VECTOR ALGEBRA

A. HAVENS

Vectors in the context of physics, engineering, and other sciences are mathematical objects representing phenomena in space that possess two attributes: *direction*, and *magnitude*.

They are the natural geometric objects for describing instantaneous linear motion or action, such as displacements, velocities, forces, and, once given an origin (a distinguished point in space), position. Thus, we will encounter *position vectors*, *displacement vectors*, *velocities*, *forces* and more, all collectively obeying algebraic rules that capture their geometric behaviors. Employing coordinates, we can give vectors a concrete life with which computations can be performed; we will nevertheless prefer geometric and coordinate-free definitions first in these notes, and derive the relevant coordinate rules from our geometric intuition whenever possible. Occasionally this kind of thinking will be relegated to the exercises and problems at the end of each section.

There is an optional section at the end which explores the mathematics behind certain geometric transformations of two and three dimensional space, and connects these to the history of vectors, via Hamilton's theory of *quaternions*.

Contents

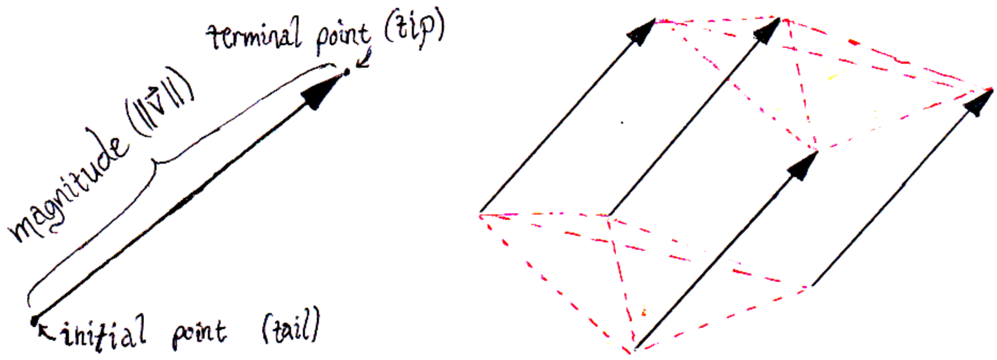
1. Vectors and Coordinates

We will give an informal definition of geometric vectors in dimensions ≤ 3, and begin to introduce the notations common for working with them, in particular in three dimensions. Thus, this section presumes familiarity with 3-dimensional rectangular coordinates. Later in the course we will see other coordinate systems (such as polar coordinates), which we will define as needed. We will not formally define vector spaces, deferring such treatment to a linear algebra course.

We'll label vectors by bolded letters, as distinguished from real (or complex numbers), which are written in italics. As our definitions of vectors as well as the algebra we will do with them are presently grounded in real numbers, it is important to be fluent in real arithmetic and algebra. The set of all reals will be denoted $\mathbb{R}$, while the real plane will be denoted $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$, and real 3-dimensional space will be denoted by $\mathbb{R}^3 := \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. We will refer to the reals as *scalars*, in keeping with standard language to distinguish the base field from the spaces of vectors built in real coordinates.

§ 1.1. Defining vectors geometrically

In practice, we wish to model vectors as arrows with a definite direction and length, sitting in our favorite (real, Euclidean) space, such as the plane $\mathbb{R}^2$ or 3-dimensional space $\mathbb{R}^3$. By direction, we mean two things: the arrow determines a line segment, and this line segment has an orientation. When we draw this, the arrow's head will tell us the orientation. More formally, we can describe a vector's anatomy as follows: there is an *initial point* in space, sometimes referred to as the *tail*, and a *terminal point*, often called the *tip*. The direction is then determined by the *oriented* line segment going from this initial point to the terminal point. See figure (1a). The initial and terminal points may coincide, in which case we have a distinguished case, to be discussed momentarily when we begin to define vector arithmetic.



(a) (b)

FIGURE 1. (a) – Anatomy of a vector. (b) – Various equivalent vectors.

We'll call two vectors parallel if their corresponding line segments are parallel, even if the orientations do not match. But for two vectors to have the same direction, we require that the line segments be parallel, and that the orientations match. Equivalently, two vectors have the same direction if and only if there is a continuous motion through similarity transformations taking one arrow to the other through parallel arrows, i.e. via translation and dilation, but without rotation or reflection.

We'll assume the usual means of measurement in space, and thus there is a length, given by a non-negative real number, associated to the line segments. The *magnitude* of a vector is this length, and for a vector $\mathbf{v}$ the magnitude is denoted $\|\mathbf{v}\|$, or sometimes simply $|\mathbf{v}|$. Some texts use the convention that the italic letter denotes the magnitude of the corresponding boldfaced vector, e.g. 'v' denotes $\|\mathbf{v}\|$, however we will prefer clarity and write the double-bar magnitude notation with few exceptions[1].

We will often say two vectors $\mathbf{u}, \mathbf{v}$ are equivalent if and only if their directions are the same and their magnitudes equal, and we will write $\mathbf{u} = \mathbf{v}$; see figure (1b). Observe that two vectors $\mathbf{u}, \mathbf{v}$ with different initial points are equivalent if and only if the translation taking the initial point of $\mathbf{u}$ to the initial point of $\mathbf{v}$ translates the terminal point of $\mathbf{u}$ to the terminal point of $\mathbf{v}$. Be aware that in applications we may have cause to prefer to place a vector with a particular initial or terminal point, but there is no preferred location for our abstract vectors in the sense of the above equivalence.

We remark that is convenient to define things like "the space of all real, 3-dimensional vectors". To do so rigorously, we would need to explore vector spaces and linear algebra in greater detail, but nevertheless it is convenient to have a notation for such things. Let $\mathscr{V}_{\mathbb{R}}^3$ be the set of all 3-dimensional vectors, up to the defined equivalence, and similarly let $\mathscr{V}_{\mathbb{R}}^2$ be the set of all 2-dimensional vectors up to the defined equivalence. Even more generally, there are higher dimensional analogs, here denoted $\mathscr{V}_{\mathbb{R}}^n$, which are sets of *n-vectors*. All these sets will have additional structure coming from the algebra of vectors we will define, making them into *vector spaces*[2], but the abstract study of vector space structure is the concern of a linear algebra course; we will but glimpse at it and focus on the calculations and their applications as are pertinent to our study of multivariate calculus. Thus, even if $\mathscr{V}_{\mathbb{R}}^n$ is written, you may imagine for ease of use that $n = 2$ or 3, and picture the comfortable world of arrows drawn on blackboards.

§ 1.2. Addition, Subtraction, and Scaling

Our geometric description of vectors as arrows, pointing from an initial point to a terminal point, allows us to give pictorial definitions for many of the concepts we will be concerned with. But to actually manipulate and compute with vectors, it is desirable to have representations that take advantage of coordinate systems. Before we do this though, we will introduce a notion of vector arithmetic. Vectors are useful precisely because they allow us to model linear motion and displacement: unlike points, which are static and admit no inherent arithmetic, vectors may be added and subtracted to describe changes in position, rates or directions of travel or acceleration, rotations in 3-space, and many other phenomena.

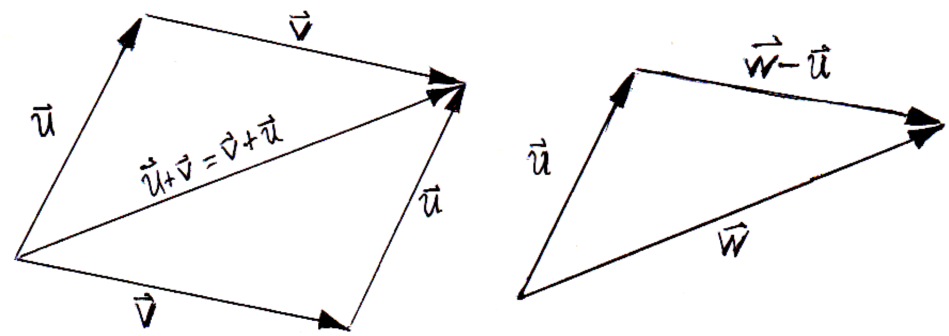
The initial and essential idea is that by placing vectors tail to tip, we locate a new terminal point and may construct a new vector from the initial point of the first vector to the terminal point of the last vector of the sum. The result of vector summation is sometimes called the *resultant*.

Consider the parallelogram in figure (2a), depicting the pair of vectors $\mathbf{u}$ and $\mathbf{v}$ emanating from a common initial point, their translates which we label as $\mathbf{u}$ and $\mathbf{v}$ as well, (in keeping with our notion of equivalence given above), and their sum, which is the diagonal vector of the parallelogram. The figure illustrates the so called parallelogram rule of vector addition.

We can derive a subtraction rule from this picture by relabeling the figure and reversing the orientation of $\mathbf{u}$. Denote by $\mathbf{w}$ the vector $\mathbf{u} + \mathbf{v}$, and by $-\mathbf{u}$ the vector obtained by reversing the

[1] In physical applications and in the context of polar coordinates, it is convenient sometimes to use a convention similar to the "bold denotes vectors, italic denotes scalar magnitudes of corresponding vectors." In such cases, symbols will be clearly defined in reference to the notations used throughout.

[2] If you want to learn about the zoo of other vector spaces with which mathematicians fascinate themselves, I highly recommend you take some linear algebra, or come bug me to motivate its study.



(a) (b)

FIGURE 2. (a) – the parallelogram rule. (b) – the vector subtraction triangle rule.

direction of . It is then clear from the preceding definitions that $\mathbf{w} + -\mathbf{u} = \mathbf{v}$. Thus, adding the *reverse of* $\mathbf{u}$ to $\mathbf{w}$ recovers $\mathbf{v}$. We then may *define* vector subtraction to be the operation of adding the reverse, and we will simply write $\mathbf{w} - \mathbf{u}$ instead of $\mathbf{w} + -\mathbf{u}$. Observe that subtraction satisfies a diagrammatic rule of"closing a triangle": if $\mathbf{w}$ and $\mathbf{u}$ are vectors emanating from a common initial point, then $\mathbf{w} - \mathbf{u}$ is the vector closing the triangle so that if its initial point is placed at the tip of $\mathbf{u}$, its terminal point will coincide with the terminal point of $\mathbf{w}$, i.e. it completes the triangle and points from $\mathbf{u}$ to $\mathbf{w}$.

An issue remains: if vectors can be added and the addition is invertible via subtraction, there ought to be an identity element! Indeed, the vector whose initial and terminal points coincide gives us such an identity. We call this vector *the zero vector*, and denote it by $\mathbf{0}$. The zero vector is the unique directionless vector and the unique vector of length zero. It is conventionally considered both parallel and perpendicular to all other vectors (a caveat we will address and exploit later). The zero vector satisfies $\mathbf{v} + \mathbf{0} = \mathbf{v}$ and $\mathbf{v} - \mathbf{v} = \mathbf{0}$, where $\mathbf{v}$ is any vector.

One may certainly consider iterated addition as one does for integers and real numbers. It is natural to write for an integer n

$$n\mathbf{v} = \underbrace{\mathbf{v} + \ldots + \mathbf{v}}_{n \text{ times}} = \sum_{i=1}^{n} \mathbf{v}\,.$$

This has an obvious geometric interpretation: $n\mathbf{v}$ is the result of stretching $\mathbf{v}$ by a factor of n, that is, the magnitude of $n\mathbf{v}$ is n times that of $\mathbf{v}$. There is however no need to restrict ourselves to only integer stretching. For any positive real number s, the vector $s\mathbf{v}$ is the vector whose direction is the same as $\mathbf{v}$ and whose length is $s\,\|\mathbf{v}\|$. For a negative real number s, we take $s\mathbf{v}$ to denote the vector whose direction is the same as the reverse of $\mathbf{v}$, and whose length is $|s|\,\|\mathbf{v}\|$, i.e., $s\mathbf{v} = -|s|\mathbf{v}$ for negative s. Naturally, for $s = 0$ or for $\mathbf{v} = \mathbf{0}$, we can take $s\mathbf{v} = \mathbf{0}$.

Proposition. *The following properties hold for any vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *in* $\mathscr{V}_{\mathbb{R}}^n$ *and any scalars* $s, t \in \mathbb{R}$.

(1) $\mathbf{u} + \mathbf{0} = \mathbf{u}$ *(vector additive identity)*
(2) $1\mathbf{u} = \mathbf{u}$ *(scaling identity)*
(3) $\mathbf{u} + (-1)\mathbf{u} = \mathbf{u} + -\mathbf{u} = \mathbf{0}$ *(vector additive inverse)*
(4) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ *(commutativity of vector addition)*
(5) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ *(associativity of vector addition)*
(6) $s(\mathbf{u} + \mathbf{v}) = s\mathbf{u} + s\mathbf{v}$ *(distributivity of scaling)*
(7) $s\mathbf{u} + t\mathbf{u} = (s + t)\mathbf{u}$ *(factoring of scalars)*
(8) $(st)\mathbf{u} = s(t\mathbf{u})$ *(associativity of scaling)*

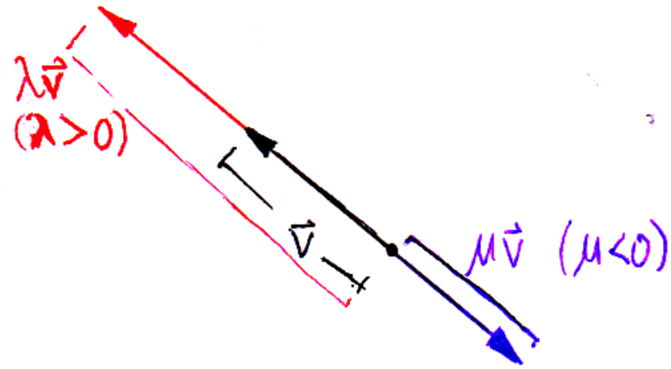


FIGURE 3. Scaling a vector.

(9) $0\mathbf{v} = \mathbf{0}$ *(annihilation by zero scaling)*
(10) $\|\mathbf{u}\| \geq 0$, *with equality* $\iff \mathbf{u} = \mathbf{0}$ *(non-degeneracy of magnitude)*
(11) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ *(triangle inequality for magnitudes)*
(12) $\|s\mathbf{v}\| = |s|\,\|\mathbf{v}\|$ *(absolute scalability of magnitude)*

A good exercise is to verify that the three dimensional vectors we have described satisfy each of the above properties, using only the geometric definitions given for vector addition and scaling, and known properties of real numbers. The first eight statements are the *axioms of an abstract vector space with real scalars*, while the last three make $\mathscr{V}_{\mathbb{R}}^n$ into a *normed vector space.* Property (9) is a simple consequence of preceding properties.

§ 1.3. Vectors in Coordinates

We start by describing *position vectors*, i.e. vectors describing the positions of points in space. Consider first a point $P(x_1, y_1, z_1) \in \mathbb{R}^3$. Then the position vector $\mathbf{p}$ associated to P is the vector whose initial point is the origin $(0,0,0) \in \mathbb{R}^3$ and whose terminal point is P. This motivates the following coordinate description of $\mathbf{p}$: we write

$$\mathbf{p} = \langle x_1, y_1, z_1 \rangle \,.$$

This notation is referred to as *angle bracket notation.* The angle bracket notation, while convenient, somewhat obscures the relation between the coordinates and vector addition. To see this relationship more concretely, we need another notation.

We define three special *coordinate vectors*, whose labels are historical[3]:

$$\hat{\mathbf{\imath}} := \langle 1,0,0 \rangle \,, \quad \hat{\mathbf{\jmath}} := \langle 0,1,0 \rangle \,, \quad \hat{\mathbf{k}} := \langle 0,0,1 \rangle \,.$$

Note that these give the position vectors for points a unit distance from the origin along the positive coordinate axes. Using the definition of vector addition, it is clear (see the figure below) that

$$\mathbf{p} = x_1\hat{\mathbf{\imath}} + y_1\hat{\mathbf{\jmath}} + z_1\hat{\mathbf{k}} \,.$$

The numbers x_1, y_1, and z_1 are called the components in the $\hat{\mathbf{\imath}}$, $\hat{\mathbf{\jmath}}$, and $\hat{\mathbf{k}}$ directions, respectively. It is not uncommon to also refer to them respectively as the x, y and z components. An expression of the form $a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3$ for scalars $a, b, c \in \mathbb{R}$ and vectors $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ is called a *linear combination* of the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$; thus writing the coordinate form of $\mathbf{p}$ amounts to expressing it as a linear combination of the coordinate vectors $\hat{\mathbf{\imath}}$, $\hat{\mathbf{\jmath}}$ and $\hat{\mathbf{k}}$.

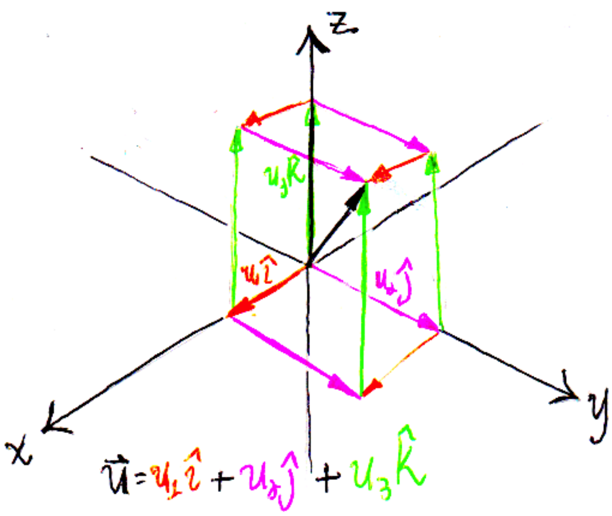


FIGURE 4. Visualizing the components of a vector in three dimensions.

A further notation, common in linear algebra, is to write the components vertically in a column enclosed in a bracket or parentheses as follows:

$$\mathbf{p} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \qquad \text{or} \qquad \mathbf{p} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}.$$

This notation is convenient when describing linear algebraic operations, such as matrix multiplication, acting on vectors. We will use this notation sparingly in these notes.

For each of the above notations, one may omit the third component (i.e. the $z_1\hat{\mathbf{k}}$ term) and obtain a notation suited to describing two-dimensional vectors. The convention of writing elements of $\mathscr{V}_{\mathbb{R}}^2$ in coordinates as linear combinations of $\hat{\boldsymbol{\imath}}$ and $\hat{\boldsymbol{\jmath}}$ conveniently also gives us a natural way to view the vector space $\mathscr{V}_{\mathbb{R}}^2$ as being embedded in $\mathscr{V}_{\mathbb{R}}^3$, just as the xy-coordinate plane sits in $\mathbb{R}^3$. If we need to strongly emphasize the difference between a planar vector $x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}}$ and a spatial vector with the same coordinates, we will write the latter as $x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + 0\hat{\mathbf{k}}$, or in angle bracket notation $\langle x, y, 0\rangle$.

Example 1.1. Figure 5 illustrates several position 2-vectors with accompanying coordinate representations. Note that for $\mathbb{R}^2$ we define $\hat{\boldsymbol{\imath}} = \langle 1, 0\rangle$ and $\hat{\boldsymbol{\jmath}} = \langle 0, 1\rangle$.

An arbitrary vector has a coordinate expression which can be found by translating the initial point to the origin, and treating it as a position vector, thus any $\mathbf{v} \in \mathscr{V}_{\mathbb{R}}^3$ can be described in coordinates using angle bracket notation, as a linear combination of the vectors $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$, and $\hat{\mathbf{k}}$, or in column notation. Each coordinate notation involves listing the same triple of numbers, just in a slightly different fashion. We can thus easily interpret the components as representing weights in a linear combination, building the vector $\mathbf{v}$ by specifying vectors parallel to the rectangular coordinate axes.

Many students first encounter the idea of components in physics, where vectors representing physical data such as velocities or forces are often decomposed as sums of vectors along directions chosen sensibly for a given physical configuration. This is not so different from our components: the essential idea of components is just to decompose vectors as linear combinations of some chosen set of vectors, and in our case we chose the (seemingly natural) set of vectors $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$, and $\hat{\mathbf{k}}$ as a *basis* to build other vectors. But can one just choose any three vectors to build all of $\mathscr{V}_{\mathbb{R}}^3$?

The answer is no. If one of the two vectors can be written as a linear combination of the other two, then there will be 3-vectors which cannot be written as linear combinations of the three chosen. This brings us to an important concept associated to set of vectors, called *linear independence*. Vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathscr{V}_{\mathbb{R}}^n$ are said to be linearly independent if the only linear combination of them equaling $\mathbf{0}$ is the trivial one, whose coefficients are all zero, i.e. if

$$\sum a_i \mathbf{v_i} = \mathbf{0}\,,$$

then $a_i = 0$ for $i = 1, \ldots k$. If a nontrivial combination vanishes, then we say the set of vectors is *linearly dependent*. Convince yourself of the following facts:

- A set of $k < n$ vectors $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathscr{V}_{\mathbb{R}}^n$ is linearly dependent if and only if there is some vector in the set which can be rewritten as a linear combination of the others.
- Two vectors are linearly independent if they are not parallel.
- Any set of 3 or more vectors in $\mathscr{V}_{\mathbb{R}}^2$ is linearly dependent, and any set of 4 or more vectors in $\mathscr{V}_{\mathbb{R}}^3$ is linearly dependent.
- Three vectors in $\mathscr{V}_{\mathbb{R}}^n$ with $n \geq 3$ are linearly independent if there is no line or plane that contains all of them (not to be confused with *their terminal points*, thought of as positions, being contained in a plane–indeed you will later see that three non-collinear *points* in $\mathbb{R}^3$ determine a plane!).

Given three linearly independent 3-vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathscr{V}_{\mathbb{R}}^3$, we can write *any other vector* in $\mathscr{V}_{\mathbb{R}}^3$ as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. This gives one a means of defining other, *non-rectangular* linear coordinate systems. We'll stick to rectangular coordinates for now, but we occasionally use the notion of linear independence in definitions and discussion of algebraic properties.

[3]The choice of the labels of the standard (rectangular) basis elements $\hat{\imath}$, $\hat{\jmath}$, and $\hat{\mathbf{k}}$ is explained in the discussion of William Rowan Hamilton's quaternions in §6.4. Hamilton is also responsible for the introduction of the term "vector". Another common notation for the standard basis of $\mathscr{V}_{\mathbb{R}}^n$ is $\mathbf{e}_i$, for integers $i = 1, \ldots, n$. Here, $\mathbf{e}$ may be remembered as standing for "elementary," though the $\mathbf{e}$ originally stands for *Einheitsvektor*, which translates as *unit vector*. This

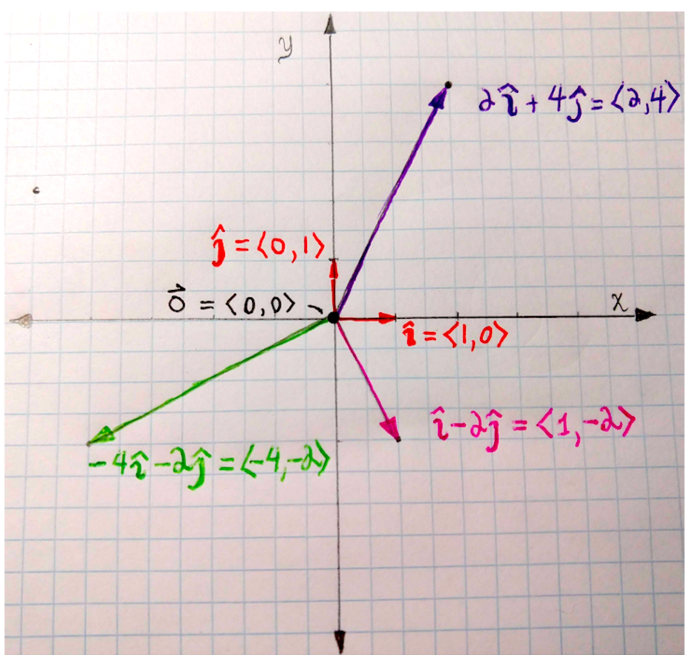


Figure 5. A few 2-vectors, labeled with both $\hat{\imath}$, $\hat{\jmath}$ notation and angle bracket notation.

An arbitrary vector drawn with initial point $P(x_1, y_1, z_1)$ and terminal point $Q(x_2, y_2, z_2)$ can be regarded as a *displacement vector* $\overrightarrow{PQ}$, modeling how Q is displaced from P. Since we wish to express such vectors in coordinates, we need to understand coordinate arithmetic.

