Chapter 1

Points and Vectors

1.1 Introduction

Points and vectors are fundamental in any discussion of geometry. They allow us to describe more complex geometric objects, and perform computations on them. For the most part, we will be discussing points and vectors in three-dimensional space. We will take the concepts of points and vectors as being intuitive, rather than trying to define them rigorously. A **point** is an object that defines a position (or location) in 3D space. A **vector** is an object that has both magnitude and direction, but no fixed location. A vector can be represented by a directed line segment or "arrow" that is floating freely in space. Alternatively, a vector can be regarded as a displacement. When programming, it is natural to represent both points and vectors as triples of numbers, which might suggest that they are the same sorts of things. But, as we will see, they are very different. Informally, a point is a dot; it has a fixed location, but no size. A vector is a free-floating arrow; it has a length and a direction, but no fixed location.

We will use upright bold-face letters, usually upper-case ones, to denote both points and vectors. So, typically, symbols like $\mathbf{u}$, $\mathbf{V}$ and $\mathbf{W}$ will represent vectors, and symbols like $\mathbf{p}$ and $\mathbf{Q}$ will represent points. But, as usual in mathematics, the symbols we use do not have any fixed universal meaning, so a symbol like $\mathbf{A}$ or $\mathbf{R}$ or $\mathbf{X}$ might denote a vector, a point, or a matrix, depending on the context.

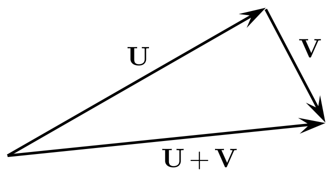
We will use the symbol $\mathcal{V}_3$ to denote the set of all 3D vectors, and the symbol $\mathcal{P}_3$ to denote the set of all 3D points. So the statement $\mathbf{V} \in \mathcal{V}_3$ just means that $\mathbf{V}$ is a 3D vector, and $\mathbf{X} \in \mathcal{P}_3$ means that $\mathbf{X}$ is a 3D point.

1.2 Equality of Points and Vectors

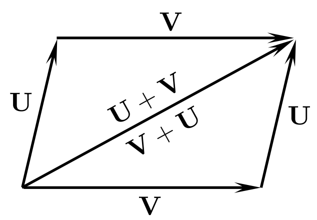
Two points are considered to be equal when they have the same location. Two vectors are considered to be equal when they have the same magnitude, direction, and sense. We mention the "sense" of the vector just for clarity. Some people would say that a vector and its reverse have opposite (and therefore unequal) directions. Others would say that these two vectors have the same direction, since they are parallel, but opposite sense. So, the word "direction" by itself, could be misinterpreted.

1.3 Addition of Vectors

Suppose we are given two vectors $\mathbf{U}$ and $\mathbf{V}$, and we wish to form their sum, $\mathbf{U} + \mathbf{V}$. We do this by placing the "tail" of $\mathbf{V}$ at the "head" of $\mathbf{U}$; the directed line segment from the tail of $\mathbf{U}$ to the head of $\mathbf{V}$ represents the sum $\mathbf{U} + \mathbf{V}$.



Alternatively, the displacement $\mathbf{U}+\mathbf{V}$ is equal to a displacement of $\mathbf{U}$ followed by a subsequent displacement of $\mathbf{V}$. Clearly $\mathbf{U}+\mathbf{V}$ is the same as $\mathbf{V}+\mathbf{U}$, since these two vectors are just two different ways of representing the diagonal of the same parallelogram, as shown below:



1.4 The Zero Vector

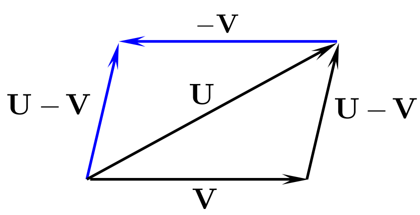
The zero vector is an arrow that has zero length. Or, alternatively, it's a displacement that involves no motion at all. We will denote the zero vector by a bold symbol $\mathbf{0}$. Clearly, adding the zero vector to a given vector $\mathbf{U}$ will have no effect, so $\mathbf{U}+\mathbf{0}=\mathbf{0}+\mathbf{U}=\mathbf{U}$.

1.5 Negation of Vectors

Given a vector $\mathbf{U}$, we define the negative of $\mathbf{U}$, denoted by $-\mathbf{U}$, to be the vector having the same magnitude and direction as $\mathbf{U}$, but opposite sense. Note that $\mathbf{U}+(-\mathbf{U})=(-\mathbf{U})+\mathbf{U}=\mathbf{0}$, where, as usual, $\mathbf{0}$ denotes the zero vector.

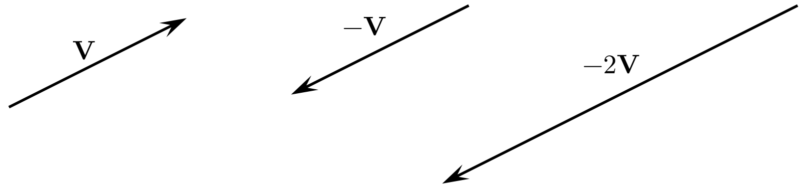
1.6 Subtraction of Vectors

We define $\mathbf{U}-\mathbf{V}$ to be $\mathbf{U}+(-\mathbf{V})$. Note that if $\mathbf{U}$ and $\mathbf{V}$ are relocated so that their "tails" are at the same point, then $\mathbf{U}-\mathbf{V}$ can be represented by the line segment from the head of $\mathbf{V}$ to the head of $\mathbf{U}$.



1.7 Multiplying a Vector by a Scalar

If $\mathbf{U}$ is a vector, and λ is a non-negative number, then we define $\lambda\mathbf{U}$ as the vector having the same direction and sense as $\mathbf{U}$, and a magnitude λ times that of $\mathbf{U}$. We call $\lambda\mathbf{U}$ the **scalar product** of λ and $\mathbf{U}$. Clearly, if $\lambda=0$, then $\lambda\mathbf{U}$ is the zero vector. If $\lambda<0$, then we define $\lambda\mathbf{U}$ to be $-(-\lambda)\mathbf{U}$.



