Basics of 2D and 3D graphics

Linear algebra

Piotr Fulmański

piotr@fulmanski.pl

November 24, 2016

Spis treści

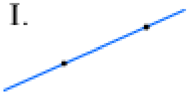
In geometry, topology and related branches of mathematics a (spatial) *point* is a primitive notion upon which other concepts may be defined. In geometry, points are zero-dimensional; i.e., they do not have volume, area, length, or any other higher-dimensional analogue.
Although there are spaces where *point* can be defined. For example, introducing Cartesian coordinates in Euclidean space a point can be defined as an ordered pair, triplet etc. of real numbers.
On the other hand one way to think of the Euclidean plane is as a set of points satisfying certain relationships, expressible in terms of distance and angle. For example, there are three fundamental operations (referred to as motions) on the plane. One is *translation*, which means a shifting of the plane so that every point is shifted in the same direction and by the same distance. The other is *rotation* about a fixed point in the plane, in which every point in the plane turns about that fixed point through the same angle. The last is *reflection*[1].

[1] In http://en.wikipedia.org/wiki/Euclidean_space

Euclidean space and euclidean geometry

Euclidean space – a space described by euclidean geometry.
Euclidean geometry is an axiomatic system, in which all theorems ("true statements") are derived from a small number of axioms. In the Elements Euclid gives five postulates (axioms) for plane geometry, stated in terms of constructions.

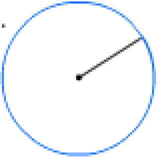
1. Any two points can be connected by a line segment.



2. Any segment can be extended indefinitely (resulting in a straight line).

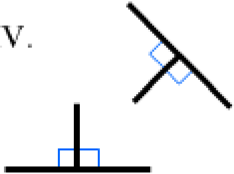
3. For a given line segment, you can describe a circle centered at one of its end points and a radius equal to its length.

III.

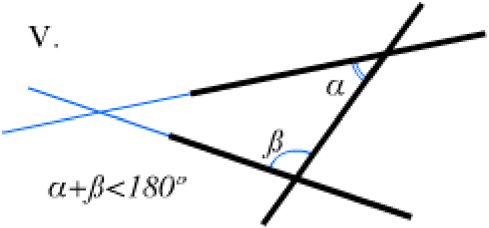


4. All right angles are congruent.

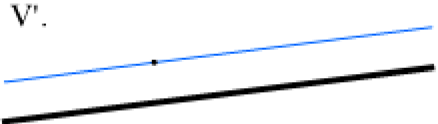
IV.



5. If two lines intersect a third in such a way that the sum of the inner angles on one side is less than two right angles, then the two lines inevitably must intersect each other on that side if extended far enough.



The last axiom can be formlated also as: through a given point which does not belong to a given straight line can be drawn one straight line disjoint with that straight line.

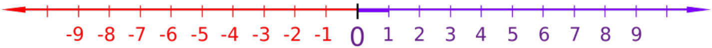


Systems of coordinates

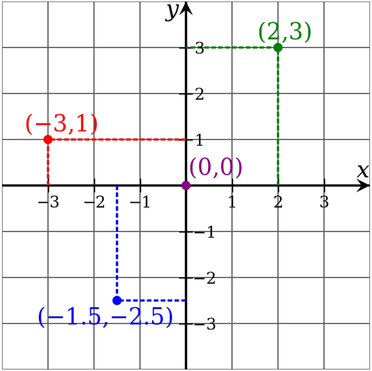
In geometry, a *coordinate system* is a system which uses one or more numbers (known as *coordinates*), to uniquely determine the position of a point or other geometric element in a space such as Euclidean space.
In this tutorial we will use the follwing coordinate systems

- number line,
- Cartesian coordinates,
- polar coordinates,
- homogeneous coordinate system.