Let $\mathbf{a} = a_1\hat{\mathbf{\imath}} + a_2\hat{\mathbf{\jmath}}\hat{\mathbf{k}}$ and $\mathbf{b} = b_1\hat{\mathbf{\imath}} + b_2\hat{\mathbf{\jmath}}$ be two arbitrary position vectors in the plane. From figure (6) and the known properties of vector arithmetic, it should be clear that the coordinate forms for vector addition and scaling are given as

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1)\hat{\mathbf{\imath}} + (a_2 + b_2)\hat{\mathbf{\jmath}} = \langle a_1 + b_1, a_2 + b_2 \rangle\,,$$

$$s\mathbf{a} = sa_1\hat{\mathbf{\imath}} + sa_2\hat{\mathbf{\jmath}} = \langle sa_1, sa_2 \rangle\,.$$

We can easily generalize this to 3-vectors, to deduce that *vector addition is performed by adding like components*, and *scaling a vector by a scalar s scales each component by s individually.*

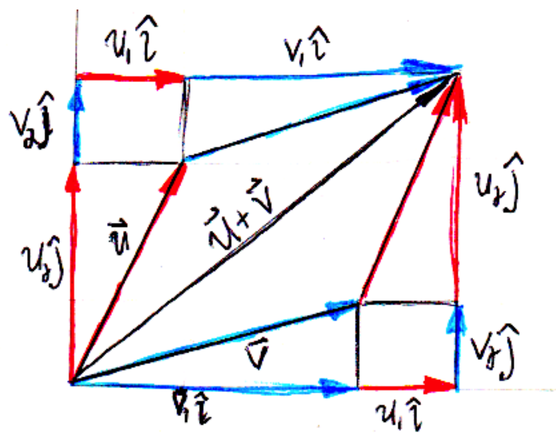


FIGURE 6. Adding 2-vectors by adding their components. Can you draw a pictorial proof for 3-vectors? What algebraic properties from above would you use to prove the analogous result at once for n-vectors with arbitrary $n \geq 3$?

Returning to the issue of representing displacement vectors in coordinates, it becomes clear (draw a picture!) that we can represent the displacement vector $\overrightarrow{PQ}$ as the vector difference $\mathbf{q} - \mathbf{p}$ of the corresponding position vectors for the terminal and initial points. Thus, if $\mathbf{p} = x_1\hat{\mathbf{\imath}} + y_1\hat{\mathbf{\jmath}} + z_1\hat{\mathbf{k}}$ and $\mathbf{q} = x_2\hat{\mathbf{\imath}} + y_2\hat{\mathbf{\jmath}} + z_2\hat{\mathbf{k}}$ then

$$\overrightarrow{PQ} = \mathbf{q} - \mathbf{p} = (x_2 - x_1)\hat{\mathbf{\imath}} + (y_2 - y_1)\hat{\mathbf{\jmath}} + (z_2 - z_1)\hat{\mathbf{k}}\,.$$

Example 1.2. The points $Q(-1, 5, 3)$ and $R(6, 7, 8)$ are vertices of a parallelogram adjacent to a vertex at $P(5, 4, 3)$. Find the fourth vertex.

Solution: Since Q and R are each adjacent to P, the displacement vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ represent sides of the parallelogram. Let S denote the fourth point of the parallelogram. By applying the parallelogram rule, we can find a displacement vector $\overrightarrow{PS}$ from P to the fourth point. Namely,

$$\overrightarrow{PS} = \overrightarrow{PQ} + \overrightarrow{PR}\,.$$

In coordinates, we have

$$\overrightarrow{PQ} = \mathbf{q} - \mathbf{p} = \langle -6, 1, 0 \rangle\,, \qquad \overrightarrow{PR} = \mathbf{r} - \mathbf{p} = \langle 1, 3, 5 \rangle\,.$$

notation was introduced by Hermann Grassmann, in his seminal work *Die Lineale Ausdehnungslehre, ein neuer Zweig der Mathematik* published in 1844.

Thus

$$\overrightarrow{PS} = \langle -6, 1, 0\rangle + \langle 1, 3, 5\rangle = \langle -5, 4, 5\rangle\,.$$

The coordinates of S can be found by adding the position vector $\mathbf{p}$ for P to the displacement vector $\overrightarrow{PS}$, since P is the initial point of the displacement vector, and the desired point S is the terminal point. Thus, the position vector $\mathbf{s}$ for the point S is

$$\mathbf{s} = \mathbf{p} + \overrightarrow{PS} = \langle 5, 4, 3\rangle + \langle -5, 4, 5\rangle = \langle 0, 8, 8\rangle\,.$$

This gives the fourth point as $S(0, 8, 8)$. In general, to solve a problem like this, one may note that

$$\overrightarrow{PS} = \mathbf{q} - \mathbf{p} + \mathbf{r} - \mathbf{p} \text{ and}$$

$$\overrightarrow{PS} = \mathbf{s} - \mathbf{p}\,, \text{ whence}$$

$$\mathbf{s} - \mathbf{p} = \mathbf{q} - \mathbf{p} + \mathbf{r} - \mathbf{p} \implies \mathbf{s} = \mathbf{q} + \mathbf{r} - \mathbf{p}\,.$$

One can now check that, indeed,

$$\langle 0, 8, 8\rangle = \langle -1, 5, 3\rangle + \langle 6, 7, 8\rangle - \langle 5, 4, 3\rangle\,.$$

The *centroid* of a collection of k points in $\mathbb{R}^n$ is defined as follows: if the points have position vectors $\mathbf{p}_1, \ldots, \mathbf{p}_k$, then the centroid C is the point whose position vector $\mathbf{c}$ is given by

$$\mathbf{c} = \frac{1}{k}\sum_{i=1}^{k} \mathbf{p}_k\,.$$

For example, the centroid of two points is the midpoint of the line segment connecting them. For three non-collinear points P, Q and R, the centroid is the point G in the interior of the triangle $\triangle PQR$ that lies at the intersection of the medians, which are line segments from the midpoints of the triangle's sides to the opposite vertices. The position vector $\mathbf{g}$ of G is

$$\mathbf{g} = \frac{1}{3}(\mathbf{p} + \mathbf{q} + \mathbf{r})\,.$$

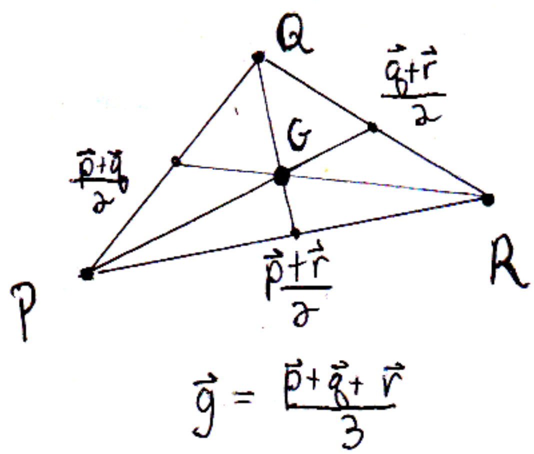


FIGURE 7. The centroid of $\triangle PQR$, which is the intersection of the medians, and the vector average of the positions of the vertices P, Q and R.

Example 1.3. Find the centroid of the points $P(\pi, e)$, $Q(1/e, -\pi)$, and $R(\sqrt{\pi}, e^\pi)$.

Solution: The position vectors are

$$\mathbf{p} = \pi\hat{\mathbf{\imath}} + e\hat{\mathbf{\jmath}}\,, \qquad \mathbf{q} = 1/e\hat{\mathbf{\imath}} - \pi\hat{\mathbf{\jmath}}\,, \qquad \mathbf{r} = \sqrt{\pi}\hat{\mathbf{\imath}} + e^\pi\hat{\mathbf{\jmath}}\,.$$

Thus the centroid has position vector

$$\begin{aligned}\mathbf{g} &= \frac{1}{3}\left(\pi\hat{\mathbf{\imath}} + e\hat{\mathbf{\jmath}} + 1/e\hat{\mathbf{\imath}} - \pi\hat{\mathbf{\jmath}} + \sqrt{\pi}\hat{\mathbf{\imath}} + e^\pi\hat{\mathbf{\jmath}}\right)\\ &= \frac{1}{3}(\pi + 1/e + \sqrt{\pi})\hat{\mathbf{\imath}} + \frac{1}{3}(e - \pi + e^\pi)\hat{\mathbf{\jmath}}\,.\end{aligned}$$

Thus

$$G\left(\frac{\pi + 1/e + \sqrt{\pi}}{3}, \frac{e - \pi + e^\pi}{3}\right)$$

is the centroid of $\triangle PQR$.

Using the Pythagorean theorem, one can readily confirm that our notion of vector magnitude should be computable by summing the squares of components, and then taking the square root of this sum. Thus if $\mathbf{a} = a_1\hat{\mathbf{\imath}} + a_2\hat{\mathbf{\jmath}} + a_3\hat{\mathbf{k}}$, we have

$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2}\,.$$

A vector is called a *unit vector* if it has length 1. To get a unit vector from another vector, we scale by the reciprocal of the magnitude:

$$\mathbf{u} \mapsto \hat{\mathbf{u}} = \frac{\mathbf{u}}{\|\mathbf{u}\|}\,.$$

Here, the hat '^' in $\hat{\mathbf{u}}$ denotes that it is a unit vector. Conventionally, if you see a vector with a hat in these notes, it has length 1.

Example 1.4. Find a unit vector in the direction of $3\hat{\mathbf{\imath}} + 4\hat{\mathbf{\jmath}}$, and a unit vector making an angle of $\pi/3$ with $\hat{\mathbf{\imath}}$.

Solution: The magnitude of $3\hat{\mathbf{\imath}} + 4\hat{\mathbf{\jmath}}$ is $\sqrt{3^2 + 4^2} = \sqrt{25} = 5$, so the unit vector in the same direction as $3\hat{\mathbf{\imath}} + 4\hat{\mathbf{\jmath}}$ is $\frac{3}{5}\hat{\mathbf{\imath}} + \frac{4}{5}\hat{\mathbf{\jmath}}$. For a unit vector making an angle of $\pi/3$ with $\hat{\mathbf{\imath}}$, we can draw a right triangle based at the origin, parallel to $\hat{\mathbf{\imath}}$, with length one hypotenuse. Then the desired unit vector has initial point at the origin and final point with coordinates $(\cos\frac{\pi}{3}, \sin\frac{\pi}{3}) = (\frac{1}{2}, \frac{\sqrt{3}}{2})$, whence the desired unit vector is $\frac{1}{2}\hat{\mathbf{\imath}} + \frac{\sqrt{3}}{2}\hat{\mathbf{\jmath}}$.

Example 1.5. Resolve the vector $\mathbf{v}$ with magnitude 2 making an angle of $\pi/12$ with the positive x-axis into x and y components.

Solution: A vector $\mathbf{v}$ making an angle θ with the positive x-axis determines an arrow the line segment of which, when placed with its tail at the origin, forms the hypotenuse of a right triangle with a horizontal leg of length $\|\mathbf{v}\|\cos\theta$ and a vertical leg of length $\|\mathbf{v}\|\sin\theta$. For our example, $\mathbf{v}$ has $\|\mathbf{v}\| = 2$ and $\theta = \pi/12$, so it resolves into components as

$$\mathbf{v} = 2\cos\left(\tfrac{\pi}{12}\right)\hat{\mathbf{\imath}} + 2\sin\left(\tfrac{\pi}{12}\right)\hat{\mathbf{\jmath}} = \left\langle 2\cos\left(\tfrac{\pi}{12}\right), 2\sin\left(\tfrac{\pi}{12}\right)\right\rangle.$$

It is tempting to merely plug the numbers into a calculator to get some decimal approximation; this might be good enough in a physics problem, for example, and may be the only recourse for many angles encountered "in the wild". However $\pi/12$ isn't as bad an angle as one might expect, and so we have an opportunity here to review the sine and cosine addition/subtraction formulae, which are used in the forthcoming discussion of dot products. We will give an exact answer for the components in this particular case, using that $\pi/12 = \pi/3 - \pi/4$.

Recall that for arbitrary angles α and β we have the identities

$$\sin(\alpha \pm \beta) = \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta)\,,$$
$$\cos(\beta \pm \alpha) = \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta)\,.$$

Using these identities we obtain

$$\begin{aligned}\mathbf{v} &= 2\cos\left(\tfrac{\pi}{3} - \tfrac{\pi}{4}\right)\hat{\imath} + \sin\left(\tfrac{\pi}{3} - \tfrac{\pi}{4}\right)\hat{\jmath} \\ &= 2\left[\cos\left(\tfrac{\pi}{3}\right)\cos\left(\tfrac{\pi}{4}\right) + \sin\left(\tfrac{\pi}{3}\right)\sin\left(\tfrac{\pi}{4}\right)\right]\hat{\imath} + 2\left[\sin\left(\tfrac{\pi}{3}\right)\cos\left(\tfrac{\pi}{4}\right) - \cos\left(\tfrac{\pi}{3}\right)\sin\left(\tfrac{\pi}{4}\right)\right]\hat{\jmath} \\ &= \frac{1}{2}(\sqrt{2} + \sqrt{6})\hat{\imath} + \frac{1}{2}(\sqrt{6} - \sqrt{2})\hat{\jmath}\,.\end{aligned}$$

Example 1.6. Suppose the 2-vector $\mathbf{u}$ makes an angle of 144° with the positive x-axis, and has magnitude $\|\mathbf{u}\| = 3$, and suppose that the 2-vector $\mathbf{v}$ makes an angle of $-18°$ with the positive x-axis and has a magnitude of $\|\mathbf{v}\| = 4$. Describe the resultant by giving its magnitude and the angle it makes with the positive x axis. Round to three decimal places.

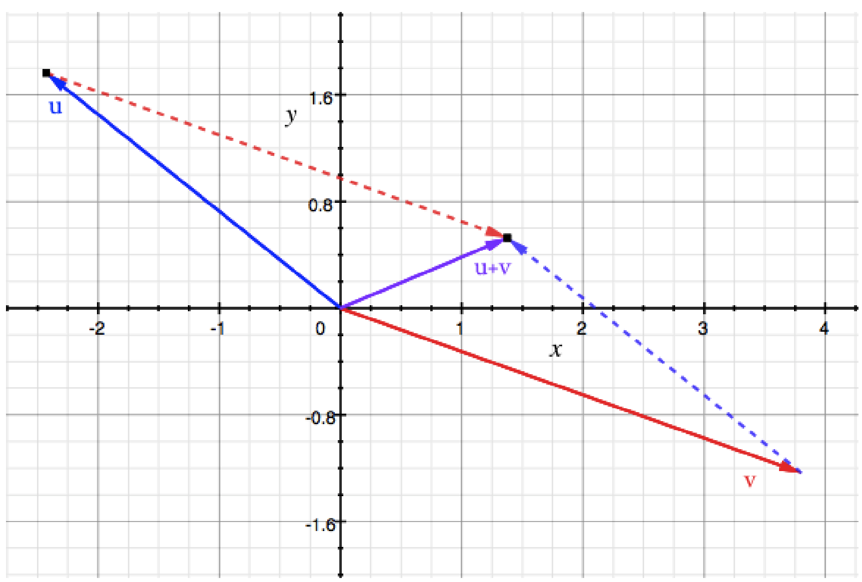


FIGURE 8. The vectors $\mathbf{u}$, $\mathbf{v}$ and their resultant $\mathbf{u}+\mathbf{v}$.

Solution: The strategy is to resolve each vector into components, then add the like components to obtain components for the resultant, and finally to compute the desired magnitude and angle from the components of the resultant.

$$\mathbf{u} = 3\langle\cos(144°), \sin(144°)\rangle = 3\langle -0.8080\ldots, 0.5877\ldots\rangle \approx \langle -2.427, 1.763\rangle\,,$$

$$\mathbf{v} = 4\langle\cos(-18°), \sin(-18°)\rangle = 4\langle 0.9510\ldots, -0.3090\ldots\rangle \approx \langle 3.804, -1.236\rangle\,.$$

Thus

$$\mathbf{u}+\mathbf{v} \approx \langle 1.377, 0.527\rangle\,.$$

The magnitude of the resultant is then approximately

$$\|\mathbf{u}+\mathbf{v}\| \approx \sqrt{1.377^2 + 0.527^2} \approx 1.475\,.$$

Observe that the angle made with the positive x-axis can be computed using an arctangent of the quotient of the $\hat{\mathbf{j}}$ component by the $\hat{\mathbf{i}}$ component:

$$\theta \approx \arctan\left(\frac{0.527}{1.377}\right) \approx 20.951^\circ\,.$$

In radians, the angle is approximately 0.366.

An intrepid student of geometry and trigonometry might recognize the angles and note that the sines and cosines are exactly computable[4]. Such a student might give the following intimidating but exact solution:

$$\mathbf{u} = \frac{-3-3\sqrt{5}}{4}\,\hat{\mathbf{i}} + \frac{3\sqrt{10-2\sqrt{5}}}{4}\,\hat{\mathbf{j}}\,,\quad \mathbf{v} = \sqrt{10-2\sqrt{5}}\,\hat{\mathbf{i}} + (1-\sqrt{5})\,\hat{\mathbf{j}}\,,$$

$$\mathbf{u}+\mathbf{v} = \frac{1}{4}\left(-3-3\sqrt{5}+4\sqrt{10+2\sqrt{5}}\right)\hat{\mathbf{i}} + \left(1-\sqrt{5}+\frac{3}{4}\sqrt{10-2\sqrt{5}}\right)\hat{\mathbf{j}}$$

$$\|\mathbf{u}+\mathbf{v}\| = \sqrt{25-6\sqrt{10+2\sqrt{5}}}\,,$$

$$\theta = \arctan\left(\frac{4-4\sqrt{5}+3\sqrt{10-2\sqrt{5}}}{-3-3\sqrt{5}+4\sqrt{10+2\sqrt{5}}}\right)\,.$$

This, together with the previous example, shows there is occasionally hope of obtaining exact answers; be warned that one will generally not often encounter such special angles in the wild, nor does one often need such exact answers in applications. Thus, for the question of resolving vectors into components, and then recovering magnitudes and angles from a resultant, expect to obtain only approximate answers, or complicated answers involving (possibly iterated) square roots.

---

§ 1.4. Problems

(1) If you are given three points, how many different ways can you complete them to a parallelogram? If the three points have position vectors $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$, what are the vector formulae for the possible position vectors of the final points of the parallelograms?

(2) Why are two linearly *dependent* vectors parallel?

(3) For each of the properties (3) - (8) of vectors listed in the proposition above, use pictures and geometric reasoning to argue the truth of the properties without appealing to coordinate algebra.

(4) Argue the absolute scalability property without using coordinates.

(5) Prove the absolute scalability property for vectors in $\mathscr{V}^3_{\mathbb{R}}$ using the coordinate formula for magnitude.

(6) Prove the coordinate formula for the magnitude of a vector in $\mathscr{V}^3_{\mathbb{R}}$ using the pythagorean theorem.

[4]One can compute the desired sines and cosines using double and half angle identities, from the knowledge that $\cos 36^\circ = (1+\sqrt{5})/4$, which happens to be half of the golden ratio and is also the apothem length of a regular pentagon inscribed in the unit circle. To see the latter, cut the pentagon into ten triangles. See problems 14 and 15 below.

(7) Prove the Pythagorean theorem. (Just kidding! But actually, if you can prove this visually, that's worth something...)

(8) For a fixed vector $\mathbf{u}$ in $\mathscr{V}^2_{\mathbb{R}}$ or $\mathscr{V}^3_{\mathbb{R}}$, what geometric object corresponds to the set of all vectors *parallel* to $\mathbf{u}$?

(9) For a fixed vector $\mathbf{n} \in \mathscr{V}^3_{\mathbb{R}}$, what is the geometric object corresponding to the set of all vectors *perpendicular* to $\mathbf{n}$?

(10) For a fixed pair of vectors $\mathbf{u}, \mathbf{v} \in \mathscr{V}^2_{\mathbb{R}}$, what geometric object corresponds to the set of all vectors $\mathbf{x}$ such that $\|\mathbf{x} - \mathbf{u}\| = \|\mathbf{x} - \mathbf{v}\|$? What if we are in $\mathscr{V}^3_{\mathbb{R}}$?

(11) What geometric object corresponds to the set of all vectors of unit length in $\mathscr{V}^2_{\mathbb{R}}$? What geometric object corresponds to the set of all vectors of unit length in $\mathscr{V}^3_{\mathbb{R}}$?

(12) For a vector $\mathbf{u} = u_1\hat{\imath} + u_2\hat{\jmath}$, give the coordinate form of the vector $\mathbf{u}' = \mathcal{R}_\theta\mathbf{u}$ obtained by rotating $\mathbf{u}$ counterclockwise by an angle θ.

(13) Use the previous exercise to describe equations relating the x, y coordinate system to an x', y' coordinate system that is obtained by rotating the coordinate axes by an angle θ.

(14) For $n = 3$, 4, 5, and 6, give exact coordinates for position vectors for the vertices of a regular n-gon (all sides equal length, all interior angles equal) inscribed in the unit circle, with one vertex on the point $(1, 0)$. Write down exact coordinates for displacement vectors giving the sides of each n-gon, oriented so that walking along the sides of the n-gon, one circles the origin counter-clockwise.

(15) For each $n = 3$, 4, 5, and 6 and regular n-gons as in the previous problem, what are the sums of the displacement vectors giving the counter-clockwise oriented perimeter? What are the perimeters? (You should give a positive scalar.) For each n-gon, what is the sum of the position vectors of the vertices?

(16) Velocity is the derivative of position, and as position can be described as a vector quantity, so too is velocity a vector quantity. If a particle is traveling around a unit circle at constant unit speed, what, intuitively, should the velocity vector of the particle be at any given time?

(17) Acceleration is the derivative of velocity. For a particle traversing a circle as in the preceding problem, what, intuitively, is the acceleration vector at any particular point in time?

(18) Suppose you are on a boat, traveling from a calm tributary into a river. The tributary meets the river at a right angle, and you maintain a constant speed relative to the surface of the water as you enter the main river. The river is flowing twice as fast as your speed relative to its surface. You wish to travel to the point straight across from where you enter the river. At what angle into the flow should you point your boat in order to get there? Why doesn't this depend on the width of the river?

2. The Dot Product

We will define an operation which takes as input two vectors and outputs a scalar:

$$\cdot : \mathscr{V}^n_{\mathbb{R}} \times \mathscr{V}^n_{\mathbb{R}} \to \mathbb{R}\,,$$
$$\mathbf{a}, \mathbf{b} \mapsto \mathbf{a} \cdot \mathbf{b}$$

It is called the *dot product* or *scalar product*. The common approach to defining and studying the dot product is to define it via a formula depending upon the coordinate representation in rectangular coordinates. This has the advantage that it is immediately clear that the result is defined in any finite number of dimensions. We will go a different route, studying first a geometric problem that motivates the coordinate expression in 2 dimensions. We will then be able to argue that the generalization of the coordinate expression continues to accomplish the geometric goal of computing such a scalar product.

§ 2.1. Motivation from Projections

Consider first two planar unit vectors $\hat{\mathbf{u}}, \hat{\mathbf{v}} \in \mathscr{V}^2_{\mathbb{R}}$. Placing them with their initial points at the origin, we have a diagram as in figure (9)[5]. Let θ denote the angle made between $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$. We can decompose $\hat{\mathbf{v}}$ into a vector along $\hat{\mathbf{u}}$ and a vector perpendicular to $\hat{\mathbf{u}}$. We call the former vector that is parallel to $\hat{\mathbf{u}}$ the *projection of* $\hat{\mathbf{v}}$ *onto* $\hat{\mathbf{u}}$, and denote it $\mathrm{proj}_{\hat{\mathbf{u}}}\,\hat{\mathbf{v}}$. Using that $\|\hat{\mathbf{v}}\| = 1 = \|\hat{\mathbf{u}}\|$, we have from the definition of cosine that $\|\mathrm{proj}_{\hat{\mathbf{u}}}\,\hat{\mathbf{v}}\| = \|\hat{\mathbf{v}}\|\,|\cos\theta| = |\cos\theta|$ and

$$\mathrm{proj}_{\hat{\mathbf{u}}}\,\hat{\mathbf{v}} = \cos\theta\,\hat{\mathbf{u}}\,.$$

Observe that if $\cos\theta < 0$ then $\hat{\mathbf{v}}$ makes an obtuse angle with $\hat{\mathbf{u}}$ and thus projects to give a vector parallel but opposite in orientation to $\hat{\mathbf{u}}$. We can write the orthogonal piece as a difference $\hat{\mathbf{v}} - \mathrm{proj}_{\hat{\mathbf{u}}}\,\hat{\mathbf{v}}$, and observe that its length is $|\sin\theta|$.