To improve appearance, we sometimes write scalar products in the opposite order (with the vector before the scalar). For example, if $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are vectors and t is a scalar, then we might write $\mathbf{A}t^2 + \mathbf{B}t + \mathbf{C}$ because it looks better than $t^2\mathbf{A} + t\mathbf{B} + \mathbf{C}$.

1.8 Properties of Vector Algebra

Let's summarize the properties of vector addition and scalar multiplication. If $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ are vectors and λ and μ are numbers, then:

- $\mathbf{U} + \mathbf{V} = \mathbf{V} + \mathbf{U}$.
- $\mathbf{U} + (\mathbf{V} + \mathbf{W}) = (\mathbf{U} + \mathbf{V}) + \mathbf{W}$.
- $\mathbf{V} + \mathbf{0} = \mathbf{V}$, where $\mathbf{0}$ is the zero vector.
- There is a vector $-\mathbf{V}$, such that $\mathbf{V} + (-\mathbf{V}) = \mathbf{0}$.
- $1.\mathbf{V} = \mathbf{V}$.
- $\lambda(\mu\mathbf{V}) = (\lambda\mu)\mathbf{V}$.
- $(\lambda + \mu)\mathbf{V} = \lambda\mathbf{V} + \mu\mathbf{V}$.
- $\lambda(\mathbf{U} + \mathbf{V}) = \lambda\mathbf{U} + \lambda\mathbf{V}$

In mathematical terms, this means that the set of all vectors forms a **vector space** over the field of real numbers, as we will explain further in **??** . You can convince yourself that all of these statements are true by drawing simple diagrams. For example, we saw a diagram earlier, in section (1.3) that illustrates the relationship $\mathbf{U} + \mathbf{V} = \mathbf{V} + \mathbf{U}$. A mathematician would not regard these pictures as proofs, but they are adequate for our purposes. And, since we have not defined vectors rigorously, formal mathematical proofs are inappropriate, anyway.

1.9 Norm or Magnitude of a Vector

The **norm** or **magnitude** of a vector $\mathbf{U}$ is just its length, and is denoted by $\|\mathbf{U}\|$. The norm function has the following obvious properties

- If $\mathbf{U}$ is any vector, then $\|\mathbf{U}\| \geq 0$.
- $\|\mathbf{U}\| = 0$ if and only if $\mathbf{U} = \mathbf{0}$.
- If $\mathbf{U}$ is a vector and λ a scalar, then $\|\lambda\mathbf{U}\| = |\lambda| \cdot \|\mathbf{U}\|$.
- If $\mathbf{U}$ and $\mathbf{V}$ are vectors, then $\|\mathbf{U} + \mathbf{V}\| \leq \|\mathbf{U}\| + \|\mathbf{V}\|$.

In mathematical jargon, this means that the set of all 3D vectors forms a normed vector space.

1.10 Unit Vectors

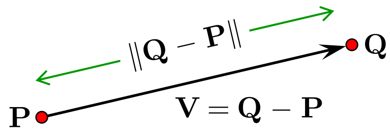
If $\|\mathbf{U}\| = 1$, then we say that $\mathbf{U}$ is a **unit vector**. If we are given a non-zero vector $\mathbf{U}$, then we can construct the vector

$$\hat{\mathbf{U}} = \frac{\mathbf{U}}{\|\mathbf{U}\|}$$

It is easy to show that $\hat{\mathbf{U}}$ is a unit vector that has the same direction and sense as $\mathbf{U}$. We say that $\hat{\mathbf{U}}$ is formed by **unitizing** $\mathbf{U}$.

1.11 Adding Points and Vectors

Suppose we are given a point $\mathbf{P}$ and a vector $\mathbf{V}$. Then we can produce a new point $\mathbf{Q}$ by applying the displacement $\mathbf{V}$ to the original point $\mathbf{P}$. We denote this new point by $\mathbf{P} + \mathbf{V}$, so we have $\mathbf{Q} = \mathbf{P} + \mathbf{V}$. Alternatively, if we are given two points $\mathbf{P}$ and $\mathbf{Q}$, then the displacement that would move $\mathbf{P}$ to $\mathbf{Q}$ is a vector, $\mathbf{V}$. Or, alternatively, $\mathbf{V}$ is the "arrow" that leads from $\mathbf{P}$ to $\mathbf{Q}$. In short, we have defined the difference of two points as a vector, so we write $\mathbf{Q} - \mathbf{P} = \mathbf{V}$. The length of this vector $\mathbf{Q} - \mathbf{P}$ then gives us the distance between the two points: $\text{Dist}(\mathbf{P}, \mathbf{Q}) = \|\mathbf{Q} - \mathbf{P}\|$.



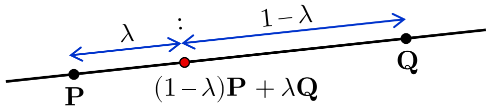
If we fix an origin point $\mathbf{O}$, then we can identify any given point $\mathbf{P}$ with the vector $\mathbf{V} = \mathbf{P} - \mathbf{O}$, which gives us a one-to-one correspondence between points and vectors. However, for many of our tasks, identifying some special origin point is somewhat artificial, and we'll usually avoid it. Identifying points and vectors in this way is attractive because it appears to give us the opportunity to define addition of points. Again, suppose we have chosen a fixed origin $\mathbf{O}$, and let $\mathbf{P}$ and $\mathbf{Q}$ be two points that we would like to add. We can identify $\mathbf{P}$ and $\mathbf{Q}$ with two vectors $\mathbf{U}$ and $\mathbf{V}$ such that $\mathbf{U} = \mathbf{P} - \mathbf{O}$ and $\mathbf{V} = \mathbf{Q} - \mathbf{O}$, and we could then declare that $\mathbf{P} + \mathbf{Q}$ is defined to be $\mathbf{U} + \mathbf{V}$. But, unfortunately, this definition is not really workable because it makes the value of $\mathbf{P} + \mathbf{Q}$ dependent on $\mathbf{O}$, which seems inappropriate. So, addition of points is not defined, except for some special cases discussed below.