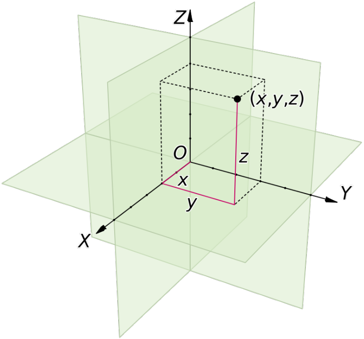
The number line is the simplest example of a coordinate system where points are identified on a line with real numbers. In this system, an arbitrary point O (the origin) is chosen on a given line. The coordinate of a point P is defined as the signed distance from O to P, where the signed distance is the distance taken as positive or negative depending on which side of the line P lies. Each point is given a unique coordinate and each real number is the coordinate of a unique point.



A Cartesian coordinate system is a coordinate system that specifies each point uniquely in a plane by a pair of numerical coordinates, which are the signed distances to the point from two fixed perpendicular directed lines, measured in the same unit of length. Each reference line is called a coordinate axis or just axis of the system, and the point where they meet is its origin, usually at ordered pair (0, 0). The coordinates can also be defined as the positions of the perpendicular projections of the point onto the two axes, expressed as signed distances from the origin.
One can use the same principle to specify the position of any point in three-dimensional space by three Cartesian coordinates, its signed distances to three mutually perpendicular planes (or, equivalently, by its perpendicular projection onto three mutually perpendicular lines).
In general, the position of any point in n-dimensional space by n Cartesian coordinates, is a signed distances to n mutually perpendicular hyperplanes (or, equivalently, it is perpendicular projection onto n mutually perpendicular hyperplanes).

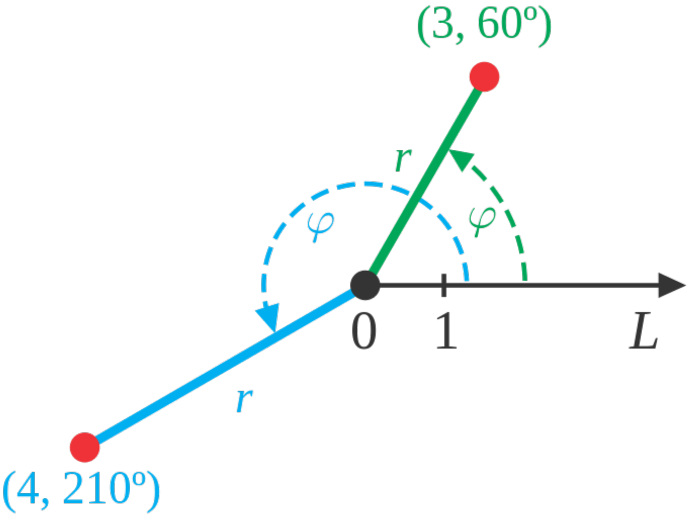


y
x
(2,3)
(−3,1)
(0,0)
(−1.5,−2.5)
−3
−2
−1
1
2
3



Z
Y
X
O
(x,y,z)
x
y
z

The polar coordinate system in two-dimension is a system in which each point on a plane is determined by a distance from a reference point and an angle from a reference direction. A point is chosen as the *pole* and a ray from this point is taken as the *polar axis*. For a given angle φ, there is a single line through the pole whose angle with the polar axis is φ (measured counterclockwise from the axis to the line). Then there is a unique point on this line whose signed distance from the origin is r for given number r. For a given pair of coordinates (r, φ) there is a single point, but any point is represented by infinite number of different polar coordinates. For example, (r, φ), $(r, \varphi + 2\pi)$ and $(-r, \varphi + \pi)$ are all polar coordinates for the same point. The pole is represented by $(0, \varphi)$ for any value of φ.


(3, 60º)
r
φ
φ
0
1
L
r
(4, 210º)

The polar coordinates r and ϕ can be converted to the Cartesian coordinates x and y by using the trigonometric functions

$$x = r\cos\varphi$$
$$y = r\sin\varphi$$

The Cartesian coordinates x and y can be converted to polar coordinates r and φ with $r \geq 0$ and φ in the interval $(-\pi, \pi]$ by

$$r = \sqrt{x^2 + y^2}$$
$$\varphi = \mathrm{atan2}(y, x)$$

To define the trigonometric functions for the angle φ, we need any right triangle that contains the angle φ. The three sides of the triangle are named as follows:

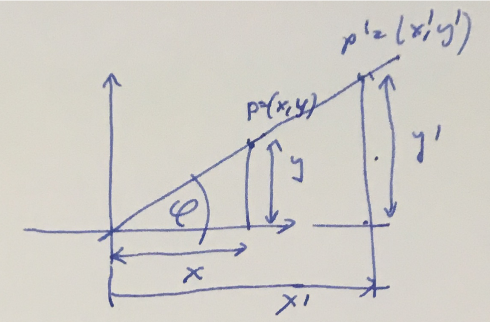
- The hypotenuse is the side opposite the right angle. The hypotenuse is always the longest side of a right-angled triangle.
- The opposite side is the side opposite to the angle we are interested in (angle φ).
- The adjacent side is the side having both the angles of interest (angle φ and right-angle).

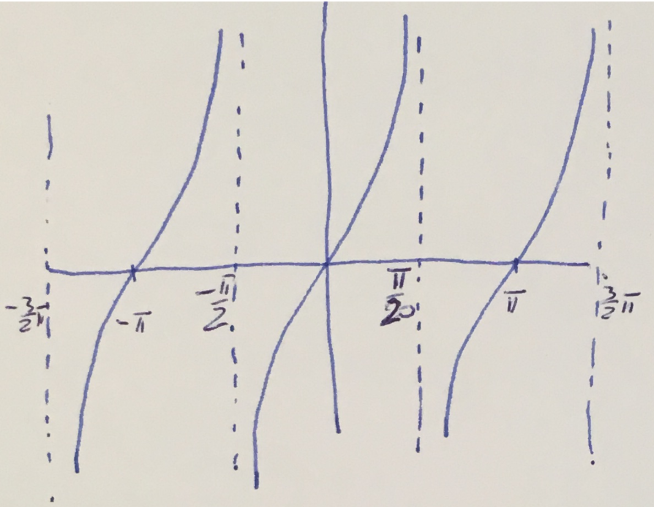
Tangent (tan or tg) is defined as

$$\tan(\varphi) = \frac{y}{x}$$

So having an angle we can find the ratio of the length of the two triangle side: opposite and adjacent to the angle.
This ratio does not depend on the size of the particular right triangle chosen, as long as it contains the angle φ, since all such triangles are similar.

$$\tan(\varphi) = \frac{y}{x} = \frac{y'}{x'}$$

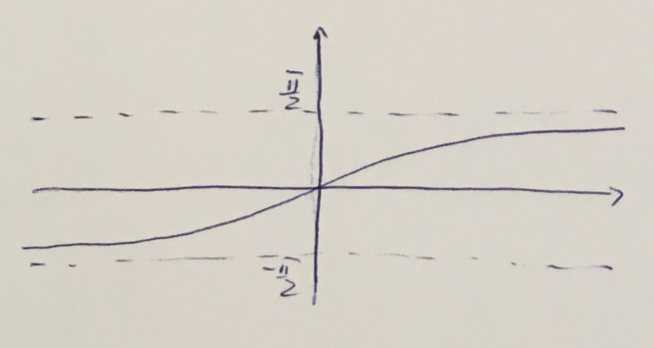




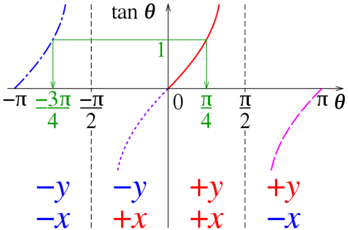
$-\frac{3}{2}\pi$
$-\pi$
$-\frac{\pi}{2}$
$\frac{\pi}{2}$
π
$\frac{3}{2}\pi$

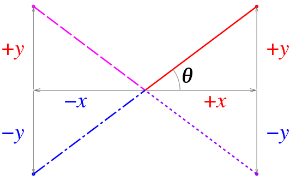
The inverse of the tangent function named arctan or atan can be defined as

$$\arctan\left(\frac{y}{x}\right) = \varphi$$



The function arctan2 is the arctangent function with two arguments. The purpose of using two arguments instead of one is to gather information on the signs of the inputs in order to return the appropriate quadrant of the computed angle, which is not possible for the single-argument arctangent function. For example, both $\frac{y}{x}$ ratio and $\frac{-y}{-x}$ ratio for fixed x and y are equal so arctangent for them returns the same value but both belongs to diferent quadrants. It also avoids the problems of division by zero.