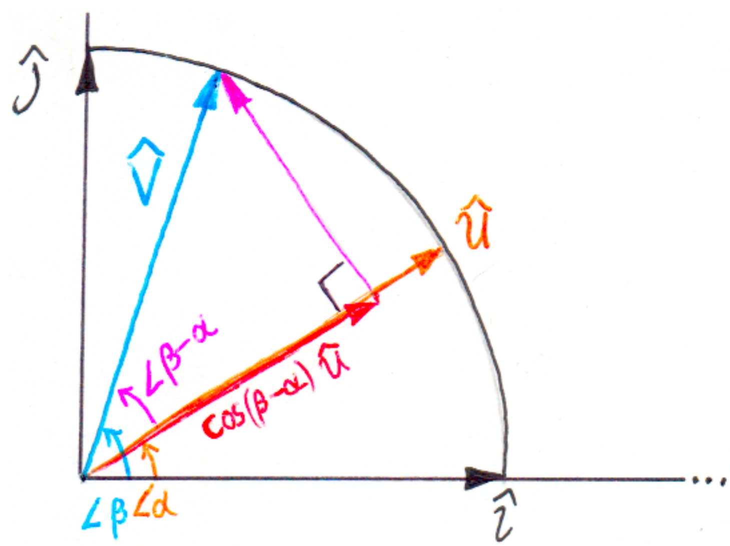


FIGURE 9. Projecting the unit vector $\hat{\mathbf{v}}$ onto the unit vector $\hat{\mathbf{u}}$ in the first quadrant of the unit circle.

[5] I've chosen to draw these vectors in the first quadrant, but the arguments make sense even if they are in different quadrants. To see this, consider the symmetries of the circle and of the corresponding relations for trigonometric functions.

We can apply the above reasoning to obtain coordinate expressions for the unit vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ in terms of trigonometric functions. If α is the angle between $\hat{\mathbf{u}}$ and $\hat{\mathbf{\imath}}$ and β is the angle between $\hat{\mathbf{v}}$ and $\hat{\mathbf{\imath}}$ then

$$\hat{\mathbf{u}} = \cos\alpha\,\hat{\mathbf{\imath}} + \sin\alpha\,\hat{\mathbf{\jmath}}\,,$$

$$\hat{\mathbf{v}} = \cos\beta\,\hat{\mathbf{\imath}} + \sin\beta\,\hat{\mathbf{\jmath}}\,.$$

Writing $\hat{\mathbf{u}} = u_1\hat{\mathbf{\imath}} + u_2\hat{\mathbf{\jmath}}$ and $\hat{\mathbf{v}} = v_1\hat{\mathbf{\imath}} + v_2\hat{\mathbf{\jmath}}$, observe that

$$\cos\theta = \cos(\beta-\alpha) = \cos\alpha\cos\beta + \sin\alpha\sin\beta = u_1v_1 + u_2v_2\,.$$

Thus, we can compute the cosine of the angle between two unit vectors from a simple expression involving only the components.

Observe that if we scale $\hat{\mathbf{v}}$ by a positive $\lambda \in \mathbb{R}$ to a non-unit vector $\mathbf{v} = \lambda\hat{\mathbf{v}} = v_1'\hat{\mathbf{\imath}} + v_2'\hat{\mathbf{\jmath}}$ then the projection $\mathrm{proj}_{\hat{\mathbf{u}}}\,\mathbf{v}$ also scales, and the expression $u_1v_1' + u_2v_2' = \lambda(u_1v_1 + u_2v_2)$ instead computes $\|\mathbf{v}\|\cos\theta$. If we scale $\hat{\mathbf{u}}$ by a positive $\mu \in \mathbb{R}$ to a non-unit vector $\mathbf{u} = \mu\hat{\mathbf{u}} = u_1'\hat{\mathbf{\imath}} + u_2'\hat{\mathbf{\jmath}}$, the projection does not change length, but the formula for it is expressible as

$$\|\mathbf{v}\|\cos(\theta)\,\hat{\mathbf{u}} = \|\mathbf{v}\|\cos(\theta)\frac{\mu^2\hat{\mathbf{u}}}{\mu^2} = \left(\frac{\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta}{\|\mathbf{u}\|}\right)\frac{\mathbf{u}}{\|\mathbf{u}\|}\,.$$

The expression in the numerator of the coefficient on $\hat{\mathbf{u}} = \mathbf{u}/\|\mathbf{u}\|$ is what we will define to be the dot product of $\mathbf{u}$ and $\mathbf{v}$. Such a definition[6] will still make sense in higher dimensions as two vectors $\mathbf{u}, \mathbf{v} \in \mathscr{V}_{\mathbb{R}}^n$ determine either a plane or a line, and one can measure the angle between them within this smaller subspace, provided it is endowed with the same geometric notion of angle as the plane $\mathbb{R}^2$.

Definition. Let $\mathbf{u}, \mathbf{v} \in \mathscr{V}_{\mathbb{R}}^n$ be two vectors separated by an angle of $\theta \in [0, \pi]$. Then the dot product $\mathbf{u}\cdot\mathbf{v}$ is the scalar quantity

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta\,.$$

It follows that the cosine of the angle θ of separation between a pair of vectors $\mathbf{u}$ and $\mathbf{v}$ whose dot product $\mathbf{u}\cdot\mathbf{v}$ is known is given by the expression

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\,.$$

However, to take advantage of this fact, we need another way to compute the dot product. We've already seen that it can be done with coordinates when $\mathbf{u}$ and $\mathbf{v}$ are planar, and it is a simple leap to generalize it.

[6]This is in some sense opposite to the usual approach, which is to define dot products algebraically, e.g. via coordinates, and then show by the law of cosines that the dot product is proportional to the cosine of the angle of separation via the product of magnitudes. One then *defines* angle in other contexts by first defining some inner product (a product satisfying the algebraic properties of the dot product) and using it to define cosines and angles. To arrest any concerns regarding the notion of angle, observe that since we are tacitly working in *euclidean normed topological vector spaces*, we can define circles via the set of unit vectors in a 2-dimensional subspace, and we can define the separation angle in radians between two vectors as the arc length of the smallest circular arc between the corresponding unit vectors. Arc-length is of course computable provided we have a notion of integration, which we always assume since this is a calculus course. We'll describe arc-length computations in the next set of notes.

§ 2.2. Coordinate Expressions

Our preceding discussion makes it clear that the dot product of $\mathbf{u}, \mathbf{v} \in \mathscr{V}_{\mathbb{R}}^2$ satisfies

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 \,,$$

where $u_1 = \mathbf{u} \cdot \hat{\mathbf{\imath}}$ is the x component of $\mathbf{u}$, $u_2 = \mathbf{u} \cdot \hat{\mathbf{\jmath}}$ is the y component of $\mathbf{u}$, and similarly $v_1 = \mathbf{v} \cdot \hat{\mathbf{\imath}}$, $v_2 = \mathbf{v} \cdot \hat{\mathbf{\jmath}}$. The nice result is that this generalizes easily:

Proposition. *If $\mathbf{u}, \mathbf{v} \in \mathscr{V}_{\mathbb{R}}^3$ are given in coordinates as $\mathbf{u} = u_1\hat{\mathbf{\imath}} + u_2\hat{\mathbf{\jmath}} + u_3\hat{\mathbf{k}}$ and $\mathbf{v} = v_1\hat{\mathbf{\imath}} + v_2\hat{\mathbf{\jmath}} + v_3\hat{\mathbf{k}}$, then the coordinate expression for their dot products is*

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3 \,.$$

More generally, if $\mathbf{u} = \langle u_1, \dots u_n \rangle$, $\mathbf{v} = \langle v_1, \dots v_n \rangle$ are vectors in $\mathscr{V}_{\mathbb{R}}^n$, then the dot product is computable as

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \ldots + u_n v_n = \sum_{i=1}^{n} u_i v_i \,.$$

If you prefer, you may take the formula given in the preceding proposition as the definition, and then prove that the geometric definition is equivalent. Armed with this definition, one can actually *compute projections and cosines of angles between vectors.*

Recall, lines are said to be *perpendicular* or *orthogonal* to each other if they intersect and the angle of separation between them at their point of intersection is a right angle. For vectors, if the angle of separation is a right angle, the cosine vanishes, and so does the dot product. In fact, the dot product is 0 if and only if either the vectors are perpendicular or at least one of the vectors is the zero vector. We can define orthogonality for vectors then by the vanishing of the dot product. The convention that $\mathbf{0}$ is orthogonal to all vectors allows us to succinctly say "two vectors have vanishing dot product if and only if they are orthogonal". When vectors $\mathbf{u}$ and $\mathbf{v}$ are perpendicular, we sometimes write $\mathbf{u} \perp \mathbf{v}$.

On the other hand, vectors are said to be *parallel* if and only if one is a scalar multiple of the other. In the parallel case, it is clear that the dot product is either the product of the two vectors' magnitudes, if they have the same direction, or the negative of the product of their magnitudes, if their directions are opposite. For parallel vectors $\mathbf{u}$ and $\mathbf{v}$, we sometimes write $\mathbf{u} \parallel \mathbf{v}$.

Example 2.1. Show that the vectors $\mathbf{u} = \hat{\mathbf{\imath}} - 8\hat{\mathbf{\jmath}} + 10\hat{\mathbf{k}}$ and $\mathbf{v} = -2\hat{\mathbf{\imath}} + \hat{\mathbf{\jmath}} + \hat{\mathbf{k}}$ are perpendicular. Show that $\mathbf{w} = -4\hat{\mathbf{\imath}} + 2\hat{\mathbf{\jmath}} - 2\hat{\mathbf{k}}$ is not parallel to $\mathbf{v}$.

Solution: To show that $\mathbf{u} \perp \mathbf{v}$, we compute the dot product:

$$\mathbf{u} \cdot \mathbf{v} = (\hat{\mathbf{\imath}} - 8\hat{\mathbf{\jmath}} + 10\hat{\mathbf{k}}) \cdot (-2\hat{\mathbf{\imath}} + \hat{\mathbf{\jmath}} + \hat{\mathbf{k}}) = (1)(-2) + (-8)(1) + (10)(1) = -2 - 8 + 10 = 0 \,.$$

To show that $\mathbf{v}$ and $\mathbf{w}$ are not parallel, we can either check that one is not a multiple of the other, or compute the angle between them from the dot product. For the first approach, one notes that if they were parallel, there would be some $\lambda \in \mathbb{R}$ such that $\mathbf{w} = \lambda\mathbf{v}$. Then

$$-4\hat{\mathbf{\imath}} + 2\hat{\mathbf{\jmath}} - 2\hat{\mathbf{k}} = \lambda(-2\hat{\mathbf{\imath}} + \hat{\mathbf{\jmath}} + \hat{\mathbf{k}}) = -2\lambda\hat{\mathbf{\imath}} + \lambda\hat{\mathbf{\jmath}} + \lambda\hat{\mathbf{k}} \,.$$

Then from the $\hat{\mathbf{\imath}}$-components, we see that $-4 = -2\lambda \implies \lambda = 2$. for the $\hat{\mathbf{\jmath}}$ component, this works, but the $\hat{\mathbf{k}}$-component equation is $-2 = \lambda$. Since λ cannot be both 2 and -2, there can be no λ such that $\mathbf{w} = \lambda\mathbf{v}$.

Instead we can calculate the cosine of the angle of separation, which yields

$$\cos\theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\,\|\mathbf{w}\|} = \frac{8 + 2 - 2}{\sqrt{4+1+1}\sqrt{16+4+4}} = \frac{8}{(\sqrt{6})(2\sqrt{6})} = \frac{2}{3} \,.$$

If $\mathbf{v} \parallel \mathbf{w}$, then the cosine must be ± 1, but instead we get $2/3$, hence they are not parallel.

---

§ 2.3. Algebraic Properties

As with vector arithmetic, one can write down a litany of identities and properties satisfied by the dot product. Arguing the truth of these identities from both the geometric and coordinate definitions proves to be a most useful exercise. We list here the essential ones.

Proposition. *Given vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *in* $\mathscr{V}^n_{\mathbb{R}}$ *and any scalar* $s \in \mathbb{R}$*, the following properties hold:*

(1) $\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2$ *(norm compatibility)*
(2) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ *(commutativity)*
(3) $(s\mathbf{u}) \cdot \mathbf{v} = s(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (s\mathbf{u})$ *(scalar/vector association)*
(4) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ *(distributivity)*
(5) $\mathbf{0} \cdot \mathbf{u} = 0$ *(universal orthogonality of zero vector)*
(6) $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$ *(Cauchy-Schwartz inequality)*

Using the symmetry coming from (2), properties (3) and (4) can be combined into a single identity as

$$\mathbf{u} \cdot (s\mathbf{v} + \mathbf{w}) = s(\mathbf{u} \cdot \mathbf{v}) + \mathbf{u} \cdot \mathbf{w}\,.$$

Generally an operator Ł on vectors satisfying

$$\text{Ł}(\mathbf{u}, s\mathbf{v} + \mathbf{w}) = s\text{Ł}(\mathbf{u}, \mathbf{v}) + \text{Ł}(\mathbf{u}, \mathbf{w})\,, \text{ and}$$
$$\text{Ł}(t\mathbf{u} + \mathbf{v}, \mathbf{w}) = t\text{Ł}(\mathbf{u}, \mathbf{w}) + \text{Ł}(\mathbf{v}, \mathbf{w})$$

is said to be *bilinear*. From the above properties it is clear that the dot product is a bilinear map from pairs of real vectors to real scalars.

§ 2.4. Simple applications of the dot product

Returning to the subject of projections, and using that $\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2$, we can now state the following general formula for projections:

Proposition. *Given two vectors* $\mathbf{u}, \mathbf{v} \in \mathscr{V}^n_{\mathbb{R}}$*, the orthogonal projection of* $\mathbf{v}$ *onto* $\mathbf{u}$*, also called the* component vector of $\mathbf{v}$ parallel to $\mathbf{u}$ *is given by the formula*

$$\operatorname{proj}_{\mathbf{u}} \mathbf{v} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} = (\hat{\mathbf{u}} \cdot \mathbf{v})\hat{\mathbf{u}} = \operatorname{proj}_{\hat{\mathbf{u}}} \mathbf{v}\,,$$

where $\hat{\mathbf{u}} = \mathbf{u}/\sqrt{\mathbf{u} \cdot \mathbf{u}} = \mathbf{u}/\|\mathbf{u}\|$. *The signed length of the projection, called the* component of $\mathbf{v}$ in the direction of $\mathbf{u}$, *is given by*

$$\operatorname{comp}_{\mathbf{u}} \mathbf{v} = \frac{\mathbf{u} \cdot \mathbf{v}}{\|u\|} = \hat{\mathbf{u}} \cdot \mathbf{v}\,.$$

The component vector of $\mathbf{v}$ *orthogonal* to $\mathbf{u}$ is then given by $\mathbf{v} - \operatorname{proj}_{\mathbf{u}} \mathbf{v}$. That this is perpendicular is easily verified with the dot product:

$$\begin{aligned}
\mathbf{u} \cdot (\mathbf{v} - \operatorname{proj}_{\mathbf{u}} \mathbf{v}) &= \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \operatorname{proj}_{\mathbf{u}} \mathbf{v} \\
&= \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}\right) \\
&= \mathbf{u} \cdot \mathbf{v} - \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right) \mathbf{u} \cdot \mathbf{u} \\
&= 0\,.
\end{aligned}$$

Example 2.2. We will decompose the vector $\mathbf{v} = \hat{\imath} + 4\hat{\jmath} + 8\hat{\mathbf{k}}$ into components parallel and perpendicular to the vector $\mathbf{u} = 2\hat{\imath} - 4\hat{\jmath} + 4\hat{\mathbf{k}}$. That is, we construct vectors $\mathbf{v}_{\parallel\mathbf{u}}$ and $\mathbf{v}_{\perp\mathbf{u}}$ such that $\mathbf{v}_{\perp\mathbf{u}} \cdot \mathbf{u} = 0$ and $\mathbf{v}_{\parallel\mathbf{u}} = \mathrm{proj}_{\mathbf{u}}(\mathbf{v})$. Following the preceding discussion, we thus compute

$$\begin{aligned}\mathbf{v}_{\parallel\mathbf{u}} &= \mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u} \\ &= \frac{2-16+32}{4+16+16}(2\hat{\imath} - 4\hat{\jmath} + 4\hat{\mathbf{k}}) = \frac{18}{36}(2\hat{\imath} - 4\hat{\jmath} + 4\hat{\mathbf{k}}) = (\hat{\imath} - 2\hat{\jmath} + 2\hat{\mathbf{k}})\end{aligned}$$

$$\begin{aligned}\mathbf{v}_{\perp\mathbf{u}} &= \mathbf{v} - \mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \hat{\imath} + 4\hat{\jmath} + 8\hat{\mathbf{k}} - (\hat{\imath} - 2\hat{\jmath} + 2\hat{\mathbf{k}}) \\ &= 6\hat{\jmath} + 6\hat{\mathbf{k}}\,.\end{aligned}$$

Observe, we can easily check that $\mathbf{v}_{\parallel\mathbf{u}} \cdot \mathbf{v}_{\perp\mathbf{u}} = (\hat{\imath} - 2\hat{\jmath} + 2\hat{\mathbf{k}}) \cdot (6\hat{\jmath} + 6\hat{\mathbf{k}}) = 0$, and

$$\mathbf{v} = \hat{\imath} + 4\hat{\jmath} + 8\hat{\mathbf{k}} = \underbrace{(\hat{\imath} - 2\hat{\jmath} + 2\hat{\mathbf{k}})}_{\mathbf{v}_{\parallel\mathbf{u}}} + \underbrace{(6\hat{\imath} + 6\hat{\jmath})}_{\mathbf{v}_{\perp\mathbf{u}}}\,.$$

Example 2.3. The projections of a vector $\mathbf{u} \in \mathscr{V}^3_{\mathbb{R}}$ onto $\hat{\imath}$, $\hat{\jmath}$ and $\hat{\mathbf{k}}$ respectively are just the coordinate components:

$$\mathbf{u}\cdot\hat{\imath} = u_1, \qquad \mathbf{u}\cdot\hat{\jmath} = u_2, \qquad \mathbf{u}\cdot\hat{\mathbf{k}} = u_3\,,$$

$$\mathbf{u} = u_1\hat{\imath} + u_2\hat{\jmath} + u_3\hat{\mathbf{k}} = (\mathbf{u}\cdot\hat{\imath})\hat{\imath} + (\mathbf{u}\cdot\hat{\jmath})\hat{\jmath} + (\mathbf{u}\cdot\hat{\mathbf{k}})\hat{\mathbf{k}} = \mathrm{proj}_{\hat{\imath}}\,\mathbf{u} + \mathrm{proj}_{\hat{\jmath}}\,\mathbf{u} + \mathrm{proj}_{\hat{\mathbf{k}}}\,\mathbf{u}\,.$$

Using the geometric interpretation of the dot product, we can compute the angles $\mathbf{u}$ makes with each axis. Let α be the angle from the positive x axis to $\mathbf{u}$, β the angle from the positive y axis to $\mathbf{u}$, and γ the angle from the positive z axis to $\mathbf{u}$. These angles, called *direction angles* satisfy

$$\cos\alpha = \frac{u_1}{\|\mathbf{u}\|} \qquad \cos\beta = \frac{u_2}{\|\mathbf{u}\|} \qquad \cos\gamma = \frac{u_3}{\|\mathbf{u}\|}\,.$$

Thus note that

$$\mathbf{u} = \|\mathbf{u}\|\Big(\cos(\alpha)\,\hat{\imath} + \cos(\beta)\,\hat{\jmath} + \cos(\gamma)\,\hat{\mathbf{k}}\Big)\,.$$

In particular, any unit vector in $\mathbb{R}^3$ has components given by the three *direction cosines*!

We can use the dot product to find vectors perpendicular to a given vector $\mathbf{u} \in \mathscr{V}^2_{\mathbb{R}}$. If we write $\mathbf{u} = \langle u_1, u_2\rangle$, then we seek some vector $\mathbf{v} = \langle v_1, v_2\rangle$ such that $\mathbf{u}\cdot\mathbf{v} = 0$. Writing out the coordinate expression for the product yields

$$u_1v_1 + u_2v_2 = 0 \implies u_1v_1 = -u_2v_2\,.$$

Since we just want to find *some vector* perpendicular (after which we can scale to get others of different lengths or in the opposite direction), we can make a choice: let's set $v_1 = u_2$. We'll see that the result of our choice is we can find a vector $\mathbf{v}$ perpendicular to $\mathbf{u}$ with the same magnitude as $\mathbf{u}$. Indeed, our choice $v_1 = u_2$ then requires that v_2 equal $-u_1$ to satisfy the condition imposed by the vanishing dot product. We can now easily check that $\mathbf{v} = \langle u_2, -u_1\rangle$ is perpendicular to $\mathbf{u} = \langle u_1, u_2\rangle$, and the magnitudes are clearly the same since $\|\mathbf{v}\|^2 = v_1^2 + v_2^2 = u_2^2 + (-u_1)^2 = \|\mathbf{u}\|^2$.

If you consider the *slopes* of the above vectors (the ratio of their y-components to their x-components) then it should come as no surprise that $\mathbf{v}$ gives a vector perpendicular to $\mathbf{u}$; the slope of $\mathbf{v}$ is $-\frac{u_1}{u_2}$, which is the negative reciprocal of the slope $\frac{u_2}{u_1}$ of $\mathbf{u}$! Observe also that we could just as well have taken $-\mathbf{v}$ to obtain another $\mathbf{u}$-orthogonal vector of equal magnitude.

Example 2.4. The points $P(-2,3)$ and $Q(4,-5)$ are opposite vertices of a square, meaning that the line segment from P to Q is a diagonal of the square. Find the other two vertices of the square.

Solution:

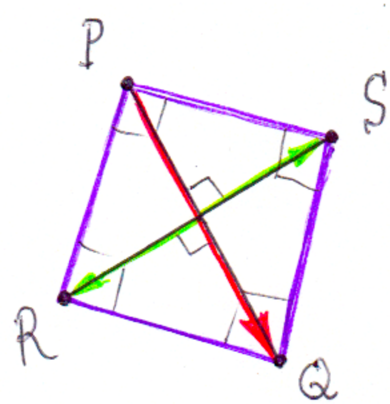


FIGURE 10. The square with known vertices P and Q, and unknown vertices R and S.

Consider figure (10), which illustrates a square like the one described, where R and S are the unknown vertices. Note that the two diagonals of the square have the same length and bisect each other perpendicularly. By finding a vector perpendicular to the diagonal $\overrightarrow{PQ}$, we get a vector parallel to $\overrightarrow{RS}$, which describes the displacement along the opposite diagonal. By finding position vector $\mathbf{m}$ of the midpoint M between P and Q, and adding one half $\overrightarrow{RS}$, we can get the position vector $\mathbf{s}$ for S, and similarly subtracting one half $\overrightarrow{RS}$ from the midpoint position vector gives the position vector $\mathbf{r}$ for R. Therefore we have

$$\begin{aligned}\mathbf{m} &= \mathbf{p} + \frac{1}{2}\overrightarrow{PQ} = \mathbf{p} + \frac{1}{2}(\mathbf{q}-\mathbf{p}) = \frac{\mathbf{p}+\mathbf{q}}{2} \\ &= \frac{1}{2}\langle 2,-2\rangle = \langle 1,-1\rangle\,.\end{aligned}$$

Note we computed the midpoint without writing down $\overrightarrow{PQ}$, but we could just as well have computed it using the first equality and the fact that $\overrightarrow{PQ} = \langle 6,-8\rangle$, giving the equivalent computation

$$\mathbf{m} = \langle -2,3\rangle + \frac{1}{2}\langle 6,-8\rangle = \langle -2,3\rangle + \langle 3,-4\rangle = \langle 1,-1\rangle\,.$$

Now, from the discussion preceding this example, we know that

$$\overrightarrow{RS} = \langle 8,6\rangle\,,$$

and so

$$\begin{aligned}\mathbf{s} &= \mathbf{m} + \frac{1}{2}\overrightarrow{RS} = \langle 1,-1\rangle + \langle 4,3\rangle = \langle 5,2\rangle\,,\text{and} \\ \mathbf{r} &= \mathbf{m} - \frac{1}{2}\overrightarrow{RS} = \langle 1,-1\rangle - \langle 4,3\rangle = \langle -3,-4\rangle\,.\end{aligned}$$

Thus the other vertices of the square are $S(5,2)$ and $R(-3,-4)$.

Projections and dot products are commonly applied in physics problems such as using Newton's laws to calculate forces in static and dynamic problems. By a velocity vector, we shall mean a vector describing the motion of an object giving the direction of its movement at that instant, with magnitude equal to the speed of the object. Velocity is the derivative of the position vector. The derivative of velocity is acceleration, and describes both changes in direction and changes in speed. Newton's second law of motion states that force is proportional to acceleration via mass: if $\mathbf{F}$ is a force vector applied to an object of mass m, and the object accelerates due to this force with acceleration vector $\mathbf{a}$[7], then

$$\mathbf{F} = m\mathbf{a}\,.$$

This assumes no other forces are acting on the object; in nature we do not encounter such circumstances. The correct way to understand this is as follows. Since force is a vector, we can really view $\mathbf{F}$ as representing the vector sum of all forces acting on the object within a particular *inertial frame of reference*[8]. By this we mean there is a coordinate system within which we measure the forces acting on objects, and if they sum to $\mathbf{0}$ in this coordinate system, there is no net acceleration of the objects to which the forces were applied. Some authors clarify this by writing Newton's second law as

$$\Sigma\mathbf{F} = m\mathbf{a}\,,$$

where the Σ reminds us that we are examining the *net force*, which is a sum of various forces acting on the object. The acceleration vector $\mathbf{a}$ is taken to mean the acceleration relative to the inertial coordinate system.