1.12 Combinations of Points and Vectors

Let $\mathbf{V}_0$ and $\mathbf{V}_1$ be two given vectors, and consider the set of vectors of the form $\mathbf{V}(\lambda) = (1-\lambda)\mathbf{V}_0 + \lambda\mathbf{V}_1$, where λ is a number. These vectors are said to be **affine combinations** of $\mathbf{V}_0$ and $\mathbf{V}_1$. Rearranging this, we see that $\mathbf{V}(\lambda) = \mathbf{V}_0 + \lambda(\mathbf{V}_1 - \mathbf{V}_0)$. This means that the head of $\mathbf{V}(\lambda)$ lies on the line joining the heads of $\mathbf{V}_0$ and $\mathbf{V}_1$. We can perform a similar construction with points. Suppose $\mathbf{P}_0$ and $\mathbf{P}_1$ are two given points, and λ is a number. Then $\mathbf{P}_1 - \mathbf{P}_0$ is a vector, and we can write $\mathbf{P}(\lambda) = \mathbf{P}_0 + \lambda(\mathbf{P}_1 - \mathbf{P}_0)$. Again, $\mathbf{P}(\lambda)$ is said to be an **affine combination** of $\mathbf{P}_0$ and $\mathbf{P}_1$. Mimicing the usual rules of arithmetic, we can write

$$\mathbf{P}_0 + \lambda(\mathbf{P}_1 - \mathbf{P}_0) = (1-\lambda)\mathbf{P}_0 + \lambda\mathbf{P}_1$$

and we can use the expression on the left as the *definition* of the otherwise meaningless expression on the right. So, even though addition and scalar multiplication of points makes no sense, in general, we can define affine combinations of points in a reasonable way. We just have to be careful whenever we see an expression of the form $\lambda\mathbf{P} + \mu\mathbf{Q}$ involving points and numbers — we must remember that this expression makes sense only when $\lambda + \mu = 1$. An affine combination $\mathbf{P}(\lambda) = (1-\lambda)\mathbf{P}_0 + \lambda\mathbf{P}_1$ where $0 \le \lambda \le 1$ is called a **convex combination** of $\mathbf{P}_0$ and $\mathbf{P}_1$. In this case, the point $\mathbf{P}(\lambda)$ lies on the line segment $\mathbf{P}_0\mathbf{P}_1$, dividing it in the ratio $\lambda : 1-\lambda$, as shown here:



Convex combinations are the foundation of many of the constructions we will see in later chapters, so this is a very important concept.

1.13 General Linear, Affine, and Convex Combinations

In the previous section, we discussed linear, affine and convex combinations of two points or vectors. Now let's generalize this to the case of n points or vectors, where $n \ge 2$. Let $\mathbf{U}_1, \dots, \mathbf{U}_n$ be a given set of n vectors, let $\lambda_1, \dots, \lambda_n$ be n scalars, and let

$$\mathbf{U} = \lambda_1\mathbf{U}_1 + \lambda_2\mathbf{U}_2 + \cdots + \lambda_n\mathbf{U}_n$$

This expression is called

- a **linear combination** if the scalars $\lambda_1, \dots, \lambda_n$ are arbitrary.
- an **affine combination** if $\lambda_1 + \cdots + \lambda_n = 1$.
- a **convex combination** if $\lambda_1 + \cdots + \lambda_n = 1$ and $\lambda_1, \dots, \lambda_n$ are all non-negative. This implies that $\lambda_1, \dots, \lambda_n$ all lie in the interval $[0, 1]$.

Using the same technique we saw above, we can also define the affine and convex combinations of n points $\mathbf{P}_1, \dots, \mathbf{P}_n$. The expression $\lambda_1\mathbf{P}_1 + \cdots + \lambda_n\mathbf{P}_n$ makes sense provided that $\lambda_1 + \cdots + \lambda_n = 1$, because it can be defined by

$$\lambda_1\mathbf{P}_1 + \cdots + \lambda_n\mathbf{P}_n = \mathbf{P}_1 + \lambda_2(\mathbf{P}_2 - \mathbf{P}_1) + \cdots + \lambda_n(\mathbf{P}_n - \mathbf{P}_1)$$

1.14 Convex Sets

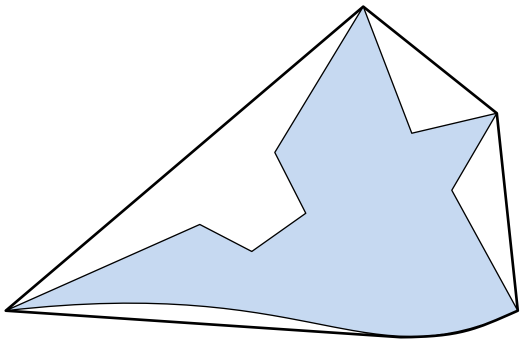
If $\mathbf{P}_0$ and $\mathbf{P}_1$ are two points, then the set

$$[\mathbf{P}_0, \mathbf{P}_1] = \{(1-\lambda)\mathbf{P}_0 + \lambda\mathbf{P}_1 : 0 \le \lambda \le 1\}$$

is called the **line segment** joining $\mathbf{P}_0$ and $\mathbf{P}_1$. A set S is said to be **convex** if whenever $\mathbf{P}_0$ and $\mathbf{P}_1$ belong to S, the line segment $[\mathbf{P}_0, \mathbf{P}_1]$ is also contained in S. It follows by induction that if the points $\mathbf{P}_1, \dots, \mathbf{P}_n$ belong to a convex set S, then all convex combinations of these points again belong to S.

