider the projection of the vector **u** onto the vector **v**, that is, drop a perpendicular from the arrowhead of $\mathbf{u}$ onto the line containing $\mathbf{v}$,



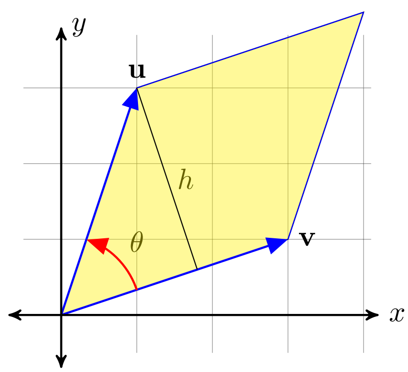
Observe that the length of the projection is clearly

$$\text{comp}_{\mathbf{v}}\,\mathbf{u} = \|\operatorname{proj}_{\mathbf{v}}\mathbf{u}\| = \|\mathbf{u}\|\cos\theta = \|\mathbf{u}\|\frac{\|\mathbf{v}\|}{\|\mathbf{v}\|}\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|} \tag{1.27}$$

Therefore, if we give it a direction, namely the unit direction $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ of $\mathbf{v}$, we get

$$\operatorname{proj}_{\mathbf{v}}\mathbf{u} = (\|\mathbf{u}\|\cos\theta)\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{v}\|^2}\mathbf{v} \tag{1.28}$$

Next, consider the parallelogram formed by the vectors $\mathbf{u}$ and $\mathbf{v}$:



We know that its area is the length of its base times its height,

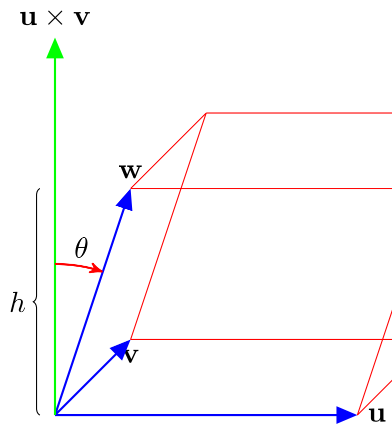
$$A = bh$$

Now, $b = |\mathbf{v}|$ and $h = |\mathbf{u}|\sin\theta$, so by equation

$$A = bh = \|\mathbf{v}\|\|\mathbf{u}\|\sin\theta = \|\mathbf{u}\times\mathbf{v}\| \tag{1.29}$$

i.e. the area of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$ is the length of their cross product!

Now consider a **parallelepiped** spanned by three vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, the 3-dimensional analog of the parallelogram, a rectangular solid whose opposite sides are all parallel.



The volume is the area of the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$ times the height h,

$$V = Ah = \|\mathbf{u} \times \mathbf{v}\| h$$

But note that h is the length of the projection of $\mathbf{w}$ onto $\mathbf{u} \times \mathbf{v}$,

$$h = \| \operatorname{comp}_{\mathbf{u}\times\mathbf{v}} \mathbf{w}\| = \frac{|\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})|}{\|\mathbf{u} \times \mathbf{v}\|} = \left| \mathbf{w} \cdot \left(\frac{\mathbf{u} \times \mathbf{v}}{\|\mathbf{u} \times \mathbf{v}\|} \right) \right| = \|\mathbf{w}\| |\cos\theta|$$

Hence,

$$V = Ah = \|\mathbf{u} \times \mathbf{v}\| \frac{|\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})|}{\|\mathbf{u} \times \mathbf{v}\|} = |(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}| \tag{1.30}$$

I.e., *the volume V is equal to the absolute value of the scalar triple product of* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{w}$, which, by our formulas for the scalar triple product from Proposition 1.28, can be computed using the determinant: if $\mathbf{u} = \langle u_1, u_2, u_3 \rangle$, $\mathbf{v} = \langle v_1, v_2, v_3 \rangle$ and $\mathbf{w} = \langle w_1, w_2, w_3 \rangle$, then

$$V = |(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}| = \text{absolute value of } \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} \tag{1.31}$$

2 Appendix: The Relationship Between Points and Vectors—the Displacement Vector

Let $O = (0, \ldots, 0) \in \mathbb{R}^n$ be the origin, thought of as a point. Given any two points $P = (x_1, \ldots, x_n)$ and $Q = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, define associated *vectors*, which we denote

$$\overrightarrow{OP} := \langle x_1, \ldots, x_n \rangle, \qquad \overrightarrow{OQ} := \langle y_1, \ldots, y_n \rangle$$