In the simplest applications, one considers the case where an object is at rest or undergoing linear motion at a constant speed within an inertial frame of reference, and thus the acceleration is $\mathbf{0}$, and the forces must balance out in all directions. This is called a *static system.*[9] We'll consider an example of a simple static system that demonstrates the use of projections.

Example 2.5. A box is sitting on a ramp, and does not slide down the ramp. The ramp is inclined at $25.0°$ relative to the horizontal. If the box has a mass of 12.0 kilograms, what is the acting force of friction keeping the box from sliding? Assume that gravity exerts a force equal to the mass of the box times $g = 9.80665\text{m/s}^2$ in a direction perpendicular to the horizontal.

Solution: The first step is to understand what forces act on the box, and in what directions. There is the downward force exerted by gravity, there is the *normal force* exerted by the ramp perpendicular to the incline, and there is a force of friction parallel to the plane of incline. The normal force must exist else the box would accelerate into the ramp, while the force of friction is what prevents the box from sliding along the ramp. To understand these, we decompose the gravitational force into components parallel and perpendicular to the ramp, and apply Newton's second law to the sum of forces, with a net acceleration of $\mathbf{0}$. A visual aid helps keep track of this information, as in figure 11 below.

[7]Recall, acceleration is the derivative of velocity, which is the derivative of position. Thus, for a particle with position $\mathbf{r}(t)$ at time t, the acceleration is $\mathbf{a}(t) = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}$. Here, the position $\mathbf{r}(t)$, the velocity $\mathbf{v}(t) = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$, and the acceleration $\mathbf{a}(t)$ are each *vector valued functions*. The next set of notes concerns vector valued functions and their derivatives, and explores velocity and acceleration in greater depth.

[8]The name *inertial reference frame* is itself a reference to Newton's first law of motion: the law of inertia. The modern statement is that an object with no net force acting upon it will experience no net acceleration in such a reference frame. The term inertia describes the tendency to resist change (in this case, in velocity): as it is often quoted "an object at rest will remain at rest until an outside force acts upon it, and an object in motion will stay in motion until an outside force acts upon it."

[9]Though the word "static" conjures an image of motionlessness, our definition permits an object moving at constant velocity. Why then call it "static"? Well, if the object has velocity $\mathbf{v}$, we can choose a new inertial reference system that moves along with the object at constant velocity $\mathbf{v}$ relative to the old reference frame. Then in this frame of reference, the object *is motionless*, i.e., *static*!

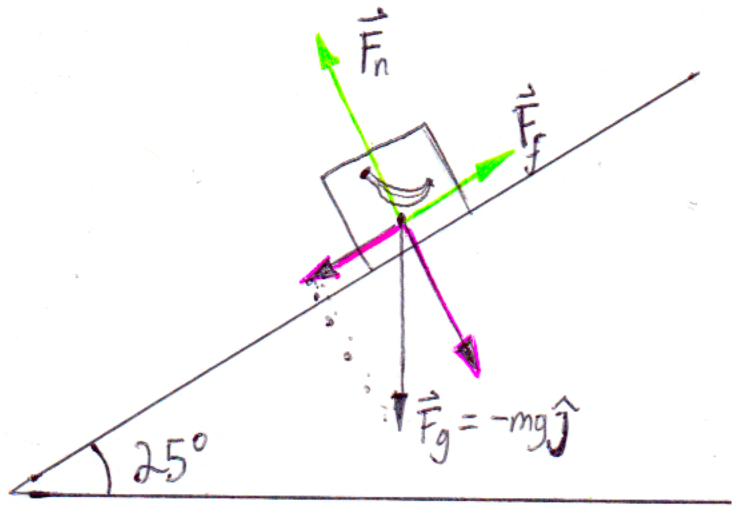


FIGURE 11. A static box on a ramp. The component vectors of gravity along and into the ramp must be met by equal and opposite forces: the normal force $\mathbf{F}_n$ and the force of friction $\mathbf{F}_f$.

First, the force of gravity is $\mathbf{F}_g = -mg\,\hat{\mathbf{j}} = -(12.0\,\text{kg})(9.80665\,\text{m/s}^2)\,\hat{\mathbf{j}} \approx 118\,\text{N}\,\hat{\mathbf{j}}$, where N is the abbreviation for the Newton force unit. One Newton is the force required to accelerate a one kilogram object at one meter per second per second.

Next, we can decompose this force into directions parallel and perpendicular to the ramp. The ramp incline is 25°, and so the plane of the ramp in our picture is parallel to the vector $\hat{\mathbf{p}} = \cos(25°)\,\hat{\imath} + \sin(25°)\,\hat{\jmath}$. Let $\hat{\mathbf{n}}$ denote the normal unit vector to the ramp. Convince yourself that this vector has component form $\hat{\mathbf{n}} = -\sin(25°)\,\hat{\imath} + \cos(25°)\,\hat{\jmath}$. From the similar triangle in the figure above, one can see that the component of $\mathbf{F}_g$ parallel to the ramp is $-\|\mathbf{F}_g\|\sin(25°)\,\hat{\mathbf{p}} = 49.7\,\text{N}\,\hat{\mathbf{p}}$, while the component perpendicular to the ramp is $\|\mathbf{F}_g\|\cos(25°)\,\hat{\mathbf{n}} \approx 107\,\text{N}\,\hat{\mathbf{n}}$. We can also realize this by applying projection operators:

$$\begin{aligned}
\text{proj}_{\hat{\mathbf{p}}}(\mathbf{F}_g) &= (\hat{\mathbf{p}}\cdot\mathbf{F}_g)\hat{\mathbf{p}} \\
&= \Big(\cos(25°)\,\hat{\imath} + \sin(25°)\,\hat{\jmath}\Big)\cdot(-mg\,\hat{\jmath})\hat{\mathbf{p}} \\
&= -mg\,\sin(25°)\,\hat{\mathbf{p}} \approx -49.7\,\text{N}\,\hat{\mathbf{p}}\,, \\
\text{proj}_{\hat{\mathbf{n}}}(\mathbf{F}_g) &= (\hat{\mathbf{n}}\cdot\mathbf{F}_g)\hat{\mathbf{n}} \\
&= \Big(-\sin(25°)\,\hat{\imath} + \cos(25°)\,\hat{\jmath}\Big)\cdot(-mg\,\hat{\jmath})\hat{\mathbf{n}} \\
&= -mg\,\cos(25°)\,\hat{\mathbf{u}} \approx -107\,\text{N}\,\hat{\mathbf{n}}\,.
\end{aligned}$$

Let $\mathbf{F}_f$ be the force of friction, and $\mathbf{F}_n$ the normal force of the ramp on the box. Putting everything together and applying Newton's second law:

$$\begin{aligned}
m\mathbf{a} = \mathbf{0} = \Sigma\mathbf{F} = \mathbf{F}_g + \mathbf{F}_f + \mathbf{F}_n& \\
\implies \text{proj}_{\hat{\mathbf{p}}}(\mathbf{F}_g) + \text{proj}_{\hat{\mathbf{n}}}(\mathbf{F}_g) + \mathbf{F}_f + \mathbf{F}_n = \mathbf{0}&
\end{aligned}$$

The force of friction, $\mathbf{F}_f$ is parallel to $\hat{\mathbf{p}}$, while the normal force $\mathbf{F}_n$ is parallel to the ramp's normal vector $\hat{\mathbf{n}}$. Thus, we can equate components in these directions to deduce

$$\mathbf{F}_n = -\text{proj}_{\hat{\mathbf{n}}}(\mathbf{F}_g) \approx 107\,\text{N}\,\hat{\mathbf{n}}$$

$$\mathbf{F}_f = -\mathrm{proj}_{\hat{\mathbf{p}}}(\mathbf{F}_g) \approx 49.7\,\mathrm{N}\,\hat{\mathbf{p}}\,.$$

In particular, the effective force of friction on the box is about 49.7 Newtons upward along the ramp.

Another physics context in which the dot product plays an important role is *work*. Work is commonly described as *force times displacement*, but when this definition is given, it is tacit that both force and displacement are vectors, and somehow the work you measure is *how much force works to produce how much linear displacement*[10]. In particular, if you apply lots of force and the object moves at an angle to the direction in which you applied the force, the work will only account for the *component* of effective force, which is parallel to the object's motion. Thus, a careful definition of work inn the setting of constant applied force is as the *dot product of force and displacement*, which equals the real product of the effective force in the direction of the displacement it causes. So, if you apply a force $\mathbf{F}$ to an object, displacing it by a vector $\mathbf{s}$, then the work is the scalar

$$W = \mathbf{F}\cdot\mathbf{s}\,.$$

If the force varies across the displacement, then one needs to account for the change of the force along the object's trajectory, and so naturally some calculus becomes necessary. In particular, one must consider *force fields along curved trajectories*, and integrate to find *net work*. The corresponding mathematical objects needed are "vector-valued functions" $\mathbf{r}:\mathbb{R}\to\mathbb{R}^3$ to describe the trajectory[11], and "vector fields" $\mathbf{F}:\mathscr{V}^3_{\mathbb{R}}\to\mathbb{R}^3$ to describe the force at different points of space. Then one defines work via a gadget called a *line integral*. This idea is explored in the final unit of this course (see the later notes on line integrals).

Example 2.6. A force of 100 Newtons is applied to a box on a floor. The force is applied at an angle, so that a component of the force is upwards, with $\mathbf{F}$ making an angle of $32°$ with the horizontal. Find the work done on the box if it is dragged 6.5 meters, in Joules (1 Joule is the work to move an object 1 meter with an effective applied force of 1 Newton, i.e. $1\,\mathrm{J} = 1\,\mathrm{N}\,\times 1\,\mathrm{m}$.)

Solution: Applying the geometric interpretation of the dot product to compute work from its definition:

$$W = \mathbf{F}\cdot\mathbf{s} = \|\mathbf{F}\|\,\|\mathbf{s}\|\cos 32° = (100\,\mathrm{N})(6.5\,\mathrm{m})(0.848\ldots) \approx 550\,\mathrm{N\,m} = 550\,\mathrm{J}\,.$$

Equivalently, by computing the projection of $\mathbf{F}$ onto $\mathbf{s}$ and then multiplying by $\|\mathbf{s}\|$, we work directly with the effective force coming from the component parallel to the displacement, and recover the same calculation.

---

§ 2.5. Problems

(1) (Do this whole problem only if you want *lots* of practice computing)
Compute all possible dot products involving the vectors $\mathbf{a} = \langle 1,1,1\rangle$, $\mathbf{b} = \langle -2,3,-3\rangle$, $\mathbf{c} = \langle -1,4,7\rangle$, $\mathbf{u} = \langle 1/2,-1/3,1/6\rangle$, $\mathbf{v} = \langle 0,9,9\rangle$, and $\mathbf{w} = \langle -7,6,0\rangle$, and compute all possible projections and the complimentary orthogonal component vectors.

[11]Here, we can identify $\mathbb{R}^3$ with $\mathscr{V}^3_{\mathbb{R}}$ by considering each point along the trajectory as being determined by a position vector; in practice we describe such functions parametrically as a triple or real valued functions $\mathbf{r}(t) = x(t)\,\hat{\mathbf{\imath}} + y(t)\,\hat{\mathbf{\jmath}} + z(t)\,\hat{\mathbf{k}}$.

(2) Give a concrete example showing that the dot product does not have the cancellation property, i.e. show that $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\cdot\mathbf{w}$ does not imply that $\mathbf{v} = \mathbf{w}$. For a fixed pair $\mathbf{u}, \mathbf{v} \in \mathscr{V}^3_{\mathbb{R}}$, describe geometrically the set of all vectors $\mathbf{w} \in \mathscr{V}^3_{\mathbb{R}}$ such that $\mathbf{u}\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{v}$.

(3) For real numbers, it is well known that multiplication satisfies an associative property: $a(bc) = abc = a(bc)$ for any $a, b, c \in \mathbb{R}$. Why is there no associative property of the dot product? What's wrong with writing $\mathbf{a}\cdot(\mathbf{b}\cdot\mathbf{c}) = \mathbf{a}\cdot\mathbf{b}\cdot\mathbf{c} = (\mathbf{a}\cdot\mathbf{b})\cdot\mathbf{c}$?

(4) By a diagonal of a cube, we mean the line segment from one vertex of a cube to the farthest vertex across the cube. By a diagonal of a cube's face, we mean the diagonal of the square face from one vertex to the opposite vertex of that face.

(a) Find the lengths of the diagonals of a cube and diagonals of faces in terms of the side length of a cube.
(b) Find the angles between a diagonal of a cube and an adjacent edge of the cube.
(c) Each diagonal of the cube is adjacent to how many face diagonals? Find the angle between a diagonal of a cube and an adjacent face diagonal.

(5) For $\mathbf{u} = u_1\hat{\mathbf{\imath}} + u_2\hat{\mathbf{\jmath}}$ and $\mathbf{v} = v_1\hat{\mathbf{\imath}} + v_2\hat{\mathbf{\jmath}}$, write the magnitude of the resultant $\mathbf{u}+\mathbf{v}$ in terms of the components of $\mathbf{u}$ and $\mathbf{v}$, and then give formulae for the angles made between the resultant and each of $\hat{\mathbf{\imath}}$, $\hat{\mathbf{\jmath}}$, $\mathbf{u}$ and $\mathbf{v}$.

(6) A 10.5 kilogram box is sliding down a ramp, which is inclined at 28.2° relative to the horizontal. Suppose the force of friction of the ramp on the sliding box has magnitude proportional to the normal force exerted by the ramp on the box, with constant of proportionality $\mu = 0.155$ (such a μ is called the *coefficient of friction*). Assuming, as in the example above, that gravity exerts a downward force equal to the mass of the box times $g = 9.80665\text{m/s}^2$, find the acceleration of the box along the ramp. Your answer should be rounded so as to contain 3 significant figures. You may of course neglect things like air resistance that would complicate the assumption of constant acceleration.

(7) A 5.25 kilogram sculpture is suspended from a gallery ceiling by three wires which meet at a single hook on the sculpture. If the origin is placed at the hook, then one wire is parallel to $1.0\hat{\mathbf{\imath}} + 1.6\hat{\mathbf{\jmath}} + 1.0\hat{\mathbf{k}}$, one wire is parallel to $2.0\hat{\mathbf{\imath}} - 3.0\hat{\mathbf{\jmath}} + 2.0\hat{\mathbf{k}}$ and the last wire is parallel to $-3.0\hat{\mathbf{\imath}} + 0.25\hat{\mathbf{\jmath}} + 3.0\hat{\mathbf{k}}$. If the sculpture hangs 1.57 meters below the ceiling of the gallery, find the lengths of each of the wires. Assuming the sculpture is non-kinetic, find the tension in each wire.

(8) Three exuberant dogs are recreationally pulling a sled around in a snow-covered field, but they are not quite coordinated. They exert forces along 3 different directions, but manage to pull the sled in a straight line briefly before it begins to veer about wildly. Suppose that during the linear motion, one dog exerts a constant force of 165 Newtons of force in the direction of the vector $-1.00\hat{\mathbf{\imath}} + 1.80\hat{\mathbf{\jmath}} + 0.60\hat{\mathbf{k}}$, one dog exerts a constant force of 184 Newtons parallel to the vector $2.00\hat{\mathbf{\imath}} + 2.20\hat{\mathbf{\jmath}} + 0.50\hat{\mathbf{k}}$, and the final dog exerts a constant force of 240 Newtons along the vector $2.40\hat{\mathbf{\imath}} + 0.70\hat{\mathbf{k}}$.

(a) Assuming the sled does not lift off of the snow, and the only forces acting upon it are those due to the dogs, gravity and friction, determine a unit vector in the direction of the sled's displacement.

(b) Compute the work each dog does if the sled is displaced 8.00 meters along the direction described in part (a).
(c) Compute the total work done by the dogs on the sled.
(d) (Challenge) If the sled is packed to weigh 65.0 kilograms, the coefficient of friction between the sled blades and the snow is $\mu = 0.007$, and the dogs accelerate so that the forces described above are the tensions maintained in the ropes during the linear motion of the sled, determine the equations of the sled's motion for the straight line displacement. How many seconds does it take to complete the 8.00 meter run? You may assume $g = 9.80665\text{m/s}^2$, and use 3 significant figures for your final answer.

(9) Let $a, b, c \in \mathbb{R}_+$ be positive real numbers giving the lengths of the sides of a triangle. Suppose the sides of length a and b are separated by an interior angle of θ. Then the Law of cosines states that

$$c^2 = a^2 + b^2 - 2ab\cos\theta\,.$$

Prove the Law of cosines using vector algebra. (Hint: describe the sides using vectors, and compute the square of c geometrically by a dot product.)

(10) Prove that, for any $\mathbf{u}, \mathbf{v} \in \mathscr{V}_{\mathbb{R}}^n$,

$$2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2 = \|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2\ ,\ \text{and}$$

$$\mathbf{u}\cdot\mathbf{v} = \frac{1}{4}\Big(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2\Big)$$

(11) Consider linearly independent vectors $\mathbf{u}, \mathbf{v} \in \mathscr{V}_{\mathbb{R}}^2$, and let $\mathscr{P}$ be the parallelogram whose sides they span. Under what conditions are the diagonals of $\mathscr{P}$ orthogonal?

(12) Demonstrate via vector algebra that the diagonals of a parallelogram always bisect each other.

(13) For a line through the origin in the direction of $\mathbf{u} \in \mathscr{V}_{\mathbb{R}}^2$, use the projection operator to give an expression for the reflection of a vector $\mathbf{x}$ through the line.

(14) Prove the Cauchy-Schwartz inequality for the dot product (property (6) in the proposition above).

(15) Describe a procedure, using the dot product, to generate from a list of vectors $\mathbf{v}_1, \dots \mathbf{v}_n \in \mathscr{V}_{\mathbb{R}}^n$, a new list $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_n \in \mathscr{V}_{\mathbb{R}}^n$ such that

$$\|\hat{\mathbf{u}}_i\| = 1\,, i = 1, \dots n\,,$$

$$\hat{\mathbf{u}}_i \cdot \hat{\mathbf{u}}_j = 0 \text{ if } i \neq j\,.$$

Such a set is called an *orthonormal basis*. The vectors $\hat{\mathbf{\imath}}$, $\hat{\mathbf{\jmath}}$, and $\hat{\mathbf{k}}$ form an orthonormal basis of $\mathscr{V}_{\mathbb{R}}^3$.

(16) Use the procedure of the preceding problem to generate an orthonormal basis from the vectors $\langle 4, -4, 7\rangle$, $\langle 1, 1, 1\rangle$, and $\langle -3, 4, -5\rangle$.

(17) Using the algebraic properties of the dot product and the geometric definition, prove the coordinate expression

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

Generalize your proof to the expression

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \ldots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

by using an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for $\mathscr{V}_{\mathbb{R}}^n$.

3. Cross Product

While we have one notion of "product" for vectors, it doesn't really meet the criteria that the average mathematician would ideally desire in a product operation. Namely, the dot product is a map from a pair of vectors to a *scalar*. But one might wonder if there is a map with useful properties taking a pair of vectors to a vector in the same vector space. By a stroke of mathematical fortune, in three dimensions such a product exists and its properties make it useful in the study of lines, planes, areas, volumes, rotational mechanics, and spatial rotations. Its origins, like the $\hat{\mathbf{\imath}}, \hat{\mathbf{\jmath}}, \hat{\mathbf{k}}$ notation for coordinate vectors, harken back to an historical attempt by William Rowan Hamilton to generalize the success of complex number arithmetic in describing planar geometry to a spatial analog via *hypercomplex numbers*. This is discussed in further depth in a later (optional) section on quaternions. Fortunately, now we can approach a definition of the product in various modern ways, all grounded in perspectives that will find use.

§ 3.1. Geometric Definition

We will rely on geometry to give a definition which is *a priori* not particularly helpful in computations, but which will nonetheless give us a map

$$\times : \mathscr{V}^3_{\mathbb{R}} \times \mathscr{V}^3_{\mathbb{R}} \to \mathscr{V}^3_{\mathbb{R}} \,.$$

Given a pair of nonzero vectors $\mathbf{u}, \mathbf{v} \in \mathscr{V}^3_{\mathbb{R}}$, then either they span a plane, or they are parallel (and thus, span only a line). In the event that they span only a line or at least one vector is the zero vector, we will define their product to return the zero vector. On the other hand, if they span a plane, we will determine a vector perpendicular to that plane whose direction will depend on the right hand rule, and whose magnitude encodes geometric information about $\mathbf{u}$ and $\mathbf{v}$. In particular, for $\mathbf{u}$ and $\mathbf{v}$ not parallel, let $\mathscr{P}$ be the parallelogram spanned by them (so if $\mathbf{u}$ and $\mathbf{v}$ are placed emanating from the origin, $\mathscr{P}$ has vertices with position vectors $\mathbf{0}$, $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{u}+\mathbf{v}$). The right hand rule furnishes a unique unit vector $\hat{\mathbf{n}}_{\mathscr{P}}$ normal to this parallelogram (and so, to both $\mathbf{u}$ and $\mathbf{v}$) such that $(\mathbf{u}, \mathbf{v}, \hat{\mathbf{n}}_{\mathscr{P}})$ is a right-handed triple (if we insisted on having a left-handed triple, we'd get $-\hat{\mathbf{n}}_{\mathscr{P}}$). Note that we could get an expression in coordinates for $\hat{\mathbf{n}}_{\mathscr{P}}$ by writing down coordinate equations from the condition $\hat{\mathbf{n}}_{\mathscr{P}} \cdot \mathbf{u} = 0 = \hat{\mathbf{n}}_{\mathscr{P}} \cdot \mathbf{v}$, solving up to a constant, then choosing the constant to ensure that the normal vector had unit length and that we had a right-handed triple. Now, let $\mathcal{A}(\mathscr{P})$ be the area of the parallelogram. You should convince yourself (draw a picture!) that this area is equal to $\|\mathbf{u}\|\,\|\mathbf{v}\| \sin\theta$ where $\theta \in (0, \pi)$ is the angle of separation of $\mathbf{u}$ and $\mathbf{v}$.

Definition. For $\mathbf{u}, \mathbf{v} \in \mathscr{V}^3_{\mathbb{R}}$, the cross product is the vector defined by the following conditions:

(1) $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ whenever $\mathbf{u}$ and $\mathbf{v}$ are parallel.