1.15 Convex Hull

Let $\mathbf{P}_1, \ldots, \mathbf{P}_n$ be a finite set of points. Then the set of all convex combinations of $\mathbf{P}_1, \ldots, \mathbf{P}_n$ is called the convex hull of $\mathbf{P}_1, \ldots, \mathbf{P}_n$. Clearly the convex hull is a convex set that contains $\mathbf{P}_1, \ldots, \mathbf{P}_n$; in fact it can be shown that the convex hull is the "smallest" convex set that contains these points. If the given points are coplanar, their convex hull is the region bordered by an elastic band that has been snapped around the points, as shown here:



The convex hulls of small sets of points are easy to describe:

- the convex hull of two points is the line segment between them
- the convex hull of three points is a triangle
- the convex hull of four points is a tetrahedron

1.16 Linear Spans

Let $\mathbf{U}_1, \ldots, \mathbf{U}_n$ be a given set of n vectors. The collection of all vectors that can be written as linear combinations of $\mathbf{U}_1, \ldots, \mathbf{U}_n$ is called the **linear span** of $\mathbf{U}_1, \ldots, \mathbf{U}_n$. In some cases, this collection might be the whole of $\mathcal{V}_3$, in which case we say that $\mathbf{U}_1, \ldots, \mathbf{U}_n$ **span** $\mathcal{V}_3$. To show that some given vector $\mathbf{V}$ lies in the linear span of the vectors $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{U}_3$, we need to find three scalars λ_0, λ_1, λ_2 such that $\mathbf{V} = \lambda_1\mathbf{U}_1 + \lambda_2\mathbf{U}_2 + \lambda_3\mathbf{U}_3$. An exercise at this end of this chapter shows that this is always possible, provided that $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{U}_3$ are not coplanar. In other words, three vectors span $\mathcal{V}_3$ provided they are not coplanar.

1.17 Linear Independence

A set of vectors $\mathbf{U}_1, \ldots, \mathbf{U}_n$ are said to be **linearly dependent** if one of them can be written as a linear combination of some of the others; otherwise, they are said to be **linearly independent**. Informally, a linearly dependent set contains some superfluous vectors. These superfluous vectors add nothing to our ability to generate new vectors via linear combinations; in other words, they do not serve to increase the linear span of the set. If $\mathbf{U}_1, \ldots, \mathbf{U}_n$ are linearly dependent then any linear combination $\lambda_1\mathbf{U}_1 + \lambda_2\mathbf{U}_2 + \cdots + \lambda_n\mathbf{U}_n$ can be rewritten as a linear combination of some subset of $\mathbf{U}_1, \ldots, \mathbf{U}_n$. The more common mathematical definition says that vectors $\mathbf{U}_1, \ldots, \mathbf{U}_n$ are linearly dependent if there exists some non-trivial linear relationship between them. In other words, there exist $\lambda_1, \ldots, \lambda_n$, not all zero, such that

$\lambda_1\mathbf{U}_1 + \lambda_2\mathbf{U}_2 + \cdots + \lambda_n\mathbf{U}_n = 0$. Then, if $\lambda_k \neq 0$, we can write

$$\mathbf{U}_k = (\lambda_1\mathbf{U}_1 + \lambda_2\mathbf{U}_2 + \cdots + \lambda_{k-1}\mathbf{U}_{k-1} + \lambda_{k+1}\mathbf{U}_{k+1} + \cdots + \lambda_n\mathbf{U}_n)/\lambda_k$$

so $\mathbf{U}_k$ is a linear combination of $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_{k-1}, \mathbf{U}_{k+1}, \dots, \mathbf{U}_n$. An equivalent definition says that $\mathbf{U}_1, \dots, \mathbf{U}_n$ are linearly independent if

$$\lambda_1\mathbf{U}_1 + \lambda_2\mathbf{U}_2 + \cdots + \lambda_n\mathbf{U}_n = 0 \quad \Longrightarrow \quad \lambda_1 = \cdots = \lambda_n = 0$$

Suppose we have written some given vector $\mathbf{V}$ in two different ways as linear combinations of vectors $\mathbf{U}_1, \dots, \mathbf{U}_n$, say

$$\mathbf{V} = \lambda_1\mathbf{U}_1 + \lambda_2\mathbf{U}_2 + \cdots + \lambda_n\mathbf{U}_n$$
$$\mathbf{V} = \mu_1\mathbf{U}_1 + \mu_2\mathbf{U}_2 + \cdots + \mu_n\mathbf{U}_n$$

Then

$$(\lambda_1 - \mu_1)\mathbf{U}_1 + (\lambda_2 - \mu_2)\mathbf{U}_2 + \cdots + (\lambda_n - \mu_n)\mathbf{U}_n = 0$$

If $\mathbf{U}_1, \dots, \mathbf{U}_n$ are linearly independent, this implies that $\lambda_1 - \mu_1, \lambda_2 - \mu_2, \dots, \lambda_n - \mu_n$ are all zero, and so $\lambda_i = \mu_i$ for $i = 1, 2, \dots, n$. So, when using a linearly independent set of vectors, the expression of any given vector as a linear combination is *unique*.

1.18 Bases for $\mathcal{V}_3$

Suppose we have some set of n vectors $S = \{\mathbf{U}_1, \dots, \mathbf{U}_n\}$. What if we would like S to span $\mathcal{V}_3$ *and* to be linearly independent. The two requirements are in conflict: if we want S to span $\mathcal{V}_3$, then we should make it larger, and if we want S to be linearly independent, we should make it smaller. A set S that manages to satisfy these two conflicting requirements is called a **basis** for $\mathcal{V}_3$.

Let $S = \{\mathbf{U}_1, \dots, \mathbf{U}_n\}$ be a basis for $\mathcal{V}_3$, and suppose we add some new vector $\mathbf{V}$ to S. Then, since S spans $\mathcal{V}_3$, we can find scalars $\lambda_1, \dots, \lambda_n$ such that $\mathbf{V} = \lambda_1\mathbf{U}_1 + \lambda_2\mathbf{U}_2 + \cdots + \lambda_n\mathbf{U}_n$. But this means that the vectors $\mathbf{V}, \mathbf{U}_1, \dots, \mathbf{U}_n$ are linearly dependent. In other words, any new vector that we add to a basis will be superfluous and will therefore destroy its linear independence. Or, saying it another way, a basis is a *maximal* linearly independent set of vectors.