Then, the **displacement vector** from P to Q is defined as

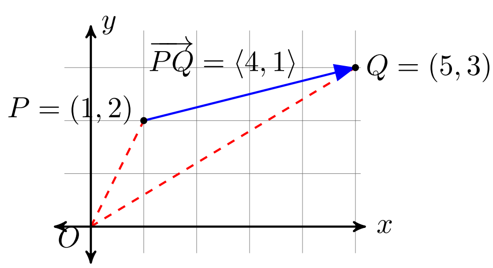
$$\begin{aligned}\overrightarrow{PQ} &= \overrightarrow{OQ} - \overrightarrow{OP} \\ &= \langle y_1, \ldots, y_n \rangle - \langle x_1, \ldots, x_n \rangle \\ &= \langle y_1 - x_1, \ldots, y_n - x_n \rangle\end{aligned} \tag{2.1}$$

We *picture* the displacement vector $\overrightarrow{PQ}$ as emanating from the point P and ending in an arrow at the point Q (even though it always, strictly speaking, emanates from the origin to its endpoint).

Example 2.1 *Let us look at the case of* $\mathbb{R}^2$. *Suppose* $P = (1, 2)$ *and* $Q = (5, 3)$. *Then,*

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \langle 5, 3 \rangle - \langle 1, 2 \rangle = \langle 5 - 1,\ 3 - 2 \rangle = \langle 4, 1 \rangle \tag{2.2}$$

and the picture is this:



Remark 2.2 In fact, $\overrightarrow{PQ}$ should be pictured as emanating from the origin, but we *want* to think of $\overrightarrow{PQ}$ as emanating from P. We should, then, if we were being rigorous, think of $\overrightarrow{PQ}$ as lying in a copy of $\mathbb{R}^n$ sitting above our position space $\mathbb{R}^n$ at the point P, that is we should think of $\overrightarrow{PQ}$ as lying in the set $\{P\} \times \mathbb{R}^n = \{(P, \mathbf{x}) | \mathbf{x} = \langle x_1, \ldots, x_n \rangle\}$, and thus

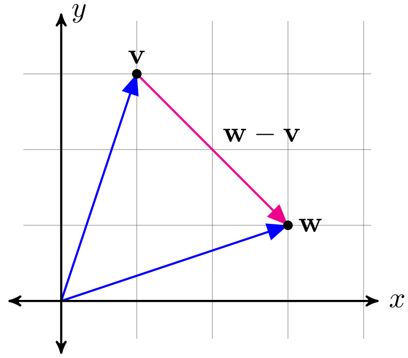
$$\overrightarrow{PQ} = (P, \overrightarrow{OQ} - \overrightarrow{OP})$$

For example, if $P = (1, 2)$ and $Q = (5, 3)$, then

$$\overrightarrow{PQ} = \Big((1, 2),\ \langle 4, 1 \rangle\Big)$$

We will not nit-pick here, and we will simply conflate points and vectors in the strict sense, but we will *picture* vectors as emanating from points in the underlying position space. ■

Now suppose we are considering not two points P and Q, but two vectors $\mathbf{v}$ and $\mathbf{w}$. Then we can consider the **displacement vector** from $\mathbf{v}$ to $\mathbf{w}$. This is, in fact $\mathbf{w}-\mathbf{v}$, which is simply due to the fact that $\mathbf{w} = \mathbf{v} + (\mathbf{w} - \mathbf{v})$:



This of course harmonizes with our previous definition, $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$. We subtract our starting position vector from our ending position vector in both cases.

The formula for the **distance between two points** $P = (x_1, \ldots, x_n)$ and $Q = (y_1, \ldots, y_n)$ is the formula gotten from a generalized Pythagorean theorem,

$$d(P,Q) \;=\; d\Big((x_1,\ldots,x_n),\ (y_1,\ldots,y_n)\Big) \;=\; \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2} \tag{2.3}$$

In the cases $n = 2$ and $n = 3$, this is literally the Pythagorean theorem. For $n > 3$ it's simply a definition which in a way takes the Pythagorean theorem as an axiom.

What is the relationship between the distance formula and the length of a vector? It is that the length of the displacement vector $\overrightarrow{PQ}$ is precisely the distance between P and Q:

$$\begin{aligned}
\|\overrightarrow{PQ}\| &= \sqrt{\overrightarrow{PQ}\cdot\overrightarrow{PQ}} \\
&= \sqrt{(\overrightarrow{Q}-\overrightarrow{P})\cdot(\overrightarrow{Q}-\overrightarrow{P})} \\
&= \sqrt{\langle y_1 - x_1, \ldots, y_n - x_n\rangle \cdot \langle y_1 - x_1, \ldots, y_n - x_n\rangle} \\
&= \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2} \\
&= d(P,Q)
\end{aligned}$$

References

[1] Gallier, J. *Notes on Differential Geometry*, online notes at http://www.cis.upenn.edu/~jean/gbooks/manif.html, Version of August 2013.

[2] Knapp, A.W., *Basic Algebra*, Birkh auser, 2000.

[3] Munkres, J.R., *Analysis On Manifolds*, Westview Press, 1997.