(2) For $\mathbf{u}$ and $\mathbf{v}$ linearly independent, with $\mathscr{P}$, $\hat{\mathbf{n}}_{\mathscr{P}}$, $\mathcal{A}(\mathscr{P})$, and θ as above,

$$\mathbf{u} \times \mathbf{v} := \mathcal{A}(\mathscr{P})\hat{\mathbf{n}}_{\mathscr{P}} = \|\mathbf{u}\|\,\|\mathbf{v}\| \sin(\theta)\, \hat{\mathbf{n}}_{\mathscr{P}} \,.$$

§ 3.2. Physical Definition

This definition is adapted from Bressoud's *Second Year Calculus*. We consider a rigid object, with center of mass at the origin, rotating around an axis. Let $\boldsymbol{\omega}$ denote an *angular velocity vector*, defined as follows: $\boldsymbol{\omega}$ points along the axis of rotation from the center of mass such that, looking back at the center of mass from the tip of $\boldsymbol{\omega}$, one sees the rotation as counter-clockwise, and $\|\boldsymbol{\omega}\|$ is the angular speed of rotation, measured in radians per second. Fix a moment in time. Now, let $\mathbf{s}$ be a position vector locating a point of the object at this moment, and let $\mathbf{v}$ be the *linear velocity*

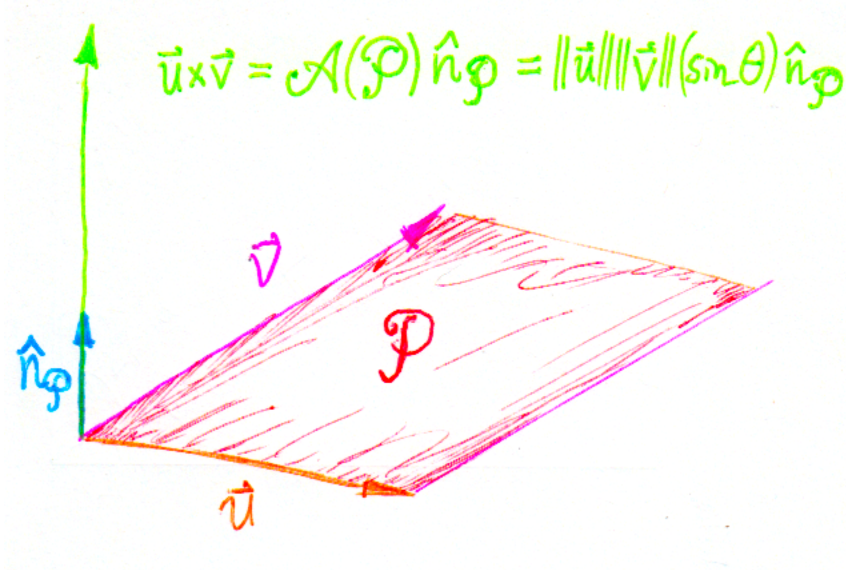


FIGURE 12

vector of this point at this particular instant. Then the cross product is the operation such that

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{s}\,.$$

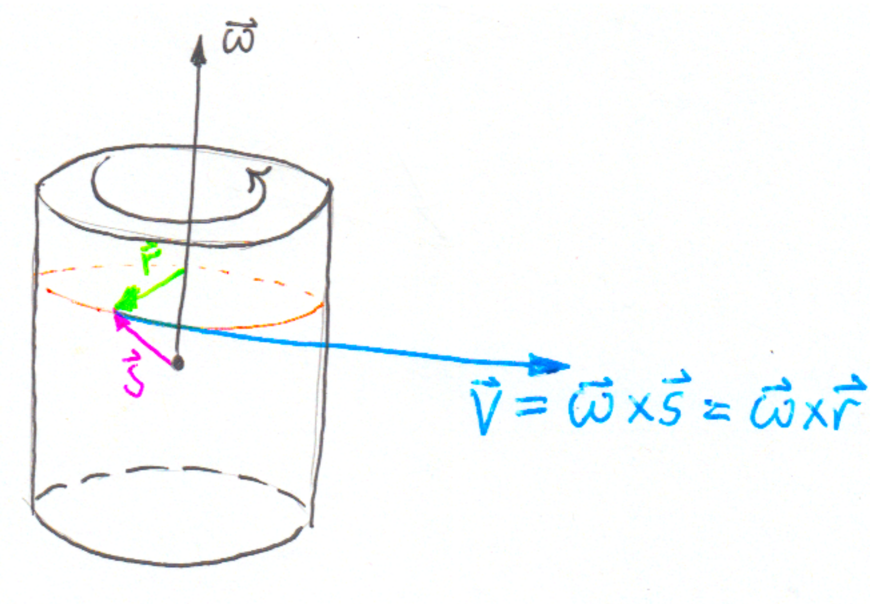


FIGURE 13

Note that it follows from our definition that

$$\|\mathbf{v}\| = \|\boldsymbol{\omega} \times \mathbf{s}\| = \|\boldsymbol{\omega}\|\,\|\mathbf{s}\|\sin\theta\,,$$

where $\theta \in [0, \pi]$ is the angle of separation of $\boldsymbol{\omega}$ and $\mathbf{s}$.

If $\mathbf{r}$ denotes a vector giving the shortest displacement from the axis of rotation to the terminus of $\mathbf{s}$, then the velocity associated is the same, and thus $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{s} = \boldsymbol{\omega} \times \mathbf{r}$. This in particular implies that the cross product is not *cancellative*, i.e. $\boldsymbol{\omega} \times \mathbf{s} = \boldsymbol{\omega} \times \mathbf{r}$ does not imply $\mathbf{s} = \mathbf{r}$. In particular, despite having a vector valued product operation for three dimensional vectors, there is *no* way to define *division* for this product.[12]

In physics, the cross product appears frequently when dealing with rotational phenomena. An elementary example is the definition of torque. Given a lever arm let $\boldsymbol{\ell}$ be a vector from the fulcrum to the end, and consider a force $\mathbf{F}$ applied to the lever arm at the terminal point of $\boldsymbol{\ell}$. Then torque is the vector quantity

$$\boldsymbol{\tau} := \boldsymbol{\ell} \times \mathbf{F}\,.$$

§ 3.3. Algebraic and Geometric properties

A real test of geometric fluency with 3-dimensional vectors is to try and use the above two definitions of the cross product to obtain a list of algebraic properties analogous to those of the dot product. There will however be some important differences. The cross product is *anti-commutative* and *non-associative*. It does however satisfy bilinearity properties like those of the dot product. The proposition below lists the most important algebraic properties of the cross product. See if you can rediscover or prove them by appealing to geometric and physical reasoning.

Proposition. *Given vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *in* $\mathscr{V}_{\mathbb{R}}^3$ *and any scalar* $s \in \mathbb{R}$*, the following properties hold:*

(1) $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$ *(anti-commutativity)*
(2) $(s\mathbf{u}) \times \mathbf{v} = s(\mathbf{u} \times \mathbf{v}) = \mathbf{u} \times (s\mathbf{u})$ *(scalar/vector association)*
(3) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ *(left distributivity)*
(4) $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$ *(right distributivity)*
(5) $\mathbf{u} \times \mathbf{u} = \mathbf{0}$ *(vanishing on co-linear pairs)*
(6) $\mathbf{0} \times \mathbf{u} = \mathbf{0}$ *(universal co-linearity of zero vector)*
(7) $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) + \mathbf{v} \times (\mathbf{w} \times \mathbf{u}) + \mathbf{w} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{0}$ *(Jacobi identity)*

The anti-commutative property is sometimes called *skew-symmetry.* The cross product determines a *skew-symmetric bilinear map* from pairs of 3-dimensional real vectors to 3-dimensional real vectors. One might wonder if there are skew-symmetric bilinear products of vectors in $\mathscr{V}_{\mathbb{R}}^n$ taking values in $\mathscr{V}_{\mathbb{R}}^n$, for n other than 3. The initially startling answer is that these products are rare, occurring only in dimensions $n = 3$ and $n = 7$ (there's of course the usual product of 1-dimensional real vectors, i.e. the usual product on real numbers, but that operation is commutative.)

§ 3.4. Deriving Coordinate Expressions

Given the properties of the cross product listed in the preceding proposition, we can apply our definitions to obtain a computational rule for calculating the cross product in coordinates.

First, we consider the cross products of the coordinate basis vectors $\hat{\mathbf{\imath}}$, $\hat{\mathbf{\jmath}}$ and $\hat{\mathbf{k}}$. Note that their are 9 possible ways to form cross products from these. Property (5) implies that 3 of these, namely $\mathbf{i} \times \mathbf{i}$, $\mathbf{j} \times \mathbf{j}$, $\mathbf{k} \times \mathbf{k}$, are all equal to $\mathbf{0}$. The remaining 6 products form three pairs, related according to property 1, by reversal. For example, since $\hat{\mathbf{\imath}}$ and $\hat{\mathbf{\jmath}}$ span a unit square with $\hat{\mathbf{k}}$ the right-handed unit normal to this square, we have from the geometric definition that

$$\hat{\mathbf{\imath}} \times \hat{\mathbf{\jmath}} = \hat{\mathbf{k}} = -\hat{\mathbf{\jmath}} \times \hat{\mathbf{\imath}}\,.$$

[12]The question of which vector spaces over $\mathbb{R}$ admit the structure of a real normed division algebra is closely connected to the existence of cross products. The only vector spaces admitting such structures occur in dimensions 1, 2, 4, and 8, corresponding to the real numbers, complex numbers, quaternions, and octonions respectively. Cross products, which fit into the scheme discussed below of *skew symmetric bilinear products taking values orthogonal to their arguments*, are produced in dimensions 3 and 7 as a by-product of the normed division algebra structures in dimensions 4 and 8.

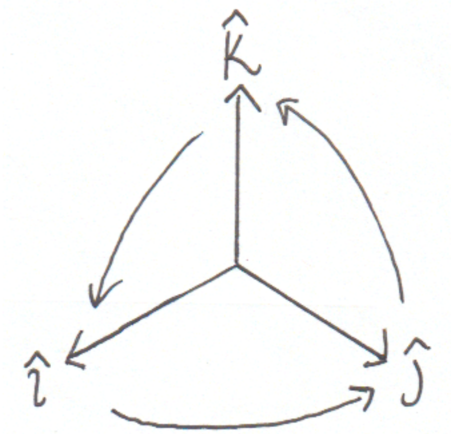


FIGURE 14. A visual mnemonic for remembering the coordinate basis cross products. Products in the order of the arrows return the next vector in the sequence; going against the arrows, one would have to introduce a negative sign to the result.

Similarly, one can deduce that

$$\hat{\mathbf{\jmath}} \times \hat{\mathbf{k}} = \hat{\mathbf{\imath}} = -\hat{\mathbf{k}} \times \hat{\mathbf{\jmath}}\,,$$

$$\hat{\mathbf{k}} \times \hat{\mathbf{\imath}} = \hat{\mathbf{\jmath}} = -\hat{\mathbf{\imath}} \times \hat{\mathbf{k}}\,.$$

If we preferred instead to apply the physical definition, it would suffice to consider 3 unit radius disks, centered at the origin, each lying in one of the coordinate planes. For the disk lying in the xy plane, suppose an angular velocity vector of $\boldsymbol{\omega} = \hat{\mathbf{k}}$, i.e. the disk spins counterclockwise in the xy plane with angular speed of 1 radian per second. Then the linear velocity at the point $(1,0,0)$ on the circumference is given by a tangential vector parallel to $\hat{\mathbf{\jmath}}$. Its magnitude must be the product of the angular speed in radians per second with the radius, $\|\boldsymbol{\omega}\|\,\|\mathbf{r}\|$ where $\mathbf{r}$ is the radial vector to the position $(1,0,0)$. But $\mathbf{r}$ is none other than $\hat{\mathbf{\imath}}$, and so this works out to $\left\|\hat{\mathbf{k}}\right\|\,\|\hat{\mathbf{\imath}}\| = (1)(1) = 1$. Thus $\hat{\mathbf{\jmath}} = \mathbf{v} = \boldsymbol{\omega} \times \mathbf{r} = \hat{\mathbf{k}} \times \hat{\mathbf{\imath}}$. Similarly, we recover the above other products above by appropriately describing the other disks and choosing points along the coordinate axes at their circumferences such that the angular momentum vector and radius give other combinations of $\pm\hat{\mathbf{\imath}}$, $\pm\hat{\mathbf{\jmath}}$ and $\pm\hat{\mathbf{k}}$.

To assemble a full coordinate expression, we now compute the cross product of two arbitrary vectors $\mathbf{u}, \mathbf{v} \in \mathscr{V}^3_{\mathbb{R}}$ by applying the distributive and scalar association properties together with our expressions for the cross products of the coordinate basis vectors. Let

$$\mathbf{u} = u_1\hat{\mathbf{\imath}} + u_2\hat{\mathbf{\jmath}} + u_3\hat{\mathbf{k}}\,, \qquad \text{and} \qquad \mathbf{v} = v_1\hat{\mathbf{\imath}} + v_2\hat{\mathbf{\jmath}} + v_3\hat{\mathbf{k}}\,.$$

Then

$$\begin{aligned}
\mathbf{u} \times \mathbf{v} &= (u_1\hat{\mathbf{\imath}} + u_2\hat{\mathbf{\jmath}} + u_3\hat{\mathbf{k}}) \times (v_1\hat{\mathbf{\imath}} + v_2\hat{\mathbf{\jmath}} + v_3\hat{\mathbf{k}}) \\
&= (u_1\hat{\mathbf{\imath}} + u_2\hat{\mathbf{\jmath}} + u_3\hat{\mathbf{k}}) \times v_1\hat{\mathbf{\imath}} + (u_1\hat{\mathbf{\imath}} + u_2\hat{\mathbf{\jmath}} + u_3\hat{\mathbf{k}}) \times v_2\hat{\mathbf{\jmath}} + (u_1\hat{\mathbf{\imath}} + u_2\hat{\mathbf{\jmath}} + u_3\hat{\mathbf{k}}) \times v_3\hat{\mathbf{k}} \\
&= u_1v_1\mathbf{i} \times \mathbf{i} + u_2v_1\mathbf{j} \times \mathbf{i} + u_3v_1\mathbf{k} \times \mathbf{i} + u_1v_2\mathbf{i} \times \mathbf{j} + u_2v_2\mathbf{j} \times \mathbf{j} + u_3v_2\mathbf{k} \times \mathbf{j} \\
&\qquad\qquad + u_1v_3\mathbf{i} \times \mathbf{k} + u_2v_3\mathbf{j} \times \mathbf{k} + u_3v_3\mathbf{k} \times \mathbf{k} \\
&= u_2v_1(-\hat{\mathbf{k}}) + u_3v_1\hat{\mathbf{\jmath}} + u_1v_2\hat{\mathbf{k}} + u_3v_2(-\hat{\mathbf{\imath}}) + u_1v_3(-\hat{\mathbf{\jmath}}) + u_2v_3\hat{\mathbf{\imath}} \\
&= (u_2v_3 - u_3v_2)\hat{\mathbf{\imath}} + (u_3v_1 - u_1v_3)\hat{\mathbf{\jmath}} + (u_1v_2 - u_2v_1)\hat{\mathbf{k}}\,.
\end{aligned}$$

§ 3.5. Determinants and Areas

A popular mnemonic for the cross product's component formula can be given in terms of an object from linear algebra called a *determinant*. Determinants are defined for square matrices, which may be thought of as collections of vectors. Determinants characterize the linear independence of collections of vectors, and give a way to compute volumes of objects determined by collections of vectors. The 2-dimensional case of the determinant characterizes when two vectors are not parallel, and gives us a formula for the area of a parallelogram in terms of the vectors spanning its sides. It shows up in explicit solutions for the intersection of two lines, and it turns out the coefficients in our cross product formula are all determinants of certain 2×2 matrices.

To see where determinants come from, you could try solving the system of equations

$$\begin{cases} ax + by &= e \\ cx + dy &= f \end{cases}.$$

After a little algebra, you should find that

$$x = \frac{de - bf}{ad - bc}\,, \quad y = \frac{af - ce}{ad - bc}\,,$$

provided the fractions given are defined. In particular, the quantity $ad - bc$ cannot be zero if you wish to have a unique solution corresponding to a point of intersection. We say that $ad - bc$ is *the determinant associated to this linear system of equations*. Traditionally, determinants are defined for square matrices, so we shall look briefly at those in the simplest cases of 2×2 and 3×3 matrices.

One often rewrites such a linear system with a *matrix-vector product*, as follows:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix},$$

where the square matrix acts on the vector $\langle x, y\rangle$ (written above in column notation) by the formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

Note that each entry of the resulting vector is the dot product of a *row vector* with the *column vector* that the matrix is acting on. By equating this with the constant vector $\langle e, f\rangle$, and thus separately equating their components, we recover the linear system of equations above.

We will define the determinant of the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

to be $ad - bc$. One often writes

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc\,.$$

This definition is more than merely a convenient shorthand. One sees from the above algebra that a linear system associated to this matrix possesses a unique solution *if and only if the determinant is nonzero*. The determinant is nonzero if and only if the vectors $\langle a, b\rangle$ and $\langle c, d\rangle$ are linearly independent. (for two vectors this means not parallel). Why might this number $ad - bc$ capture that fact?

It turns out, $ad - bc$ is the *signed area* of the parallelogram with one vertex the origin and sides given by the vectors $\langle a, b\rangle$ and $\langle c, d\rangle$. One can prove this using a little bit of trigonometry (try writing the vectors in polar coordinates, compute the determinant, then try to relate it via a trigonometric identity to the formula for parallelogram area.) Of course, if the vectors are parallel, they do not determine a true parallelogram, and one can think of the 'area enclosed" as 0.

It should then be unsurprising that determinants show up in our coordinate expression for the cross product, since we defined it so that its length was the area of a parallelogram! The popular mnemonic involves 3×3 determinants, though. Thankfully, a 3×3 determinant is built as a weighted sum of three 2×2 determinants.

Define a 3×3 determinant to be

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$
$$= a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1\,.$$

By an abuse of notation, we place along the first row $\hat{\mathbf{\imath}}$, $\hat{\mathbf{\jmath}}$, and $\hat{\mathbf{k}}$ instead of a_1, a_2 and a_3, and along the second and third rows we place the components of the vectors we wish to cross, with the first vector of the product having its entries in the second row, and the last vector having its entries in the final row. This gives us

$$\mathbf{u}\times\mathbf{v} = \begin{vmatrix} \hat{\mathbf{\imath}} & \hat{\mathbf{\jmath}} & \hat{\mathbf{k}} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}\hat{\mathbf{\imath}} - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}\hat{\mathbf{\jmath}} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}\hat{\mathbf{k}}$$
$$= (u_2v_3 - u_3v_2)\hat{\mathbf{\imath}} + (u_3v_1 - u_1v_3)\hat{\mathbf{\jmath}} + (u_1v_2 - u_2v_1)\hat{\mathbf{k}}\,.$$

We'll shortly connect this mnemonic with the volume interpretation of determinants, but for the moment, we present an example involving the geometric interpretation of the cross product in terms of area.

Example 3.1. Find the area of the triangle with vertices $P(-1,-2,-3)$, $Q(1,-1,-3)$ and $R(1,2,0)$.

Solution: The triangle's sides are spanned by displacement vectors $\overrightarrow{PQ} = \mathbf{q}-\mathbf{p} = \langle 2,1,0\rangle$ and $\overrightarrow{PR} = \mathbf{r}-\mathbf{p} = \langle 2,4,3\rangle$. The area of the triangle is half of the area of the parallelogram spanned by these vectors. Thus, we compute $\|\overrightarrow{PQ}\times\overrightarrow{PR}\|/2$:

$$\begin{vmatrix} \hat{\mathbf{\imath}} & \hat{\mathbf{\jmath}} & \hat{\mathbf{k}} \\ 2 & 1 & 0 \\ 2 & 4 & 3 \end{vmatrix} = 3\hat{\mathbf{\imath}} - 6\hat{\mathbf{\jmath}} + 6\hat{\mathbf{k}}\,.$$
$$\frac{1}{2}\|\overrightarrow{PQ}\times\overrightarrow{PR}\| = \frac{1}{2}\sqrt{9+36+36} = \frac{9}{2}\,.$$

Thus the area of the triangle is $9/2$.

---

§ 3.6. Triple Products

Consider a triple of linearly independent vectors $\mathbf{a},\mathbf{b},\mathbf{c}\in\mathscr{V}^3_{\mathbb{R}}$, placed with tails at the origin of $\mathbb{R}^3$. Each pair taken from this triple spans a parallelogram with one vertex at $\mathbf{0}$, and viewed as position vectors, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{0}$ together with the sums $\mathbf{a}+\mathbf{b}$, $\mathbf{b}+\mathbf{c}$, $\mathbf{c}+\mathbf{a}$, and $\mathbf{a}+\mathbf{b}+\mathbf{c}$, form the vertices of a solid called a parallelepiped. The points on the interior of the parallelepiped form the set $\mathcal{W} = \{t_1\mathbf{a}+t_2\mathbf{b}+t_3\mathbf{c}\,|\,0\le t_i\le 1, i=1,2,3\}$. We wish to give an expression for the volume of the parallelepiped.

Reasoning geometrically, this volume should equal the area of any one of the parallelogram faces times an *altitude* perpendicular to that face, i.e. a line segment starting with the chosen face and

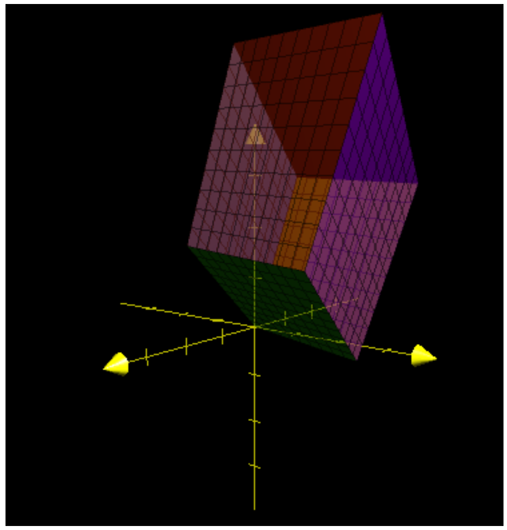


FIGURE 15. A view of a parallelepiped in $\mathbb{R}^3$.

ending in the opposite face, meeting each orthogonally. We can use a cross product of two of our vectors to compute the area of one of the faces, and a little geometric reasoning allows us to see that scaling the cross product by the length of the remaining vector times the cosine of the angle of separation between the cross product and this remaining vector gives us, up to a sign, an expression equal to the product of the area of a face with the length of an altitude. Thus, the volume of the parallelepiped is given by the absolute value of a *scalar triple product*

$$\mathcal{V} = |(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}|\,.$$

If the triple is right handed, the scalar triple product is positive, and gives the volume straight away. Convince yourself of the following:

Proposition. *For a given triple of vectors* $\mathbf{a},\mathbf{b},\mathbf{c}\in\mathscr{V}^3_{\mathbb{R}}$*, the possible rearrangements of scalar triple products yield only two values, one of which is the volume of the associated parallelepiped, and the other the negative of this volume. The positive arrangements for a right-handed triple* $(\mathbf{a},\mathbf{b},\mathbf{c})$ *are*

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}\,,\qquad (\mathbf{c}\times\mathbf{a})\cdot\mathbf{b}\,,\qquad (\mathbf{b}\times\mathbf{c})\cdot\mathbf{a}\,.$$

The negative arrangements are

$$(\mathbf{a}\times\mathbf{c})\cdot\mathbf{b}\,,\qquad (\mathbf{b}\times\mathbf{a})\cdot\mathbf{c}\,,\qquad (\mathbf{c}\times\mathbf{b})\cdot\mathbf{a}\,.$$

Let's examine the triple product with our coordinate expressions for the dot and cross products. Consider, e.g. $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$ with $\mathbf{a}=a_1\hat{\imath}+a_2\hat{\jmath}+a_3\hat{\mathbf{k}}$, $\mathbf{b}=b_1\hat{\imath}+b_2\hat{\jmath}+b_3\hat{\mathbf{k}}$, and $\mathbf{c}=c_1\hat{\imath}+c_2\hat{\jmath}+c_3\hat{\mathbf{k}}$. Then

$$\begin{aligned}
\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= \big(a_1\hat{\imath}+a_2\hat{\jmath}+a_3\hat{\mathbf{k}}\big)\cdot\Big((b_2c_3-b_3c_2)\hat{\imath}-(b_1c_3-b_3c_1)\hat{\jmath}+(b_1c_2-b_2c_1)\hat{\mathbf{k}}\Big)\\
&= a_1(b_2c_3-b_3c_2)-a_2(b_1c_3-b_3c_1)+a_3(b_1c_2-b_2c_1)\\
&= a_1\begin{vmatrix} b_2 & b_3\\ c_2 & c_3\end{vmatrix} - a_2\begin{vmatrix} b_1 & b_3\\ c_1 & c_3\end{vmatrix} + a_3\begin{vmatrix} b_1 & b_2\\ c_1 & c_2\end{vmatrix}\\
&= \begin{vmatrix} a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}\,.
\end{aligned}$$

This demonstrates that 3×3 determinants compute volumes, up to sign!

Example 3.2. Find the volume of the parallelepiped with edges spanned by the vectors $\langle -2, 1, 2\rangle$, $\langle 2, 1, -2\rangle$ and $\langle -1, 2, -1\rangle$.

Solution: we compute the determinant

$$\begin{vmatrix} -2 & 1 & 2 \\ 2 & 1 & -2 \\ -1 & 2 & -1 \end{vmatrix} = 2 + 2 + 8 - (-2) - (8) - (-2) = 8\,.$$

Here, the 3×3 determinant was calculated using the formula

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - c_1b_2a_3 - a_2b_3c_1 - c_3b_1a_2\,.$$

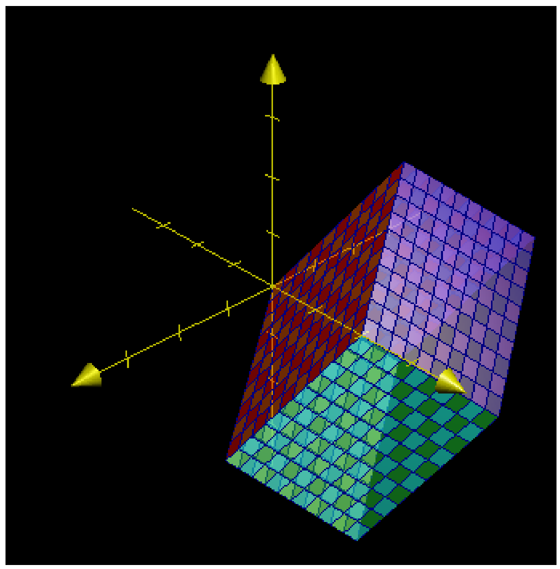


FIGURE 16. The volume of this parallelepiped is 8.

There is another type of triple product to consider, the *vector triple product*. For example, one might consider products of the form

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c})\,, \qquad \text{and} \qquad (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}\,.$$

These are not in general equal.