On the other hand, suppose we remove some vector $\mathbf{U}_k$ from a basis $S = \{\mathbf{U}_1, \dots, \mathbf{U}_n\}$. Since S was linearly independent, it is not possible to express $\mathbf{U}_k$ as a linear combination of the remaining vectors $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_{k-1}, \mathbf{U}_{k+1}, \dots, \mathbf{U}_n$. This means that $\mathbf{U}_k$ is not in the linear span of these remaining vectors, so they do not span $\mathcal{V}_3$. So, if we remove a vector from a basis, we destroy its ability to span $\mathcal{V}_3$. In other words, a basis is a *minimal* spanning set for $\mathcal{V}_3$.

So, we see that a basis is an exactly balanced compromise between the two requirements of spanning $\mathcal{V}_3$ and being linearly independent. If we add one more vector, we destroy linear independence, and if we remove one, we destroy the ability to span $\mathcal{V}_3$. Another way to summarize the properties of a basis S is as follows:

- because S spans $\mathcal{V}_3$, any vector in $\mathcal{V}_3$ can be written as a linear combination of vectors from S, and

- because S is linearly independent, no vector in $\mathcal{V}_3$ can be written as a linear combination of vectors from S in more than one way

So, a basis is large enough to provide a representation of every vector in $\mathcal{V}_3$ (as a linear combination of its members), but not so large that duplicate redundant representations exist.

1.19 Examples

If $\mathbf{U}$ and $\mathbf{V}$ are two non-zero vectors, then their linear span is the plane through the origin that is parallel to $\mathbf{U}$ and $\mathbf{V}$. If we transplant $\mathbf{U}$ and $\mathbf{V}$ so that their tails are at the origin, then they will lie within this plane.

If $\mathbf{P}$ and $\mathbf{Q}$ are two distinct points, then affine combinations of $\mathbf{P}$ and $\mathbf{Q}$ are vectors of the form $\mathbf{P}(\lambda) = (1-\lambda)\mathbf{P} + \lambda\mathbf{Q}$. These points lie on the (infinite) line through $\mathbf{P}$ and $\mathbf{Q}$. If $0 \le \lambda \le 1$, then $\mathbf{P}(\lambda)$ lies on the line *segment* $\mathbf{PQ}$ — in other words, between $\mathbf{P}$ and $\mathbf{Q}$.

Let $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$ be three vectors. Suppose the three vectors are not coplanar, which implies, in particular, that they are all different and that none of them is zero. Then their linear span is all of $\mathcal{V}_3$, and so they form a basis for $\mathcal{V}_3$.

Let $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$ be three distinct points, and $\lambda_0 + \lambda_1 + \lambda_2 = 1$. Then the affine combination $\mathbf{P}(\lambda_0, \lambda_1, \lambda_2) = \lambda_0\mathbf{P}_0 + \lambda_1\mathbf{P}_1 + \lambda_2\mathbf{P}_2$ is, by definition, given by the formula

$$\mathbf{P}(\lambda_0, \lambda_1, \lambda_2) = \mathbf{P}_0 + \lambda_1(\mathbf{P}_1 - \mathbf{P}_0) + \lambda_2(\mathbf{P}_2 - \mathbf{P}_0)$$

This is a point on the plane that contains the point $\mathbf{P}_0$ and is parallel to the two vectors $\mathbf{P}_1 - \mathbf{P}_0$ and $\mathbf{P}_2 - \mathbf{P}_0$. In other words, it's a point on the plane containing the three points $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$. If $\lambda_0, \lambda_1, \lambda_2 \ge 0$, then our affine combination is a convex combination and it lies inside the triangle $\mathbf{P}_0\mathbf{P}_1\mathbf{P}_2$.

Let $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ be four distinct points. Any other point can be written as an affine combination of $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ provided they are not coplanar. Convex combinations of $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$ lie inside the tetrahedron $\mathbf{P}_0\mathbf{P}_1\mathbf{P}_2\mathbf{P}_3$.

1.20 Orthogonal and Orthonormal Vectors

Two vectors $\mathbf{U}$ and $\mathbf{V}$ whose directions are perpendicular are said to be **orthogonal**. We sometimes denote this by writing $\mathbf{U}\perp\mathbf{V}$. Three vectors, $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ are said to be (mutually pairwise) orthogonal if $\mathbf{U}\perp\mathbf{V}$, $\mathbf{V}\perp\mathbf{W}$, and $\mathbf{W}\perp\mathbf{U}$. If three orthogonal vectors have unit length, they are said to be **orthonormal**. Any set of three orthogonal vectors (hence certainly any three orthonormal vectors) forms a basis for $\mathcal{V}_3$. In practice, the bases we use are almost always orthonormal.

1.21 Dot Product of Two Vectors

Let $\mathbf{U}$ and $\mathbf{V}$ be two vectors. The **dot product** of $\mathbf{U}$ and $\mathbf{V}$ is defined by

$$\mathbf{U}\cdot\mathbf{V} = \|\mathbf{U}\| \,.\, \|\mathbf{V}\| \,.\, \cos\theta$$

where θ is the angle between $\mathbf{U}$ and $\mathbf{V}$, with $0 \le \theta \le \pi$.

1.22 Properties of Dot Products

Let $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$, be three vectors and let λ be a scalar. Then

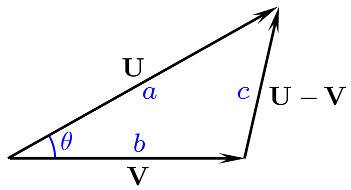
- $\mathbf{U}\cdot\mathbf{V} = \mathbf{V}\cdot\mathbf{U}$
- $\mathbf{U}\cdot(\mathbf{V}+\mathbf{W}) = \mathbf{U}\cdot\mathbf{V}+\mathbf{U}\cdot\mathbf{W}$
- $(\lambda\mathbf{U})\cdot\mathbf{V} = \mathbf{U}\cdot(\lambda\mathbf{V}) = \lambda(\mathbf{U}\cdot\mathbf{V})$
- $\mathbf{U}\cdot\mathbf{U} = \|\mathbf{U}\|^2$
- $|\mathbf{U}\cdot\mathbf{V}| \le \|\mathbf{U}\|.\|\mathbf{V}\|$
- If $\mathbf{U}\cdot\mathbf{V} = 0$, then either $\mathbf{U} = \mathbf{0}$, or $\mathbf{V} = \mathbf{0}$, or $\mathbf{U}$ and $\mathbf{V}$ are orthogonal.

Moreover, if θ is the angle between $\mathbf{U}$ and $\mathbf{V}$ $(0 \le \theta \le \pi)$, then

- $\mathbf{U}\cdot\mathbf{V} > 0 \iff \theta < \pi/2$
- $\mathbf{U}\cdot\mathbf{V} < 0 \iff \theta > \pi/2$

1.23 Dot Products and the Cosine Rule

Suppose the sides of a triangle are labelled as shown in the figure below. The symbols a, b, c denote the lengths of the sides.



Then, from the properties of dot products, we have

$$\|\mathbf{U}-\mathbf{V}\|^2 = (\mathbf{U}-\mathbf{V})\cdot(\mathbf{U}-\mathbf{V}) = \mathbf{U}\cdot\mathbf{U}+\mathbf{V}\cdot\mathbf{V}-2\mathbf{U}\cdot\mathbf{V} = \|\mathbf{U}\|^2+\|\mathbf{V}\|^2-2\mathbf{U}\cdot\mathbf{V}$$

But $\|\mathbf{U}\| = a$, $\|\mathbf{V}\| = b$, and $\|\mathbf{U}-\mathbf{V}\| = c$, so we get

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

which of course is the Law of Cosines familiar from trigonometry.

1.24 Dot Products of Orthonormal Vectors

If $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ are orthonormal vectors, then

$$\begin{array}{ccccc} \mathbf{E}_1\cdot\mathbf{E}_1 = 1 & ; & \mathbf{E}_1\cdot\mathbf{E}_2 = 0 & ; & \mathbf{E}_1\cdot\mathbf{E}_3 = 0 \\ \mathbf{E}_2\cdot\mathbf{E}_1 = 0 & ; & \mathbf{E}_2\cdot\mathbf{E}_2 = 1 & ; & \mathbf{E}_2\cdot\mathbf{E}_3 = 0 \\ \mathbf{E}_3\cdot\mathbf{E}_1 = 0 & ; & \mathbf{E}_3\cdot\mathbf{E}_2 = 0 & ; & \mathbf{E}_3\cdot\mathbf{E}_3 = 1 \end{array}$$

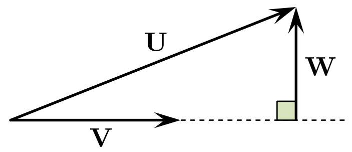
or, more briefly

$$\mathbf{E}_i \cdot \mathbf{E}_j = \delta_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } i \neq j \end{cases}$$

The symbol δ_{ij} is often called the **Kronecker delta** symbol.

1.25 Projections & Components

Let $\mathbf{U}$ and $\mathbf{V}$ be two given vectors. We wish to express $\mathbf{U}$ as the sum of two vectors, one of which is parallel to $\mathbf{V}$, and the other perpendicular to $\mathbf{V}$. In other words, we require $\mathbf{U} = \lambda\mathbf{V} + \mathbf{W}$ where $\mathbf{W}$ is perpendicular to $\mathbf{V}$.



Clearly $\mathbf{W} = \mathbf{U} - \lambda\mathbf{V}$, so we only need to find $\lambda\mathbf{V}$ (or λ). Now, the required magnitude for $\lambda\mathbf{V}$ is given by

$$\|\mathbf{U}\| \cos\theta = \frac{\|\mathbf{U}\| \, (\mathbf{U}\cdot\mathbf{V})}{\|\mathbf{U}\| \, . \, \|\mathbf{V}\|} = \frac{\mathbf{U}\cdot\mathbf{V}}{\|\mathbf{V}\|}$$

We now multiply this by a unit vector in the direction of $\mathbf{V}$, giving

$$\lambda\mathbf{V} = \frac{\mathbf{U}\cdot\mathbf{V}}{\|\mathbf{V}\|} \cdot \frac{\mathbf{V}}{\|\mathbf{V}\|} = \frac{\mathbf{U}\cdot\mathbf{V}}{\mathbf{V}\cdot\mathbf{V}}\mathbf{V}$$

and hence

$$\lambda = \frac{\mathbf{U}\cdot\mathbf{V}}{\mathbf{V}\cdot\mathbf{V}}$$

So the required decomposition of $\mathbf{U}$ is

$$\mathbf{U} = \frac{\mathbf{U}\cdot\mathbf{V}}{\mathbf{V}\cdot\mathbf{V}}\mathbf{V} + \left(\mathbf{U} - \frac{\mathbf{U}\cdot\mathbf{V}}{\mathbf{V}\cdot\mathbf{V}}\mathbf{V}\right)$$

The scalar λ (or sometimes the vector $\lambda\mathbf{V}$) is called the component of $\mathbf{U}$ parallel to $\mathbf{V}$, and $\mathbf{W}$ is called the component of $\mathbf{U}$ perpendicular to $\mathbf{V}$. It is easy to verify that $\mathbf{V}\cdot\mathbf{W} = 0$. In the case when $\mathbf{V}$ is a unit vector, we have

$$\lambda = \mathbf{U}\cdot\mathbf{V} = \|\mathbf{U}\| \cdot \|\mathbf{V}\| \cos\theta = \|\mathbf{U}\| \cos\theta$$

Informally, when the "reference" vector $\mathbf{V}$ has unit length, we can calculate components simply by taking dot products.