Proposition. *For* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathscr{V}^3_{\mathbb{R}}$, *the vector triple product satisfies*

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

This is sometimes written as $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$, leading to the mnemonic name "BAC−CAB rule" (pronounced "back minus cab rule").

We offer a geometric argument for this identity, with a key part left as an exercise. First, convince yourself that the following special case holds: if $\mathbf{u} \cdot \mathbf{v} = 0$, then $\mathbf{u} \times (\mathbf{v} \times \mathbf{u}) = \|\mathbf{u}\|^2\,\mathbf{v}$ (without reference to the rule we seek to prove, of course – see problem (13) below). Now, the first observation is that if $\mathbf{b} \parallel \mathbf{c}$, then the vector triple product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{0}$. So let's assume $\mathbf{b}$ and $\mathbf{c}$ are linearly independent.

Now, since $(\mathbf{b} \times \mathbf{c})$ is perpendicular to both $\mathbf{b}$ and $\mathbf{c}$, the product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is perpendicular to $(\mathbf{b} \times \mathbf{c})$ and thus lies in the plane spanned by $\mathbf{b}$, $\mathbf{c}$. Anything in this plane can be written as a linear combination of the linearly independent vectors $\mathbf{b}$ and $\mathbf{c}$, whence there are scalars λ and μ such that

$$(*) \qquad \mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \lambda\mathbf{b} + \mu\mathbf{c}\,.$$

Now we wish to solve for λ and μ. If we can dot both sides of the equation $(*)$ with a vector that is perpendicular to $\mathbf{c}$ which has an understood dot product with the left hand side of $(*)$ and with $\mathbf{b}$, then we could solve for λ. We can't use $\mathbf{b}\times\mathbf{c}$, as this is perpendicular to $\mathbf{b}$, $\mathbf{c}$, and the triple product $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$ (and so dotting with this would would annihilate all terms). Note however that if we cross $\mathbf{c}$ with $\mathbf{b}\times\mathbf{c}$, we obtain a vector that is perpendicular to $\mathbf{c}$ and to $\mathbf{b}\times\mathbf{c}$, but not with $\mathbf{b}$ (as it lies in the plane spanned by $\mathbf{b}$ and $\mathbf{c}$). Taking a dot product with $\mathbf{c}\times(\mathbf{b}\times\mathbf{c})$ on both sides of $(*)$ yields

$$\begin{aligned}\big(\mathbf{a}\times(\mathbf{b}\times\mathbf{c})\big)\cdot\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big) &= \big(\lambda\mathbf{b}+\mu\mathbf{c}\big)\cdot\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big)\\ &= \lambda\mathbf{b}\cdot\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big) + \mu\mathbf{c}\cdot\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big)\\ &= \lambda\mathbf{b}\cdot\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big)\,.\end{aligned}$$

We will use the fact that for the scalar triple product, $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = (\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}$ to simplify both the left and right sides. On the left, we have

$$\begin{aligned}\big(\mathbf{a}\times(\mathbf{b}\times\mathbf{c})\big)\cdot\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big) &= \mathbf{a}\cdot\Big((\mathbf{b}\times\mathbf{c})\times\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big)\Big)\\ &= \mathbf{a}\cdot\big(\,\|\mathbf{b}\times\mathbf{c}\|^2\,\mathbf{c}\big) = (\mathbf{a}\cdot\mathbf{c})\,\|\mathbf{b}\times\mathbf{c}\|^2\,,\end{aligned}$$

which relies on the fact that $\mathbf{c}\perp\mathbf{b}\times\mathbf{c}$ implies that the triple product

$$(\mathbf{b}\times\mathbf{c})\times\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big) = \big(\,\|\mathbf{b}\times\mathbf{c}\|^2\,\mathbf{c}\big)\,,$$

by the special case above (left to the reader as problem (12) below). Meanwhile, on the right-hand side we have

$$\lambda\mathbf{b}\cdot\big(\mathbf{c}\times(\mathbf{b}\times\mathbf{c})\big) = \lambda(\mathbf{b}\times\mathbf{c})\cdot(\mathbf{b}\times\mathbf{c}) = \lambda\,\|\mathbf{b}\times\mathbf{c}\|^2\,.$$

Equating the left and right sides and canceling, we have

$$(\mathbf{a}\cdot\mathbf{c})\,\|\mathbf{b}\times\mathbf{c}\|^2 = \lambda\,\|\mathbf{b}\times\mathbf{c}\|^2\,,$$

$$\lambda = \mathbf{a}\cdot\mathbf{c}\,.$$

Similarly, by dotting both sides of $(*)$ with $\mathbf{b}\times(\mathbf{b}\times\mathbf{c})$, we can show $\mu = -(\mathbf{a}\cdot\mathbf{b})$, which proves the BAC−CAB identity for the vector triple product $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$.

---

§ 3.7. Problems

(1) Compute all possible pairs of cross products and triple products for the vectors $\mathbf{u} = \langle 1,-2,3\rangle$, $\mathbf{v} = \langle -4,2,-2\rangle$, $\mathbf{w} = \langle 7,6,5\rangle$, and $\mathbf{y} = \langle 8,0,9\rangle$.

(2) If you are given non-collinear three points, they determine a triangle whose edges are line segments connecting pairs of the points. How many different displacement vectors can you find modeling these edges? From these, how many different cross products can be computed? How are their magnitudes related?

(3) Referring to a picture and by using trigonometry and vectors, argue that the determinant of a 2×2 matrix gives the signed area of a parallelogram. How is the sign determined?

(4) At what angle relative to a lever arm described by $\boldsymbol{\ell} \in \mathscr{V}^3_{\mathbb{R}}$ should you apply a force $\mathbf{F}$ to maximize the torque $\boldsymbol{\tau}$ on $\boldsymbol{\ell}$? If you reduce the force by a factor of 2/3, how much longer would $\boldsymbol{\ell}$ need to be for the new torque to equal the previous torque?

(5) A tetrahedron is a solid with four triangular faces. Give a vector-based formula for the volume of a tetrahedron spanned by vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathscr{V}^3_{\mathbb{R}}$ (Hint: relate the tetrahedron to the parallelepiped associated to the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$).

(6) The points $P(4, 0, -4)$, $Q(3, -3, 0)$, $R(0, -4, 4)$, $S(-3, 0, 3)$, $T(-4, 4, 0)$ and $U(0, 3, -3)$ form an irregular hexagon in $\mathbb{R}^3$. Draw the hexagon in 3-space. What is the unit normal to the plane in which this hexagon lies? Find the perimeter and area of the hexagon.

(7) Suppose a cube sits atop the xy-plane with one vertex at the origin, and with a vertex at $\hat{\mathbf{\imath}} + \hat{\mathbf{\jmath}} + 2\hat{\mathbf{k}}$ that is diagonally across a face from the origin. The cube's vertex opposite the origin is directly above the origin, so that the cube's diagonal is along the z-axis. Find the other vertices of the cube. Then compute the surface area and volume of the cube using vector products, and check that these answers agree with what you get from the standard formulae.

(8) For each of the properties (1) – (4) of the cross product listed in the proposition above, use pictures and geometric reasoning to argue the truth of the properties without appealing to coordinate algebra.

(9) Suppose $* : \mathscr{V}^n_{\mathbb{R}} \times \mathscr{V}^n_{\mathbb{R}} \to \mathscr{V}^n_{\mathbb{R}}$ is a skew symmetric bilinear operator on n-dimensional vectors, i.e. it satisfies properties (1) – (4) with '$\times$' replaced by '$*$'. Show that $\mathbf{u} * \mathbf{u} = \mathbf{0}$ for any $\mathbf{u} \in \mathscr{V}^n_{\mathbb{R}}$.

(10) Give an example of vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathscr{V}^3_{\mathbb{R}}$ for which $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$. Then, geometrically characterize when $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$ holds.

(11) Give a proof of the Jacobi identity $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) + \mathbf{v} \times (\mathbf{w} \times \mathbf{u}) + \mathbf{w} \times (\mathbf{u} \times \mathbf{v}) = \mathbf{0}$ without using coordinate expressions.

(12) Show that $\|\mathbf{a} \times \mathbf{b}\|^2 = \|\mathbf{a}\|^2 \|\mathbf{b}\|^2 - (\mathbf{a} \cdot \mathbf{b})^2$ using the geometric definitions of dot and cross products.

(13) Prove the special case of the vector triple product, without using the BAC$-$CAB rule:

$$\mathbf{u} \cdot \mathbf{v} = 0 \implies \mathbf{u} \times (\mathbf{v} \times \mathbf{u}) = \|\mathbf{u}\|^2 \mathbf{v}\,.$$

(14) Let $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathscr{V}^3_{\mathbb{R}}$. Show by geometric/algebraic arguments without appealing to coordinate calculations that

$$(\mathbf{a} \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{c}) = \big(\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})\big)\mathbf{a}\,, \text{ and}$$
$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})\,.$$

(15) For a non-associative operator $*: \mathscr{V} \to \mathscr{V}$ on a vector space $\mathscr{V}$, one can define the *associator bracket* $[\cdot,\cdot,\cdot] : \mathscr{V} \times \mathscr{V} \times \mathscr{V} \to \mathscr{V}$ given by $[\mathbf{a},\mathbf{b},\mathbf{c}] = (\mathbf{a} * \mathbf{b}) * \mathbf{c} - \mathbf{a} * (\mathbf{b} * \mathbf{c})$. What is the associator of the cross product? Can you give a geometric interpretation of the cross product's associator? (Hint: Consider the Jacobi identity, and think about linear independence.)

(16) Find a matrix $\mathrm{A}_{\mathbf{u}\times}$ such that the matrix vector product $\mathrm{A}_{\mathbf{u}\times}\mathbf{v} = \mathbf{u}\times\mathbf{v}$, and another matrix $\mathrm{P}_{\mathrm{proj}_{\mathbf{u}}}$ such that the matrix vector product $\mathrm{P}_{\mathrm{proj}_{\mathbf{u}}}\mathbf{v} = \mathrm{proj}_{\mathbf{u}}\,\mathbf{v}$. Use these to reverify the identities proven in problems (7), (10), and (13).

4. Lines

§ 4.1. Vector Equations and Parametric Equations

Recall that in two dimensions we can describe a line by an equation of the form

$$y = mx + b\,,$$

where b is the height of the y-intercept, and $m = \frac{\Delta y}{\Delta x}$ is the slope. One can instead describe a line in two dimensions by a vector equation. Take any vector $\mathbf{v}$ with slope m, e.g. $\mathbf{v} = \hat{\mathbf{\imath}} + m\hat{\mathbf{\jmath}}$, for m defined, or $\mathbf{v} = \hat{\mathbf{\jmath}}$ if the line is vertical. Then the line can be written in vector form as $\langle x, y\rangle = t\mathbf{v} + b\hat{\mathbf{\jmath}}$, where $t \in \mathbb{R}$ is called a parameter. Indeed, note that when $t = 0$, we get $x = 0$, $y = b$, which is the intercept, and for $t = 1$ we have

$$x = 1\,, \qquad y = m + b\,.$$

Generally, setting $t = x_1$ gives $x = x_1, y = mx_1 + b$. More generally, if we have a point (x_0, y_0) on the line, and we know a vector $\mathbf{v}$ parallel to the line, we can describe the line by a vector equation

$$\mathbf{r} = x\hat{\mathbf{\imath}} + y\hat{\mathbf{\jmath}} = t\mathbf{v} + \mathbf{r}_0\,.$$

We can generalize this to higher dimensions: as long as we know at least one point $\mathbf{r}_0$ on a line, and a vector $\mathbf{v}$ parallel to the line, we can give a vector equation for the line of the form

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}\,.$$

Writing this out in components for the three dimensional case we have

$$x\,\hat{\mathbf{\imath}} + y\,\hat{\mathbf{\jmath}} + z\,\hat{\mathbf{k}} = x_0\,\hat{\mathbf{\imath}} + y_0\,\hat{\mathbf{\jmath}} + z_0\,\hat{\mathbf{k}} + tv_1\,\hat{\mathbf{\imath}} + tv_2\,\hat{\mathbf{\jmath}} + tv_3\,\hat{\mathbf{k}}\,,$$

which gives us the *parametric equations*

$$x = x_0 + tv_1\,, \qquad y = y_0 + tv_2\,, \qquad z = z_0 + tv_3$$

for the line.

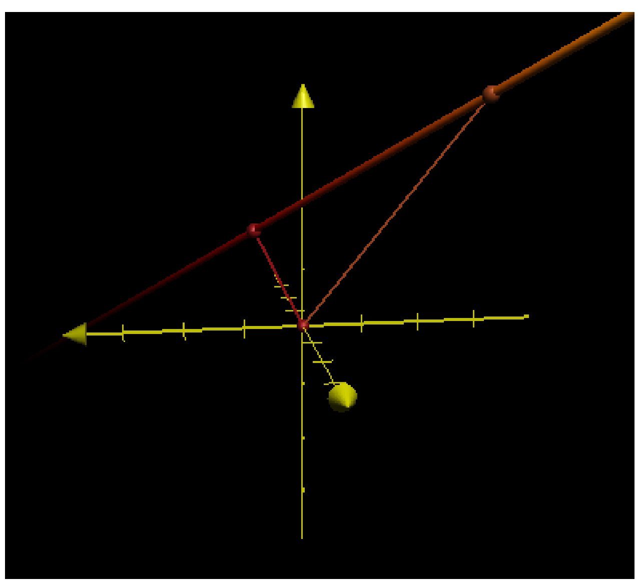


FIGURE 17. A line in $\mathbb{R}^3$, together with two positions along it. It may be parameterized by the vector equation $\mathbf{r}(t) = \langle 1 - t, 2 + 2t, 2 + t\rangle$.

Example 4.1. Find parametric equations for the line from $A(1, 2, 1)$ to $B(-4, 3, 6)$.

Solution: We need a point on the line, and a direction. Since we know two points on the line, we can work with either one as our initial point, and we can use the displacement vector from one to the other as the direction. For example, we could set $\mathbf{r}_0 = \mathbf{a} = \langle 1, 2, 1\rangle$, and use the displacement vector

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a} = \langle -5, 1, 5\rangle$$

as the velocity vector $\mathbf{v}$. The vector equations read as

$$\mathbf{r} = \langle x, y, z\rangle = \langle 1, 2, 1\rangle + t\langle -5, 1, 5\rangle\,,$$

which gives us parametric equations

$$x = 1 - 5t\,, \qquad y = 2 + t\,, \qquad z = 1 + 5t\,.$$

Swapping in $\mathbf{r}_0 = \mathbf{b}$ instead of $\mathbf{a}$ will change the initial constants, but you should convince yourself that this corresponds to changing the above parametric equations by shifting t to $t + 1$. And if we instead chose $\overrightarrow{BA}$ as the direction vector, the equations would differ from the those above by an exchange of $-t$ for t. In particular, the different constant speed parametric equations for a line differ only by a linear change of the parameter.

Let us look at the above example symbolically. We were given two points A and B, to which we associate position vectors $\mathbf{a}$ and $\mathbf{b}$. Then we can obtain a vector equation for the line connecting A and B by

$$\mathbf{r} = \mathbf{a} + t\overrightarrow{AB} = \mathbf{a} + t(\mathbf{b} - \mathbf{a}) = (1 - t)\mathbf{a} + t\mathbf{b}\,.$$

Observe that $t = 0$ gives us a position of $\mathbf{a}$, while $t = 1$ gives us a position of $\mathbf{b}$. For any t between 0 and 1, we obtain a point between A and B along the line segment connecting A and B. Thus, we can describe the line segment as

$$\mathbf{r} = \mathbf{a} + t(\mathbf{b} - \mathbf{a}) = (1 - t)\mathbf{a} + t\mathbf{b}\ , \quad 0 \le t \le 1\,.$$

We can swap out t and *reparameterize* to suite certain applications. For example, if we want to describe linear motion from a point A to B with a certain fixed *speed* (the magnitude of velocity), we can reparameterize to alter the length of $\mathbf{v}$, as in the following example.

Example 4.2. Parameterize the line segment from $A(-3, 4, 5)$ to $B(1, -3, 1)$ so that the speed is 2.

Solution: From the above formula, we can parametrize the line segment from A to B as

$$\mathbf{r}(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a}) = (-3 + 4t)\hat{\imath} + (4 - 7t)\hat{\jmath} + (5 - 4t)\hat{\mathbf{k}}\,, \quad 0 \le t \le 1\,.$$

The velocity vector is $\mathbf{v} = 4\hat{\imath} - 7\hat{\jmath} - 4\hat{\mathbf{k}}$, which has length $\|\mathbf{v}\| = \sqrt{16 + 49 + 16} = \sqrt{81} = 9$. Thus, if we scale the velocity by $\frac{2}{9}$ then the new velocity will have magnitude 2, as desired. But scaling $\mathbf{v}$ can be accomplished by setting $t = \frac{2}{9}s$, and rescaling the domain. Since the original line segment is defined by the bounds $0 \le t \le 1$, we obtain new bounds

$$0 \le \tfrac{2}{9}s \le 1 \implies 0 \le s \le \tfrac{9}{2}\,.$$

The desired parameterization for the line segment is therefore

$$\mathbf{r}(s) = \left(-3 + \frac{8}{9}s\right)\hat{\imath} + \left(4 - \frac{14}{9}s\right)\hat{\jmath} + \left(5 - \frac{8}{9}s\right)\hat{\mathbf{k}}\,, \qquad 0 \le s \le \frac{9}{2}\,.$$

---

§ 4.2. Symmetric Equations

Note that we can solve each of the parametric equations for t, giving

$$t = \frac{x - x_0}{v_1}\,, \qquad t = \frac{y - y_0}{v_2}\,, \qquad t = \frac{z - z_0}{v_3}\,.$$

Setting these equal gives *symmetric equations of the line*:

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2} = \frac{z - z_0}{v_3}\,.$$

The components of the velocity vector appearing in the denominators in the symmetric equations are called *direction numbers*. Observe that given symmetric equations, one can easily recover a vector equation.

Example 4.3. For the points $P(-3, 2, 5)$, $Q(-1, -5, -7)$ and $R(4, -3, -1)$, find symmetric equations for the line perpendicular to the triangle $\triangle PQR$ passing through the triangle's centroid.

Solution: First, we find the centroid, whose position vector will serve the role of $\mathbf{r}_0 = \langle x_0, y_0, z_0 \rangle$:

$$\mathbf{r} = \frac{1}{3}(\mathbf{p} + \mathbf{q} + \mathbf{r}) = 0\hat{\imath} - 2\hat{\jmath} - \hat{\mathbf{k}}\,.$$

We need direction numbers v_1, v_2 and v_3, which come from a velocity vector. Since we want the normal direction to $\triangle PQR$ we can cross the displacement vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ to get the direction numbers:

$$\overrightarrow{PQ} \times \overrightarrow{PR} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{\mathbf{k}} \\ 2 & -7 & -12 \\ 7 & -5 & -6 \end{vmatrix} = -18\hat{\imath} - 72\hat{\jmath} + 39\hat{\mathbf{k}}$$
$$= -3(9\hat{\imath} + 24\hat{\jmath} - 13\hat{\mathbf{k}})\,.$$

The direction numbers are only defined up to multiplication of them all by some scalar, so we can take

$$v_1 = 9\,, \qquad v_2 = 24\,, \qquad v_3 = 13\,.$$

Thus, we can describe the line perpendicular to $\triangle PQR$ through its centroid by the symmetric equations

$$\frac{x}{9} = \frac{y + 2}{24} = \frac{z + 1}{-13}\,.$$

---

§ 4.3. Problems

(1) Find vector equations for a line parallel to the line $\mathbf{r} = \hat{\imath} - \hat{\jmath} + t(\hat{\jmath} + 2\hat{\mathbf{k}})$ such that it passes through the point $P(7, 7, 7)$.

(2) Find parametric equations for a a line normal to the triangle with vertices $P(0, -2, 0)$, $Q(2, 0, 2)$ and $R(-1, -1, -1)$, such that at $t = 1$ the line is at the point above the triangle along the normal direction at a distance of 1 from the centroid of $\triangle PQR$.

(3) Parametrize the line between the points $P\left(\frac{1}{2}, \frac{2}{3}, \frac{3}{4}\right)$ and $Q\left(-\frac{4}{5}, \frac{5}{6}, -\frac{6}{7}\right)$ so that the parameter is the height above the xy-plane (i.e. use $z = t$ as the parameter, for an appropriate choice of velocity vector).

(4) Give a vector parametrization of the line with equations
$$x = 3\,, \quad \frac{y+2}{-1} = \frac{2z-3}{4}\,.$$

(5) Use a vector equation to parametrize the diameter of the radius 6 sphere centered at $C(1,-2,6)$, such that the diameter contains the point $P(3,-1,8)$.

(6) Find symmetric equations for the line, a segment of which gives the shortest distance between the surface of the sphere of radius 2 centered at $P(-1,-2,-3)$ and the surface of the sphere of radius 4 centered at $Q(6,7,8)$. What is this shortest distance?

(7) Find a parametrization of a line whose velocity is perpendicular to the line with parametric equations
$$x = 1-6t\,, \qquad y = -6+t\,, \qquad z = 4t\,,$$
such that the two lines never intersect (be sure to show they never intersect.) Generally, lines which are not parallel but never intersect are called *skew*.

(8) Parametrize the line segments of edges of the hexagon with vertices $P(4,0,-4)$, $Q(3,-3,0)$, $R(0,-4,4)$, $S(-3,0,3)$, $T(-4,4,0)$ and $U(0,3,-3)$ by unit-speed parameterizations. Arrange your parameterizations into a piecewise function describing continuous motion around the perimeter. How long would it take for a particle to traverse the entire perimeter if it moved along each line segment according to these unit speed parameterizations?

(9) Recall that the medians of a triangle are line segments from the midpoints of the triangle's sides to opposite vertices. Use vector algebra and the vector description of line segments to prove that the centroid of a triangle is the intersection of its medians.

(10) Recall that a *chord* of a circle is a line segment connecting two points P and Q that lie on the circle's circumference. Find a parametrization $\mathbf{r}(\theta)$ of the chord of the unit circle from $(1,0)$ to $(0,1)$ with parameter equal to the angle θ made by $\mathbf{r}$ with the x-axis. Give the speed of motion along the chord under this parametrization.

(11) Find the a formula for the point on the line $\mathbf{r}(t) = \mathbf{r}_0 + t\mathbf{v}$ that is closest to the origin.

(12) Given a line with parametric equations
$$x = x_0 + tv_1\,, \qquad y = y_0 + tv_2\,, \qquad z = z_0 + tv_3$$
and a vector $\mathbf{u}$ not parallel to the given line, describe a procedure for generating a line parallel to $\mathbf{u}$ which is skew to the given line, and write a general form for the parametric equations of the new line.

(13) Given two non-parallel lines L_1 and L_2 with direction vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ respectively, which meet at a point $\mathbf{r}_0$, give a unit-speed vector equation for the line through $\mathbf{r}_0$, perpendicular to both L_1 and L_2, such that $t = 0$ gives the point on this line nearest the origin.

5. Planes

§ 5.1. Vector Equations and Linear forms

A vector equation of a plane Π in $\mathbb{R}^3$ can be described directly from the dot product, provided one knows a point on the plane and a normal direction to the plane. If $\mathbf{r}_0$ is a known point contained in the plane Π, and $\mathbf{r}$ an arbitrary point of Π, then for any normal vector $\mathbf{n}$ to Π, the vector $\mathbf{r}-\mathbf{r}_0$ must be orthogonal to $\mathbf{n}$, which means

$$\mathbf{n}\cdot(\mathbf{r}-\mathbf{r}_0)=0\,.$$

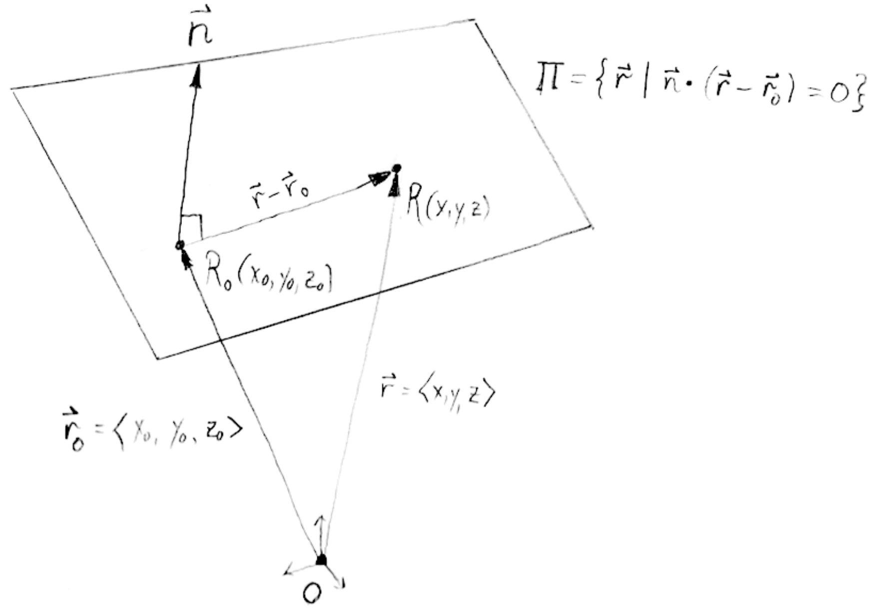


FIGURE 18. A plane in $\mathbb{R}^3$ consists of all points at positions $\mathbf{r}=\langle x,y,z\rangle$ displaced from a given point $R_0(x_0,y_0,z_0)$ with position $\mathbf{r}_0$ by displacement vectors perpendicular to a given normal vector $\mathbf{n}$, yielding a formula $\mathbf{n}\cdot(\mathbf{r}-\mathbf{r}_0)=0$.