1.26 Example: Normal from Point to Line

1.27 Cross Product of Two Vectors

Let $\mathbf{U}$ and $\mathbf{V}$ be two vectors, and let θ be the angle between them, with $0 \leq \theta \leq \pi$. Then the cross product of $\mathbf{U}$ and $\mathbf{V}$ is defined by

$$\mathbf{U} \times \mathbf{V} = \|\mathbf{U}\| \, . \, \|\mathbf{V}\| \, .(\sin\theta).\mathbf{N}$$

where $\mathbf{N}$ is a unit vector perpendicular to the plane of $\mathbf{U}$ and $\mathbf{V}$ whose sense is that of a right-hand screw being turned from $\mathbf{U}$ to $\mathbf{V}$ through the angle θ.

1.28 Properties of Cross Products

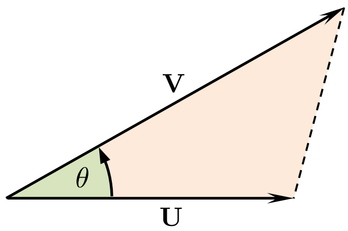
Let $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ be three vectors and let λ be a scalar. Then

- $\mathbf{U} \times \mathbf{V} = -\mathbf{V} \times \mathbf{U}$
- $\mathbf{U} \times (\mathbf{V} + \mathbf{W}) = \mathbf{U} \times \mathbf{V} + \mathbf{U} \times \mathbf{W}$
- $(\lambda\mathbf{U}) \times \mathbf{V} = \mathbf{U} \times (\lambda\mathbf{V}) = \lambda(\mathbf{U} \times \mathbf{V})$
- $\mathbf{U} \times \mathbf{V} = \mathbf{0} \iff \mathbf{U} = \mathbf{0}$, or $\mathbf{V} = \mathbf{0}$, or $\mathbf{U}$ and $\mathbf{V}$ are parallel.
- $\mathbf{U} \times \mathbf{U} = \mathbf{0}$
- $\mathbf{U} \cdot (\mathbf{U} \times \mathbf{V}) = \mathbf{V} \cdot (\mathbf{U} \times \mathbf{V}) = 0$

The first relationship needs special attention — the order of the vectors in a cross product is important, unlike most multiplication operations that you may have seen before.

1.29 Cross Product as an Area

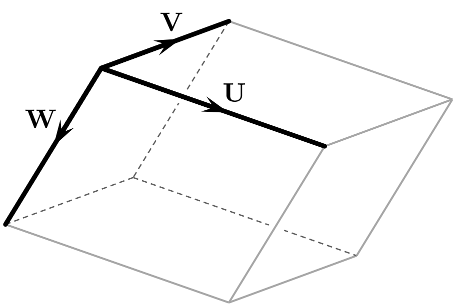
Let $\mathbf{U}$ and $\mathbf{V}$ be two vectors. The cross product $\mathbf{U} \times \mathbf{V}$ can be interpreted as a (signed) measure of the area of the triangle area "enclosed" by $\mathbf{U}$ and $\mathbf{V}$, as shown below



$$\text{Area} = \tfrac{1}{2} \|\mathbf{U} \times \mathbf{V}\| = \tfrac{1}{2} \|\mathbf{U}\| . \|\mathbf{V}\| . \sin\theta$$

1.30 The Scalar Triple Product

Let $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ be three vectors. Then $\mathbf{V} \times \mathbf{W}$ is a vector, so we can form its dot product with $\mathbf{U}$ to get $\mathbf{U} \cdot (\mathbf{V} \times \mathbf{W})$. This is known as the **scalar triple product** of $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$. In fact, $\mathbf{U} \cdot (\mathbf{V} \times \mathbf{W})$ represents the (signed) volume of the parallelipiped with sides $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$. In the picture below $\mathbf{U}$ points (roughly) to the right and $(\mathbf{V} \times \mathbf{W})$ points (roughly) to the left, so their dot product $\mathbf{U} \cdot (\mathbf{V} \times \mathbf{W})$ will be negative.



We can divide the parallelipiped into six identical tetraheda whose edges are described by the vectors $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$. Then, the volume of each of these tetrahedra is given by

$$\text{Volume} = \tfrac{1}{6}\left|\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W})\right| \tag{1.1}$$

The volume of the parallelipiped will be non-zero provided that $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ are not coplanar (which implies, in particular, that they are all different and that none of them is zero). So, $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ form a basis for $\mathcal{V}_3$ if and only if $\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W}) \neq 0$.

1.31 Dual Bases

Let $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ be three vectors that form a basis for $\mathcal{V}_3$. Then, given any vector $\mathbf{R}$, we know that there exist unique numbers λ, μ, ν such that

$$\mathbf{R} = \lambda\mathbf{U} + \mu\mathbf{V} + \nu\mathbf{W}$$

Our goal is to derive explicit formulae for λ, μ, and ν. We know that $\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W}) \neq 0$, so we can define the **dual basis** $\tilde{\mathbf{U}}$, $\tilde{\mathbf{V}}$, $\tilde{\mathbf{W}}$ by the equations

$$\tilde{\mathbf{U}} = \frac{\mathbf{V}\times\mathbf{W}}{\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W})} \quad ; \quad \tilde{\mathbf{V}} = \frac{\mathbf{W}\times\mathbf{U}}{\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W})} \quad ; \quad \tilde{\mathbf{W}} = \frac{\mathbf{U}\times\mathbf{V}}{\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W})}$$

It is easy to verify the following:

- If $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ are orthonormal, then $\mathbf{U} = \tilde{\mathbf{U}}$, $\mathbf{V} = \tilde{\mathbf{V}}$, $\mathbf{W} = \tilde{\mathbf{W}}$.
- $\mathbf{U}\cdot\tilde{\mathbf{U}} = \mathbf{V}\cdot\tilde{\mathbf{V}} = \mathbf{W}\cdot\tilde{\mathbf{W}} = 1$
- $\mathbf{V}\cdot\tilde{\mathbf{U}} = \mathbf{W}\cdot\tilde{\mathbf{U}} = 0 \quad ; \quad \mathbf{W}\cdot\tilde{\mathbf{V}} = \mathbf{U}\cdot\tilde{\mathbf{V}} = 0 \quad ; \quad \mathbf{U}\cdot\tilde{\mathbf{W}} = \mathbf{V}\cdot\tilde{\mathbf{W}} = 0$