Writing $\mathbf{n}=\langle a,b,c\rangle$, $\mathbf{r}=\langle x,y,z\rangle$ and $\mathbf{r}_0=\langle x_0,y_0,z_0\rangle$, the dot product equation above becomes

$$a(x-x_0)+b(y-y_0)+c(z-z_0)=0\,.$$

Since $\mathbf{r}_0$ is presumed known, we can set $d:=-\mathbf{n}\cdot\mathbf{r}_0=ax_0+by_0+cz_0$, and write

$$ax+by+cz+d=0\,.$$

This is a more standard way to write the equation of a plane in $\mathbb{R}^3$. Such an equation is called a *linear equation in three variables.* If $d\neq 0$, it is said to be *inhomogeneous.* Observe that a homogeneous equation, with $d=0$ corresponds to a plane which contains the origin.

Example 5.1. Find the equation of the plane containing the three points $P(5,4,3)$, $Q(-1,5,3)$ and $R(6,7,8)$.

Solution: Recall in example 1.1 we had computed

$$\overrightarrow{PQ}=\langle -6,1,0\rangle\,,\qquad \overrightarrow{PR}=\langle 1,3,5\rangle\,.$$

Their cross product gives us a possible normal to the plane, and we can use the position vector of any of the given points in place of $\mathbf{r}_0=\langle x_0,y_0,z_0\rangle$. We'll settle on using $\mathbf{p}=\langle 5,4,3\rangle$. Then the equation of the plane can be given by setting equal to zero the scalar triple product

$$\langle x-5, y-4, z-3\rangle \cdot (\overrightarrow{PQ} \times \overrightarrow{PR})\,.$$

This gives

$$0 = \left| \begin{matrix} x-5 & y-4 & z-3 \\ -6 & 1 & 0 \\ 1 & 3 & 5 \end{matrix} \right] = 5(x-5) + 30(y-4) - 19(z-3)\,.$$

Simplifying this yields

$$5x + 30y - 19z - 88 = 0\,.$$

Example 5.2. Give parametric equations for the line of intersection of the two planes

$$\Pi_1 : x + y + z = 2\,,$$

$$\Pi_2 : 3x - 3y + 2z = 1\,.$$

Solution: We give two ways to accomplish this: first, by vector algebra, and then, by linear algebra. The vector algebra way to do it is as follows: first, we need a velocity vector for the line. The cross product of the normals to the planes will necessarily be parallel to each plane, and thus parallel to the line of their intersection. Let $\mathbf{n}_1$ and $\mathbf{n}_2$ be the normals of the planes Π_1 and Π_2 respectively, and take $\mathbf{v}$ to be their cross product:

$$\mathbf{v} = \mathbf{n}_1 \times \mathbf{n}_2 = \langle 1,1,1\rangle \times \langle 3,-3,2\rangle = 5\hat{\mathbf{\imath}} + \hat{\mathbf{\jmath}} - 6\hat{\mathbf{k}}\,.$$

Now, we need to find a point on the line. For this, any point on the line will do. Since neither plane is of the form $z = k$ for some constant k, and since and the velocity of our line has nonzero component in the z direction, we are free to choose $z = 0$ (as the line will contain a point with z coordinate equal to 0). Then the equation for Π_1 gives us $x + y = 2$. Then setting $y = 2 - x$ in the equation for Π_2 gives $3x - 3(2-x) = 6x - 6 = 1$, so $x = \frac{7}{6}$, $y = 2 - \frac{7}{6} = \frac{5}{6}$. This tells us that the point $\left(\frac{7}{6}, \frac{5}{6}, 0\right)$ is a point on the line. Then our parametric equations are

$$x = \frac{7}{6} + 5t \qquad y = \frac{5}{6} + t\,, \qquad z = -6t\,.$$

The linear algebra solution is to use elimination to arrive at an equation dependent on one of the variables. For example, we can add 3 times the equation of Π_1 to the equation of Π_2 to obtain

$$6x + 5z = 7\,.$$

solving for x, we get $x = \frac{1}{6}(7-5z) = \frac{7}{6} - \frac{5}{6}z$. Using the equation of Π_1 we can now find y in terms of z as well:

$$\frac{1}{6}(7-5z) + y + z = 2 \implies y = \frac{5}{6} - \frac{1}{6}z\,.$$

Setting $z = s$, we can write parametric equations

$$x = \frac{7}{6} - \frac{5}{6}s\,, \qquad y = \frac{5}{6} - \frac{1}{6}s\,, \qquad z = s\,.$$

Note that by setting $s = -6t$ we recover the parametric equations found by vector algebra.

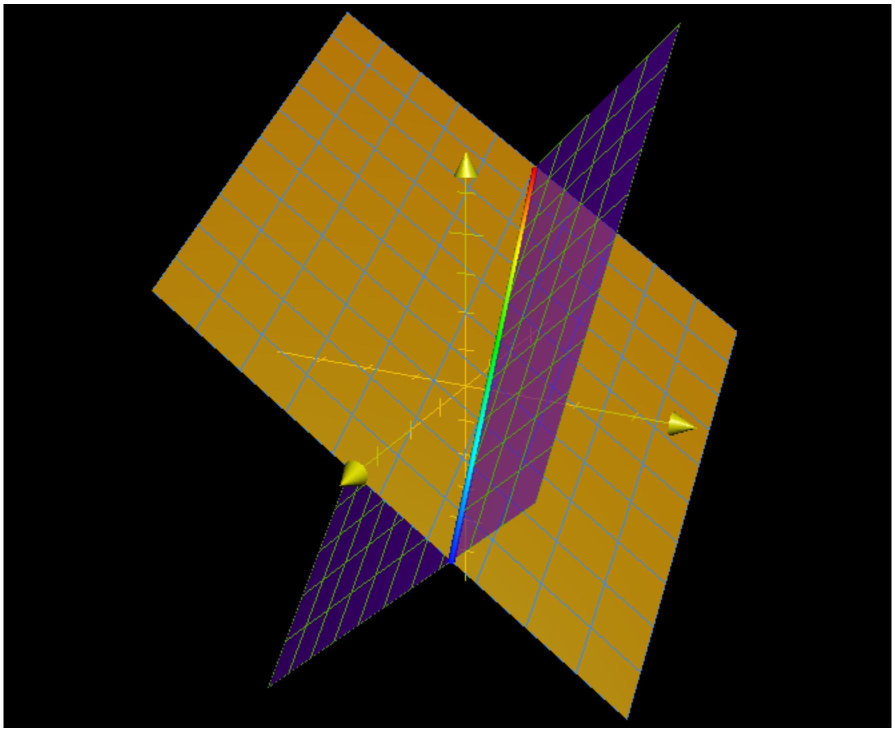


Figure 19. Portions of the planes Π_1 and Π_2, and the corresponding portion of their line of intersection.

§ 5.2. Distances

Let Π be a plane with normal $\mathbf{n} = \langle a, b, c\rangle$ containing the point R_0 with position vector $\mathbf{r}_0$, and suppose $P(x_1, y_1, z_1)$ is a point not on the plane. Then observe that the projection $\mathrm{proj}_{\mathbf{n}}(\mathbf{p} - \mathbf{r}_0)$ is a vector normal to the plane whose length is the shortest distance to the plane from P. Recalling that the magnitude of the projection is the absolute value of the *component* $\mathrm{comp}_{\mathbf{n}}(\mathbf{p} - \mathbf{r}_0)$, we have that the distance D from P to Π is

$$\begin{aligned} D &= |\mathrm{comp}_{\mathbf{n}}(\mathbf{p} - \mathbf{r}_0)| \\ &= \frac{|\mathbf{n}\cdot(\mathbf{p} - \mathbf{r}_0)|}{\|\mathbf{n}\|} \\ &= \frac{|ax_1 + by_1 + cz_1 + d|}{\sqrt{a^2 + b^2 + c^2}}\,, \end{aligned}$$

where $d = -\mathbf{n}\cdot\mathbf{r}_0$, which is the very same constant appearing in the plane equation when it is in the form $ax + by + cz + d = 0$.

Example 5.3. Find the distance between the skew lines L_1 and L_2 with the given vector equations

$$L_1 : 4\hat{\mathbf{\i}} - 7\hat{\mathbf{\j}} + 16\hat{\mathbf{k}} + t\left(\frac{1}{2}\hat{\mathbf{\i}} + \frac{1}{4}\hat{\mathbf{\j}} + \frac{1}{7}\hat{\mathbf{k}}\right)\,,$$

$$L_2 : -6\hat{\mathbf{\i}} + 5\hat{\mathbf{\j}} + 30\hat{\mathbf{k}} + s(28\hat{\mathbf{\i}} - 14\hat{\mathbf{\j}} + 32\hat{\mathbf{k}})\,.$$

Solution: The pair of skew lines lie in a pair of parallel planes, a normal of which is $\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2$, where $\mathbf{v}_i$ is a vector parallel to L_i, $i = 1, 2$. We can read such $\mathbf{v}_i$ off from the above equations and compute the normal $\mathbf{n}$, after which it is a matter of using a point in each plane (e.g. a point from

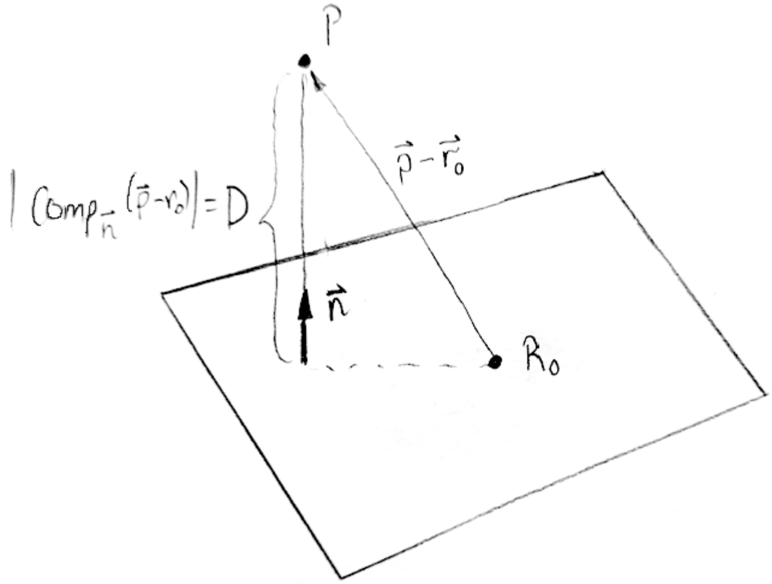


FIGURE 20. The distance from a point $P(x, y, z)$ with position $\mathbf{p}$ to the plane Π containing R_0 is the length of the projection of the displacement from R_0 to P onto any normal to the plane: $|\mathrm{comp}_{\mathbf{n}}(\mathbf{p} - \mathbf{r}_0)|$.

each line suffices) to apply the above formula for the distance.

$$\mathbf{n} = \left(\frac{1}{2}\hat{\imath} + \frac{1}{4}\hat{\jmath} + \frac{1}{7}\hat{\mathbf{k}}\right) \times (28\hat{\imath} - 14\hat{\jmath} + 32\hat{\mathbf{k}}) = 10\hat{\imath} - 12\hat{\jmath} - 14\hat{\mathbf{k}}\,.$$

$$\begin{aligned} D &= \frac{\left|\mathbf{n} \cdot \left[-6\hat{\imath} + 5\hat{\jmath} + 30\hat{\mathbf{k}} - (4\hat{\imath} - 7\hat{\jmath} + 16\hat{\mathbf{k}})\right]\right|}{\|\mathbf{n}\|} \\ &= \frac{|10(-10) - 12(12) - 14(-14)|}{\sqrt{10^2 + 12^2 + 14^2}} = 2\sqrt{110}\,. \end{aligned}$$

---

§ 5.3. Problems

(1) Find a normal vector $\mathbf{n}$ to the plane with equation $4x - 3y = 12 - 6z$, and find a vector $\mathbf{v}$ parallel to the plane. Verify that $\mathbf{n}$ is orthogonal to the vector parallel to the plane. How does $\mathbf{v} \times \mathbf{n}$ relate geometrically to $\mathbf{v}$ and $\mathbf{n}$?

(2) Find the angle between the planes

$$\Pi_1 : 12x + 18y - 17z = 32\,, \qquad \Pi_2 : -8x + 3y + 5z = 24\,.$$

(3) Find the equation of the plane through $P(1, 0, -1)$ perpendicular to the line

$$x = 9 - 4t\,, \quad y = -1 + 9t\,, \quad z = t\,.$$

(4) Find the point where the line $\mathbf{r}(t) = \langle 3 + t, 4 + 4t, -2 + 3t\rangle$ intersects the plane containing the point $P(8, 6, 3)$ with normal $\mathbf{n} = \langle -7, 1, 1\rangle$.

(5) Find the equation of the plane through $Q(0,-4,12)$ containing the line

$$\frac{x-5}{-2}=\frac{y-3}{5}=\frac{z+1}{6}\,.$$

(6) By recognizing the symmetric equations of a line as an intersection of two planes, re-derive the vector equation of the line in terms of the normals of the associated planes.

(7) Describe the equation of the plane containing three given points P, Q and R using a 3×3 determinant.

(8) Given a point P and a plane Π with normal vector $\mathbf{n}$ and containing a point $\mathbf{r}_0$ but not containing P, find an expression for the point on Π nearest to P. Writing $\mathbf{r}_0 = x_0\hat{\mathbf{\imath}}+y_0\hat{\mathbf{\jmath}}+z_0\hat{\mathbf{k}}$, $\mathbf{p} = x_1\hat{\mathbf{\imath}}+y_1\hat{\mathbf{\jmath}}+z_1\hat{\mathbf{k}}$, and $\mathbf{n} = a\hat{\mathbf{\imath}}+b\hat{\mathbf{\jmath}}+c\hat{\mathbf{k}}$, write the formula out explicitly in coordinates.

(9) Find a formula for the reflection of a three dimensional vector $\mathbf{r}$ through a plane Π, with normal vector $\mathbf{n}$ containing a point $\mathbf{r}_0$. As in the previous problem, using $\mathbf{r}_0 = x_0\hat{\mathbf{\imath}}+y_0\hat{\mathbf{\jmath}}+z_0\hat{\mathbf{k}}$ and $\mathbf{n} = a\hat{\mathbf{\imath}}+b\hat{\mathbf{\jmath}}+c\hat{\mathbf{k}}$, write the formula out explicitly in coordinates. You may also describe it via a 3×3 matrix.

(10) Find the equation of a plane such that the region in the first octant, bounded between the plane and the coordinate planes forms a tetrahedron of volume 4 whose edges along the coordinate axes are all of equal length.

(11) Find two vectors of length 3 whose cross product is the vector $\mathbf{v} = 4\hat{\mathbf{\imath}}+4\hat{\mathbf{\jmath}}+7\hat{\mathbf{k}}$. What must the scalar triple product of these vectors with $\mathbf{v}$ be, regardless of which two vectors you find?

(12) Express the distance between the parallel planes $ax+by+cz+d=0$ and $ax+by+cz+e=0$ in terms of a, b, c, d, and e.

(13) Let $\mathbf{r}_1(t) = t\mathbf{v}_1+\mathbf{b}_1$ and $\mathbf{r}_2(s) = s\mathbf{v}_2+\mathbf{b}_2$ be skew lines in $\mathbb{R}^3$. Find explicit expressions for the pair of nearest points on these lines in terms of the vectors $\mathbf{v}_1$, $\mathbf{b}_1$, $\mathbf{v}_2$, and $\mathbf{b}_2$. Note that your expressions are explicitly independent of the distance between the two lines.

(14) Use vector algebra and geometric reasoning to prove that a system of three linear equations in 3 variables has a unique solution if and only if the determinant of the matrix of the equations' coefficients is nonzero. (Hint: think about normal vectors to planes, and try to envision ways a system with no solutions can arise.)

6. Geometry of rigid motions in the plane and 3-Space*

This is an optional section which builds up to the historical origins of vectors, via Hamilton's *quaternions*. It is heavier on algebraic thinking than the preceding material, but there is ample room for geometric intuition. One should at every opportunity try to draw pictures to capture the ideas laid out here. The section also has exercises mixed in, and then some challenge problems at the end. The exercises within the section are meant to encourage you to check claims being made and to actively engage with the ideas before you move on. The problems are more challenging and require greater mathematical maturity than those of preceding sections – though they are perhaps best suited to an honors project, any ambitious student interested in the geometry of space may find some portion of the problems quite interesting and useful.

§ 6.1. Translations, Rotations, and Reflections in the plane

Presently we consider how to use vectors and a little "light" linear algebra to model the so called *rigid motions* of the plane (in the spirit of Euclidean geometry.) That is, we wish to describe maps of the plane $\mathbb{R}^2$ to itself which preserve the Euclidean notions of angles, areas, and lengths, and thus give us the classical notion of *congruence* of planar objects in Euclidean geometry.

The simplest such motion to describe is a *translation*, which relocates all points by moving a constant distance in a given direction. Suppose for example we wanted to shift everything 3 units to the right (along the x-direction), and 2 units upwards (along the y-direction). If $\mathbf{r} = x\,\hat{\imath} + y\,\hat{\jmath}$ is a generic position vector, then its new position is $\mathbf{r} + 3\hat{\imath} + 2\hat{\jmath} = (x+3)\,\hat{\imath} + (y+2)\,\hat{\jmath}$. Thus, we see that translation is accomplished by adding a displacement vector to the position vector, and we can view this as determining a map $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(\mathbf{r}) = \mathbf{r} + \mathbf{b}$ where $\mathbf{b}$ is a constant vector giving the displacement of the translation.

Suppose we wanted to rotate a vector $\mathbf{r} \in \mathbb{R}^2$ by an angle φ. If $\mathbf{r} = x\,\hat{\imath} + y\,\hat{\jmath}$ makes an angle of θ with $\hat{\imath}$, then we can write $\mathbf{r} = r\cos\theta\,\hat{\imath} + r\sin\theta\,\hat{\jmath}$ where $r = \|\mathbf{r}\|$. It follows that the rotated vector has expression $\mathscr{R}_\varphi(\mathbf{r}) = r\cos(\theta+\varphi)\,\hat{\imath} + r\sin(\theta+\varphi)\,\hat{\jmath}$. We can re-express this in terms of x, y, and φ as follows: using the trigonometric addition identities, we can re-write $\mathscr{R}_\varphi(\mathbf{r})$:

$$\begin{aligned}
\mathscr{R}_\theta(\mathbf{r}) &= r\cos(\theta+\varphi)\,\hat{\imath} + r\sin(\theta+\varphi)\,\hat{\jmath} \\
&= \big(r\cos(\theta)\cos(\varphi) - r\sin(\theta)\sin(\varphi)\big)\,\hat{\imath} + \big(r\sin(\theta)\cos(\varphi) + r\cos(\theta)\sin(\varphi)\big)\,\hat{\jmath} \\
&= \big(x\cos(\varphi) - y\sin(\varphi)\big)\,\hat{\imath} + \big(y\cos(\varphi) + x\sin(\varphi)\big)\,\hat{\jmath} \\
&= \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.
\end{aligned}$$

The final line above reveals a 2×2 matrix which acts on $\mathbf{r}$ to achieve the rotation: $\mathscr{R}_\varphi(\mathbf{r}) = \mathrm{M}\mathbf{r}$ where M is the matrix whose first row is $\langle\cos(\varphi), -\sin(\varphi)\rangle$ and whose second row is $\langle\sin(\varphi), \cos(\varphi)\rangle$. (Recall, the matrix product acts by dotting these row vectors with $\mathbf{r}$ to obtain the entries of the transformed vector.) This map is a special example of a *linear map* of $\mathbb{R}^2$, and in fact, every linear map from the plane to itself can be represented using a 2×2 matrix. We can see that this is the correct transformation from a picture as well! See the following exercise.

Exercise: Draw an appropriate picture and use a right triangle to deduce the above formula for planar rotation about the origin. This was the content of problems (12) and (13) at the end of the first section. (Hint: By interpreting the columns of the rotation matrix as vectors, give a geometric interpretation of their meaning.)

Exercise: For each of the angles in $[0, \pi]$ corresponding to special right triangles, write down the corresponding rotation matrices.

We can combine translations and rotations to create more sophisticated motions of the plane. Try to convince yourself, however, that no combination of translations and rotations can achieve a reflection. To build up to reflection in a generic line, we first consider reflection in a coordinate axis, and build up to the general case.

Thus, consider the task of reflecting a point $\mathbf{r}$ through the x axis. The reflected vector $\mathfrak{Ref}_{\hat{\imath}}(\mathbf{r})$ is the vector obtained by negating the y component, which is the component perpendicular to the x axis;

$$\mathfrak{Ref}_{\hat{\imath}}(x\,\hat{\imath} + y\,\hat{\jmath}) = x\,\hat{\imath} - y\,\hat{\jmath} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Composing this transformation with appropriate rotations allows one to get a formula for a general reflection matrix through a line $\ell \subset \mathbb{R}^2$ through the origin: assuming the line makes an angle φ with the x-axis, one obtains (see the exercise below)

$$\begin{aligned}\mathfrak{Ref}_{\ell}(\mathbf{r}) &= \begin{bmatrix} \cos 2\varphi & \sin 2\varphi \\ \sin 2\varphi & -\cos 2\varphi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \big(x\cos(2\varphi) + y\sin(2\varphi)\big)\,\hat{\imath} + \big(x\sin(2\varphi) - y\cos(2\varphi)\big)\,\hat{\jmath}\,.\end{aligned}$$

Exercise: Describe a composition of appropriate rotations with reflection across the x-axis to prove the above formula for reflection across the line ℓ making angle φ with the x-axis. Then, show that any rotation of $\mathbb{R}^2$ about the origin can be achieved by computing a composition of two reflections across lines through the origin. How does the angle of rotation relate to the angles of separation between the lines?

By drawing a picture, one can also approach this problem purely from a vector algebra perspective (without relying on rotations and matrix algebra). Indeed, one can show that reflection through the line ℓ is given by

$$\mathfrak{Ref}_{\ell}(\mathbf{r}) = 2(\hat{\mathbf{u}} \cdot \mathbf{r})\hat{\mathbf{u}} - \mathbf{r}\,,$$

where $\hat{\mathbf{u}}$ is a unit vector parallel to ℓ.

Exercise: Draw an appropriate abstract picture and use vector algebra to prove the above formula.

Exercise: Explicitly describe the reflections through the lines $y = \sqrt{3}x$ and $y = \frac{3}{4}x$.

Exercise: Draw an appropriate abstract picture and use vector algebra to find a formula for reflection across the line with equation $ax + by = c$. (Hint: rewrite the line's equation using a vector product, with $\mathbf{r} = x\hat{\imath} + y\hat{\jmath}$. What's the geometric interpretation of this equation? Use this interpretation to describe a solution. By parametrizing the line obtain a solution in coordinates. If you get seemingly different expressions, reconcile them!)

---

§ 6.2. Complex multiplication and planar geometry

Recall that by i we denote a solution of the equation $x^2 + 1 = 0$, and call i *the imaginary unit.* Thus, we regard i as $\sqrt{-1}$, whence $(\pm i)^2 = -1$. We wish to consider the set of *complex numbers* $\mathbb{C}$ which can be regarded as all real linear-combinations of 1 and i:

$$\mathbb{C} := \{a + bi \,|\, a, b \in \mathbb{R}\}\,.$$

We define *the real and imaginary parts* of $a + bi$ in the obvious way: the real part is a and the imaginary part is b.

Thinking of 1 and i as being vectors (playing the roles of $\hat{\mathbf{\imath}}$ and $\hat{\mathbf{\jmath}}$), we get a vector space structure on $\mathbb{C}$, with complex numbers adding in precisely the same way as two dimensional real vectors. This one-to-one correspondence will be denoted by writing $a\,\hat{\mathbf{\imath}}+b\,\hat{\mathbf{\jmath}} \sim a+bi$. We can draw a complex number just as we would a real position vector sitting in $\mathbb{R}^2$ with its tail at the origin (the arrow head is conventionally omitted, and sometimes the line segment itself, too, except for pedagogical reasons). The x-axis becomes "the real axis" $\mathbb{R}$ sitting inside $\mathbb{C}$ as a subset (and as a vector subspace, whatever that means!) and the y-axis becomes "the imaginary axis" $i\mathbb{R}$.