So, if $\mathbf{R} = \lambda\mathbf{U} + \mu\mathbf{V} + \nu\mathbf{W}$, then, taking dot products with $\tilde{\mathbf{U}}$, we get

$$\mathbf{R}\cdot\tilde{\mathbf{U}} = \lambda\mathbf{U}\cdot\tilde{\mathbf{U}} + \mu\mathbf{V}\cdot\tilde{\mathbf{U}} + \nu\mathbf{W}\cdot\tilde{\mathbf{U}} = \lambda$$

Taking dot products with $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{W}}$ in a similar way shows that $\mathbf{R}\cdot\tilde{\mathbf{V}} = \mu$ and $\mathbf{R}\cdot\tilde{\mathbf{W}} = \nu$. So, finally, we have

$$\mathbf{R} = (\mathbf{R}\cdot\tilde{\mathbf{U}})\mathbf{U} + (\mathbf{R}\cdot\tilde{\mathbf{V}})\mathbf{V} + (\mathbf{R}\cdot\tilde{\mathbf{W}})\mathbf{W}$$

So, the components of $\mathbf{R}$ with respect to the basis $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ are $\mathbf{R}\cdot\tilde{\mathbf{U}}$, $\mathbf{R}\cdot\tilde{\mathbf{V}}$, $\mathbf{R}\cdot\tilde{\mathbf{W}}$ respectively. We already know that we can obtain components with respect to an orthonormal basis simply by calculating dot products. Here, we have shown that similar calculations work with *any* basis (orthonormal or not), provided we take dot products with the vectors in the dual basis.

1.32 The Vector Triple Product

Let $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$ be three vectors. Then we have

$$\mathbf{U}\times(\mathbf{V}\times\mathbf{W}) = (\mathbf{U}\cdot\mathbf{W})\mathbf{V} - (\mathbf{U}\cdot\mathbf{V})\mathbf{W}$$

$$(\mathbf{U}\times\mathbf{V})\times\mathbf{W} = (\mathbf{U}\cdot\mathbf{W})\mathbf{V} - (\mathbf{V}\cdot\mathbf{W})\mathbf{U}$$

1.33 Lagrange's Identity

Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ be four vectors. Then we have

$$(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{C}\times\mathbf{D}) = (\mathbf{A}\cdot\mathbf{C})(\mathbf{B}\cdot\mathbf{D}) - (\mathbf{A}\cdot\mathbf{D})(\mathbf{B}\cdot\mathbf{C}) = \begin{bmatrix} \mathbf{A}\cdot\mathbf{C} & \mathbf{A}\cdot\mathbf{D} \\ \mathbf{B}\cdot\mathbf{C} & \mathbf{B}\cdot\mathbf{D} \end{bmatrix} \tag{1.2}$$

An important special case is $\mathbf{A}=\mathbf{C}$ and $\mathbf{B}=\mathbf{D}$. Then we get

$$\|\mathbf{A}\times\mathbf{B}\|^2 = \|\mathbf{A}\|^2\cdot\|\mathbf{B}\|^2 - (\mathbf{A}\cdot\mathbf{B})^2$$

Exercises

[1.1] Draw a diagram that "proves" that $\mathbf{U}+(\mathbf{V}+\mathbf{W}) = (\mathbf{U}+\mathbf{V})+\mathbf{W}$.

[1.2] Use vector methods to show that the angle in a semi-circle is a right angle. Is there an analogous result for spheres ?

[1.3] A median of a triangle is a line that joins a vertex to the mid-point of the opposite side. Use vector methods to show that the three medians of any triangle meet at a point.

[1.4] A line segment joining a vertex of a tetrahedron with the centroid of the opposite face is called a median. A line segment joining the midpoints of two opposite edges is called a bimedian of the tetrahedron. Hence there are four medians and three bimedians in a tetrahedron. Show that these seven line segments all meet at a point. Hint: this meeting point is called the centroid of the tetrahedron.

[1.5] Let $ABCD$ be any quadrilateral. Let P, Q, R, S be the mid-points of AB, BC, CD, AD respectively. Show that $PQRS$ is a parallelogram.

[1.6] Show that $(\mathbf{U}+\mathbf{V})\cdot(\mathbf{U}-\mathbf{V}) = \|\mathbf{U}\|^2 - \|\mathbf{V}\|^2$. What does this mean geometrically?

[1.7] Let $\mathbf{Q}$ be a point, let $\mathbf{U}$ be a unit vector, and let $r>0$. Consider the set of points $\mathbf{X}$ with the property that $\|(\mathbf{X}-\mathbf{Q})\times\mathbf{U}\| = r$. What kind of geometric object is this set (is it a line, a circle, a sphere, or what) ? What if $r=0$?

[1.8] Show that

$$(\mathbf{U}\times\mathbf{V})\cdot[(\mathbf{V}\times\mathbf{W})\times(\mathbf{W}\times\mathbf{U})] = [\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W})]^2$$

Hint: We know that for any vector $\mathbf{X}$,

$$\mathbf{X}\times(\mathbf{W}\times\mathbf{U}) = (\mathbf{X}\cdot\mathbf{U})\mathbf{W} - (\mathbf{X}\cdot\mathbf{W})\mathbf{U}$$

Apply this formula with $\mathbf{X} = \mathbf{V}\times\mathbf{W}$.

[1.9] If $\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W}) = k$, show that the dual basis vectors $\tilde{\mathbf{U}}$, $\tilde{\mathbf{V}}$, $\tilde{\mathbf{W}}$ defined in section (1.31) satisfy $\tilde{\mathbf{U}}\cdot(\tilde{\mathbf{V}}\times\tilde{\mathbf{W}}) = 1/k$. Hint: Use the result from the previous problem above.