We can now use complex arithmetic to immediately recover many of the previous rigid motions! For example, translations are accomplished simply by complex addition!

How about rotations? Multiplying by i sends 1 to i, and sends $x + yi$ to $xi - y \sim -y\,\hat{\mathbf{\imath}} + x\hat{\mathbf{\jmath}} = \mathscr{R}_{\pi/2}(\mathbf{r})$, which is merely rotation by $\pi/2$ radians. For a general rotation, one merely needs to multiply by $\cos\varphi + i\sin\varphi = e^{i\varphi}$. This final equality is a remarkable fact (known as Euler's formula), which you are encouraged to investigate in the exercises below. This allows us also to express complex numbers in *polar form*: $z = x + yi = r\cos\theta + ir\sin\theta = re^{i\theta}$, where $r^2 = |z|^2 = x^2 + y^2$ is the *square of the modulus* $r = |z|$. The angle θ is called *the argument of* z: $\theta = \arg(z)$.

Exercise: Use the Maclaurin series for the natural exponential function $f(x) = e^x$ to derive Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$, assuming that such expressions are well defined. How does one determine convergence properties for such a series? Can you argue (at least informally) that the series converges for all $\theta \in \mathbb{R}$?

Exercise: Check that the usual rules of exponentiation work for imaginary powers of e by using Euler's formula and appropriate trigonometric identities. What happens when you take roots of imaginary powers, i.e., what happens with fractional exponents? (Hint: You must take care to account for the fact that different values θ_1, θ_2 may yield the same complex number $e^{i\theta_1} = e^{i\theta_2}$. How do such values differ?)

Exercise: Argue Euler's formula in the following way: consider the differential equation $\frac{\mathrm{d}z}{\mathrm{d}t} = iz$. Here, the complex number z represents position in $\mathbb{C}$. By appealing to geometry of motion and the correspondence between $\mathbb{C}$ and $\mathscr{V}^2_{\mathbb{R}}$, argue that the solution to the initial value problem $\frac{\mathrm{d}z}{\mathrm{d}t} = iz$, $z(0) = 1$ is necessarily $z(t) = \cos t + i\sin t$, and then appeal to the notion that $f(t) = \exp(kt) = e^{kt}$ is definable as the unique solution to the initial value problem $\frac{\mathrm{d}f}{\mathrm{d}t} = kf$. In particular, you should check that such a solution satisfies the known properties of an exponential function, as does the complex trigonometry expression for the proposed solution $z(t)$.

The polar form confers an immediate advantage in understanding complex multiplication geometrically: the product of two complex numbers z and w has length equal to the product $|z||w|$ and has an argument equal the sum of the arguments: $\arg(zw) = \arg(z) + \arg(w)$. That is, we can geometrically interpret the complex-linear map $z \mapsto wz$ for a fixed w as a rotation and dilation of the complex plane, with rotation angle $\arg(w)$ and dilation factor $|w|$.

Exercise: For $z = x + iy = re^{i\theta}$ and $w = u + iv = se^{i\varphi}$, verify that $|zw| = |z||w|$ and $\arg(zw) = \arg(z) + \arg(w)$, using the polar form and properties of exponents. Then check by multiplying out the expressions $(x + iy)(u + iv)$ and $(r\cos\theta + ir\sin\theta)(s\cos\varphi + is\sin\varphi)$.

The operation of complex conjugation acts on a complex number by negating its imaginary component: $\overline{x+iy} = x - iy$, and thus geometrically is a reflection through the real axis: $\mathfrak{Ref}_{\mathbb{R}}(z) = \bar{z}$. It is now quite easy to recover the formula for general reflection across a line through the origin: if the line is represented by $\ell(t) = te^{i\varphi}$ for real parameter t and constant φ representing the angle made by ℓ with the real axis, and if $z = x + iy$ is a generic point of $\mathbb{C}$, then

$$\begin{aligned}
\mathfrak{Ref}_\ell(z) &= \mathscr{R}_\varphi \circ \mathfrak{Ref}_{\mathbb{R}} \circ \mathscr{R}_{-\varphi}(z) \\
&= \mathscr{R}_\varphi \circ \mathfrak{Ref}_{\mathbb{R}}(e^{-i\varphi} z) \\
&= \mathscr{R}_\varphi\left(\overline{e^{-i\varphi} z}\right) = e^{i\varphi}\left(\overline{e^{-i\varphi} z}\right) \\
&= e^{i\varphi}(e^{i\varphi}\bar{z}) = e^{2i\varphi}\bar{z} \\
&= x\cos(2\varphi) + y\sin(2\varphi) + i\big(x\sin(2\varphi) - y\cos(2\varphi)\big)\,.
\end{aligned}$$

Exercise: Show that $z\bar{z} = |z|^2$. Use this to check that complex division is well defined, in particular, by calculating an expression for the real and imaginary parts of $1/z$ for $z = x + iy$.

Exercise: Check that the real part of $z = x+iy$ is equal to $\frac{z+\bar{z}}{2}$, and find a similar expression for the imaginary part (remember, the imaginary part is a real number!); then draw a pair of pictures clarifying the geometric reasons that these formulae work.

Exercise: Recall, a similarity transformation of the plane preserves the magnitude of angles and proportions of plane figures, but not necessarily sizes. Argue that any similarity transformation $T : \mathbb{C} \to \mathbb{C}$ can be described by one of the following formulae, for *complex numbers* a and b:

$$T(z) = az + b \quad \text{if} \quad T \text{ is conformal}$$

$$T(z) = a\bar{z} + b \quad \text{if} \quad T \text{ is anti-conformal}$$

(See problem 1 below for the definition of conformal; for this exercise, it is merely a question of whether or not the transformation is *orientation preserving.*)

Exercise: Argue, first with a geometric visualization and then algebraically, that the product $\bar{z}w$, for $z = x + yi \sim x\hat{\mathbf{i}} + y\hat{\mathbf{j}} = \mathbf{z}$ and $w = u + vi \sim u\hat{\mathbf{i}} + v\hat{\mathbf{j}} = \mathbf{w}$, satisfies the equation

$$\bar{z}w = \mathbf{z}\cdot\mathbf{w} + i(\mathbf{z}\times\mathbf{w})\cdot\hat{\mathbf{k}}\,.$$

Exercise: Give geometric interpretations of the maps $z \mapsto 1/\bar{z}$ and $z \mapsto 1/z$ by considering the previous exercises. (Hint: use the polar form of z, and consider what sets are invariant under each map.)

---

§ 6.3. Spatial motions via vector algebra

The spatial motions we wish to describe are the rigid geometric motions of $\mathbb{R}^3$ which determine the notion of congruence of surfaces and solids in 3-dimensional Euclidean geometry. As in the two dimensional case, we will encounter translations, reflections, and rotations.

As before, translations merely involve the addition of a constant vector. The group of translations can thus be identified with the space of 3-dimensional vectors $\mathscr{V}^3_{\mathbb{R}}$, and thus with $\mathbb{R}^3$ itself. As in the 2-dimensional case, the reflections can generate rotations. Nevertheless, we'll take some time. to study the algebra of rotations separately, after investigating reflections.

Let Π be a plane in $\mathbb{R}^3$ with normal vector $\mathbf{n}$ and containing the point $\mathbf{r}_0$. If $\mathbf{r}$ is the position vector for a point $R(x, y, z)$ not on the plane, then to reflect R across Π, we must subtract $2\mathrm{proj}_{\mathbf{n}}(\mathbf{r}-\mathbf{r}_0)$ from $\mathbf{r}$ to obtain the position of the reflected point $\mathfrak{Ref}_\Pi(R)$. Thus, we have the vector formula

$$\mathfrak{Ref}_\Pi(\mathbf{r}) = \mathbf{r} - 2\mathrm{proj}_{\mathbf{n}}(\mathbf{r}-\mathbf{r}_0)\,.$$

Exercise: If Π has scalar equation $ax+by+cz+d=0$, rewrite the above formula for $\mathfrak{Ref}_\Pi(\mathbf{r})$ in terms of a, b, c, and d by writing the projection operator as a matrix dependent only on the plane equation and subsequently expressing the reflection operator as a matrix. In particular, express the result of reflection on $R(x, y, z)$ in coordinates, using only x, y, z, a, b, c, and d. C.f., problem (9) of §5.3.

We now consider the problem of studying spatial rotations via vector algebra. First, consider what is needed to describe a spatial rotation centered about the origin: we need an axis of rotation, which can be described via any vector $\mathbf{u} \in \mathscr{V}^3_{\mathbb{R}}$, and we need an angle φ. By convention, since we use right handed coordinates where the established positive orientation for angles is counterclockwise, we let $\varphi > 0$ be measured as a counterclockwise rotation as viewed from the tip of the axial vector, looking towards the center of rotation. We can assume without loss of generality that the axial vector is a unit vector $\hat{\mathbf{u}}$.

Let $\mathbf{r}$ be an arbitrary position vector, and $\mathscr{R}^{\hat{\mathbf{u}}}_\varphi(\mathbf{r})$ be the corresponding rotated position vector, after rotating by an angle φ about the axis $\mathbf{a}(t) = t\hat{\mathbf{u}}$, $t \in \mathbb{R}$.

Exercise: Draw a picture relating $\mathbf{r}$ to the axis, and determine an appropriate decomposition of $\mathbf{r}$ relative to the axis. It may be helpful to draw the radius $\|\mathbf{r}\|$ sphere centered at the origin, and imagine how the rotation acts on the sphere. USe this decomposition to deduce the following formula:

$$\mathscr{R}^{\hat{\mathbf{u}}}_\varphi(\mathbf{r}) = \big(1-\cos(\varphi)\big)\mathrm{proj}_{\hat{\mathbf{u}}}(\mathbf{r}) + \cos(\varphi)\mathbf{r} + \sin(\varphi)\hat{\mathbf{u}}\times\mathbf{r}\,.$$

This is the Rodriguez formula for spatial rotation in modern notation, named for Olinde Rodriquez, who worked out the components for angle-axis rotation and published the results in the 1840 paper *Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire*, which was largely ignored until long after Rodriguez's death.[13]

---

§ 6.4. Quaternions

The 19th century mathematician William Rowan Hamilton wished to replicate for 3 dimensions the success of complex numbers in algebraically encoding the Euclidean geometry of the plane. He

[13]The title of Rodriquez's paper may be translated as "Geometrical laws that govern the displacements of a solid system in space, and the variation of the coordinates resulting from these displacements considered independently of the causes that can produce them."

thus sought a collection of *hypercomplex numbers*. After years of trying and failing to turn $\mathbb{R}^3$ into something analogous to $\mathbb{C}$, he realized that 3 dimensions was doomed to fail, but 4 dimensions could work! Better yet, the algebra of 3 dimensional spatial rotations *is encoded* in the four dimensional system that he uncovered.[14]

Let us first unravel what it is that makes $\mathbb{C}$ special algebraically, and just why we can't have a $3D$ system of hypercomplex numbers. In the modern language, what Hamilton desired was the structure of a *normed division algebra over the real numbers.* Just what does that mean?

Observe that $\mathbb{C}$ can be regarded as a vector space over $\mathbb{R}$, its elements behaving isomorphically like the 2-vectors of $\mathbb{R}^2$. Additionally, $\mathbb{C}$ is endowed with a *norm*, namely the modulus function $z \mapsto |z| = \sqrt{z\bar{z}}$, which in this case corresponds to the Euclidean notion of length. Finally, it earns the title of division algebra because we can multiply any two complex numbers, and there exists a well defined inverse for each nonzero complex number. This multiplication is compatible with addition: it distributes and factors as one might hope. Additionally, the product structure is commutative and associative, but as we shall see, one can't always guarantee such things about a division algebra.

The geometry we recover via complex algebra is that of *dilative rotations*, also called *similarity transformations*, as they take planar figures to similar figures (meaning, angles and proportions are preserved). Thus, an ideal hypercomplex algebra modeling geometry in 3 dimensions would also capture dilative rotations via its hypercomplex multiplication.

Exercise: Show that there can be no 3-dimensional hypercomplex number system whose product acts geometrically by dilative rotations. (HInt: by considering the discussion of rotations above in §6.3, calculate the degrees of freedom necessary to specify an arbitrary dilative rotation.)

Hamilton's idea for the Quaternions was born on October 16, 1843, while Hamilton walked with his wife along the canals of Dublin, Ireland. After so long musing on the potentials of hypercomplex numbers in 3-dimensions, he realized an equation for a system with three hypercomplex units, thus making a 4-dimensional real division algebra. The fundamental relation that came to him on the walk was

$$i^2 = j^2 = k^2 = ijk = -1\,.$$

Here, he uses i, j, and k as three independent stand-ins for square roots of -1. Thus, there would be *six* distinct such roots, $\pm i, \pm j, \pm k$. The story goes that Hamilton was so excited at this breakthrough that he whipped out his pen knife and carved the fundamental relation of the quaternions into the Brougham Bridge crossing the canal. A commemorative plaque with these equations sits on the bridge in the present day.

Exercise: Use the fundamental relation to deduce a multiplication table for the quaternionic units $\pm i, \pm j, \pm k, \pm 1$. Hint: try taking pairs of equations from the relation, such as $ijk = -1$ and $i^2 = -1$; multiplying both sides by i on the left, deduce a relationship between j times k and i. Observe that the product is *not commutative*, like matrix products and unlike complex and real multiplication. Compare your multiplication table to the rules for taking cross products of the vectors $\hat{\mathbf{\imath}}$, $\hat{\mathbf{\jmath}}$, and $\hat{\mathbf{k}}$.

[14] Credit for the historical understanding of spatial rotations also belongs to Olinde Rodriguez and Leonhard Euler, but the quaternions sprouted many developments in notation and algebra across the nineteenth century. For an interesting account of the history and controversy around quaternions, see the article *Hamilton, Rodrigues, and the Quaternion Scandal* by Simon Altmann, published in 1989 in the MAA's *Mathematics Magazine*.

We can define the full algebra of real quaternions as the associative normed division algebra whose elements are real-linear combinations of 1, i, j, and k:

$$\mathbb{H} := \{t + xi + yj + zk : t, x, y, z \in \mathbb{R}\},$$

which has the natural vector space structure isomorphic to $\mathscr{V}_{\mathbb{R}}^4$, with norm coinciding with the euclidean norm, together with the associative and distributive product structure determined by linearly extending the product determined by the fundamental relation. In particular, if $q_1 = s + ui + vj + wk$ and $q_2 = t + xi + yj + zk$, then

$$\begin{aligned} q_1 q_2 = st &- ux - vy - wz \\ &+ (sx + ut + vz - wy)\, i \\ &+ (sy - uz + vt + wx)\, j \\ &+ (sz + uy - vx + wt)\, k\,. \end{aligned}$$

Exercise: Using the multiplication table built in the previous exercise, together with assumptions of associativity and distributivity for quaternions and the same plus commutativity for scalars, derive the above product expression by multiplying out $q_1q_2 = (s + ui + vj + wk)(t + xi + yj + zk)$.

Exercise: Compute the reverse order product q_2q_1, and compare it with the above, that is compute the difference $[q_1, q_2] := q_1q_1 - q_2q_1$, which is called the *commutator bracket of* q_1 and q_2. A product structure on a vector space is of course commutative if and only if all commutator brackets are zero. Determine under what conditions two quaternions commute.

Quaternions, like complex numbers, can be endowed with a conjugation operation: if $q = t + x\,i + y\,j + z\,k$, then $\bar{q} = t - x\,i - y\,j - z\,k$. There is a splitting of a quaternion into a real or *scalar part* and a hypercomplex, or *vector, part*:

$$q_s = \frac{q + \bar{q}}{2} = t\,, \quad \mathbf{q}_v = \frac{q - \bar{q}}{2} = x\,i + y\,j + z\,k \sim x\,\hat{\mathbf{\imath}} + y\,\hat{\mathbf{\jmath}} + z\,\hat{\mathbf{k}}\,.$$

Observe that the hypercomplex part is best treated as being analogous to a vector, and is not realizable as a scalar, unlike the imaginary part of a complex number.

Hamilton introduced the term vector to describe the spatial, hypercomplex piece of a quaternion, before the advent of modern vector algebra. It is from this piece that we acquire the notation $\hat{\mathbf{\imath}}, \hat{\mathbf{\jmath}}, \hat{\mathbf{k}}$ for the standard basis of $\mathscr{V}_{\mathbb{R}}^3$. One can exploit the idea of a quaternion as "a sum of a scalar and a vector" to better understand the quaternion product.

Exercise: Let $q = q_s + \mathbf{q}_v$ and $p = p_s + \mathbf{p}_v$ be two arbitrary quaternions. Show that

$$qp = q_s p_s - \mathbf{q}_v \cdot \mathbf{p}_v + q_s \mathbf{p}_v + p_s \mathbf{q}_v + \mathbf{q}_v \times \mathbf{p}_v\,,$$

where the dot and cross products perform as usual on 3-vectors, using the equivalence $\mathbf{q}_v \sim x\,\hat{\mathbf{\imath}} + y\,\hat{\mathbf{\jmath}} + z\,\hat{\mathbf{k}}$ for $q = t + x\,i + y\,j + z\,k$. Reinterpret the previous exercise on commutator brackets using this formula. What does the non-vanishing of the commutator of quaternions tell us, geometrically, about *how* they fail to be commutative?

The norm, or quaternionic modulus, is computed as one expects:

$$|q| = \sqrt{q\bar{q}} = \sqrt{q_s^2 + \|\mathbf{q}_v\|^2} = \sqrt{t^2 + x^2 + y^2 + z^2}\,.$$

Now, we need to check that division is actually well defined. We will avoid using fraction notation with quaternions, except when dividing by scalars, because the non-commutative nature of the product makes this notation ambiguous. Thus, we wish to show that to each $q \in \mathbb{H}$ there exists a q^{-1} such that $qq^{-1} = q^{-1}q = 1$. This is accomplished below in the exercises, in analogy to the construction of an inverse for complex multiplication.

Exercise: Demonstrate the ambiguity mentioned above in using fraction notation.

Exercise: Show that q^{-1} may be defined as $\bar{q}/|q|^2$.

The problems below explore the use of complex numbers and quaternions to understand geometry in 2 and 3 dimensions.

---

§ 6.5. Problems

(1) (Hard – requires familiarity with plane curves, partial derivatives and comfort with multivariable limits; see the subsequent batch of notes)

A map is called *conformal* if it preserves both the magnitude and sense (clockwise or counter-clockwise) of all angles of tangency between curves. In particular, a transformation $T : \mathbb{C} \to \mathbb{C}$ is said to be conformal at a point p if and only if whenever two differentiable curves $\gamma_1, \gamma_2 \subset \mathbb{C}$ intersect at p, the angle between the tangent vectors of the curves at p (measured in the parallelogram they span) equals the angle of separation of the tangent vectors of the image curves $T(\gamma_1)$ and $T(\gamma_2)$ at the image point of intersection $T(p)$ (again measured in the parallelogram they span), with the same sense (meaning, if the angle is swept counterclockwise from the tangent vector of γ_1 to that of γ_2, then so too is the angle swept counterclockwise from the tangent vector of $T(\gamma_1)$ to that of $T(\gamma_2)$.) If a transformation $T : \mathbb{C} \to \mathbb{C}$ preserves the magnitudes of angles but reverses their sense, it is called *anti-conformal*.

(a) Let $T(z) = T(x, y)$ be a transformation of $\mathcal{C}$ that is continuously real-differentiable with respect to x and y at a point p. Suppose further that $\partial_x T(p) + i\partial_y T(p) = 0$, and neither T_x nor T_y is zero at p. Show that T is conformal at p.

(b) Show that $\bar{z}$ is anti-conformal throughout $\mathbb{C}$.

(c) Define
$$T'(z_0) = \lim_{z \to z_0} \frac{T(z) - T(z_0)}{z - z_0}\,.$$
Show that if $T'(z_0)$ exists and is nonzero, then T is *conformal at* z_0, and conversely, if T is a conformal transformation on a set $D \subset \mathbb{C}$ with continuous real partial derivatives, then $T'(z)$ exists and is nonzero throughout D.

(2) Prove that every spatial rotation is the composition of two reflections. For a rotation by φ centered at a point with position $\mathbf{p} \in \mathbb{R}^3$, about an axis $\mathbf{u}$, give an explicit formula for the normals of a pair of planes Π_1 and Π_2 such that the rotation is realized by the composition $\mathfrak{Ref}_{\Pi_2} \circ \mathfrak{Ref}_{\Pi_1}$. Do this exercise without quaternions or the results of later problems. See also problem (10).

(3) Re-express the Rodriquez formula using a 3×3 matrix. Observe that the columns form an *orthonormal basis*: they are each unit vectors, and mutually orthogonal. Now argue that any orthonormal 3×3 matrix with positive determinant determines a unique spatial rotation, and describe how one can determine an axis and angle of the rotation from the columns of this *special orthogonal matrix*.

(4) Prove or disprove the following claim: given a pair of skew lines in $\mathbb{R}^3$, there is a rigid motion swapping the two lines.

(5) Taking the scalar-plus-vector derived formula for quaternion multiplication as definition, show that the product is indeed associative, despite the fact the cross products are not associative.

(6) This problem explores how one can encode spatial rotations using quaternions.
   (a) A *versor* is a unit quaternion, i.e., a quaternion q with $|q| = 1$. A *pure quaternion* is a quaternion whose scalar part is 0, and thus points of $\mathbb{R}^3$ can be identified with pure quaternions in the obvious way, via their position vectors. Deduce first that any versor can be expressed as
$$\cos(\varphi/2) + \sin(\varphi/2)\hat{\mathbf{u}}$$
for some angle $\varphi \in [0, 2\pi)$, and where $\hat{\mathbf{u}}$ is a unit pure quaternion, i.e., a pure versor.
   (b) Let p be a pure quaternion, and q a versor as above. Multiply out qpq^{-1} and compare to the Rodriquez formula to deduce that the map $p \mapsto qpq^{-1}$ models a spatial rotation. How does one determine the axis and angle for the transformation? Is the corresponding rotation uniquely represented?
   (c) Show that computing the spatial rotation with quaternions is more computationally efficient than using an orthogonal matrix, by showing that there are fewer elementary operations (addition/subtractions, multiplications and divisions) necessary to compute the effect of a spatial rotation on a vector.
   (d) Derive an expression for a rotation matrix in terms of the components of a versor representing the spatial rotation.

(7) These problems depend upon the previous collection of results on rotation.
   (a) Give a geometric interpretation of the space of all versors, and the space of unit versors inside it.
   (b) Prove a polar form for quaternions, in analogy to the polar form of complex numbers provided by Euler's famous formula $e^{i\theta} = \cos\theta + i\sin\theta$.
   (c) Let q be an arbitrary quaternion, and p a pure quaternion. Show that the map $p \mapsto qp\bar{q}$ represents a dilative rotation.

(8) Consider a product of rotations, represented by a pair of versors v_1 and v_2. Derive formulae for the axis of rotation and the angle of the composite rotation in terms of the axes and angles of the original versors. By geometrically interpreting the algebraic result in terms of the unit parts of the versors, according to the result of (a) of the preceding problem, rephrase the result in terms of triangles in spherical geometry.

(9) Let p and q be pure quaternions, corresponding to 3-vectors $\mathbf{p}$ and $\mathbf{q}$.
   (a) Show that $pq + qp = 0$ if and only if p and q represent orthogonal vectors.
   (b) Show that if q is additionally a unit quaternion orthogonal to p, then $p = qpq$.
   (c) Show that the map on $\mathbb{R}^3$ induced by sending an arbitrary pure quaternion p to upu for some fixed unit pure quaternion u acts geometrically by reflection in a plane. In particular, you must check that the map $p \mapsto upu$ induces a well defined map from $\mathbb{R}^3$ to itself (so upu must also be a pure quaternion).

(10) Reprove that any spatial rotation is a composition of two reflections, using quaternions. Explicitly demonstrate the relationship between the axis of rotation and the quaternions representing the reflections.

(11) (For students of linear algebra) Find a way to associate a matrix to any quaternion such that, for quaternions q and q' with Q and Q' the respective matrices,
- the quaternion sum $q + q'$ has matrix $Q + Q'$,
- the quaternion product qq' is realized by the matrix product QQ'

Can you instead find a matrix where qq' is represented by $Q'Q$?

(12) (Hard - requires some topological background and intuition) Argue (informally) that the space of all spatial rotations can be identified with the space of all lines through the origin in $\mathbb{R}^4$. That is, the special orthogonal group SO(3) of 3×3 orthogonal matrices with determinant 1, is *diffeomorphic* to the real projective space $\mathbb{RP}^3$. How does the space of versors relate to the space of rotations?