arctan2(y, x) is the angle in radians between the positive x-axis of a plane and the point given by the coordinates (x, y) on it. The angle is positive for counter-clockwise angles (upper half-plane, $y > 0$), and negative for clockwise angles (lower half-plane, $y < 0$).
In terms of the standard arctan function, that is with range of $(-\frac{\pi}{2}, \frac{\pi}{2})$, it can be expressed as follows:

$$\text{arctan2}(y, x) = \begin{cases} \arctan(\frac{y}{x}) & x > 0 \\ \arctan(\frac{y}{x}) + \pi & y \geq 0\ ,\ x < 0 \\ \arctan(\frac{y}{x}) - \pi & y < 0\ ,\ x < 0 \\ \frac{\pi}{2} & y > 0\ ,\ x = 0 \\ -\frac{\pi}{2} & y < 0\ ,\ x = 0 \\ \text{undefined} & y = 0\ ,\ x = 0 \end{cases}$$

In mathematics, homogeneous coordinates or projective coordinates, are a system of coordinates used in projective geometry, as Cartesian coordinates are used in Euclidean geometry. With this type of coordinates we can easyly represent also points at infinity, using finite coordinates.

Definition

Given a point $p = (x, y)$ on the Euclidean plane, for any non-zero real number w, the triple (xw, yw, w) is called a set of homogeneous coordinates for the point p. By this definition, **multiplying the three homogeneous coordinates by a common, non-zero factor gives a new set of homogeneous coordinates for the same point**. In particular, $(x, y, 1)$ is such a system of homogeneous coordinates for the point (x, y).

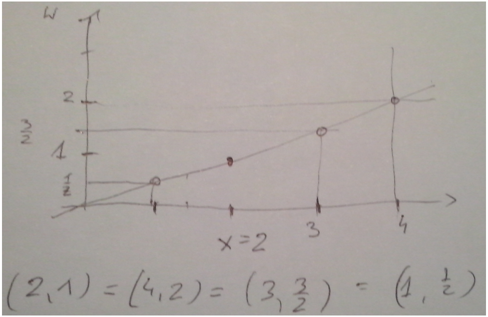
For example, the Cartesian point $(1, 2)$ can be represented in homogeneous coordinates as $(1, 2, 1)$ or $(2, 4, 2)$. The original Cartesian coordinates are recovered by dividing the first two positions by the third. Thus unlike Cartesian coordinates, **a single point can be represented by infinitely many homogeneous coordinates**.

Another definition of the real projective plane can be given in terms of equivalence classes.

Definition

For non-zero element of R^3, define $(x_1, y_1, z_1) \sim (x_2, y_2, z_2)$ to mean there is a non-zero λ so that $(x_1, y_1, z_1) = (\lambda x_2, \lambda y_2, \lambda z_2)$. Then $\sim$ is an equivalence relation and the projective plane can be defined as the equivalence classes of $R^3 \setminus \{0\}$. If (x, y, z) is one of the elements of the equivalence class p then these are taken to be homogeneous coordinates of p.

Of course both previous definition are equivalent, but focus on different aspects of homogeneous coordinates. Let's consider a following example.



The point $p_{Cart} = (x) = (2)$ is represented by any point $p_{Hom} = (x, w)$ on the line $w = \frac{1}{2}x$. However, before we interpret p_{hom} as a conventional Cartesian coordinate, we have to ,,divide" it first (divide all its coordinates) by w to get the form when last coordinate is equal to 1

$$p_{hom} = (x, w) = \left(\frac{x}{w}, \frac{w}{w}\right) = \left(\frac{x}{w}, 1\right)$$

Because $w = \frac{1}{2}x$ then $x = 2w$ and finaly
$p_{hom} = \left(\frac{x}{w}, 1\right) = \left(\frac{2w}{w}, 1\right) = (2, 1)$

1. Any point in the (2D) projective plane is represented by a triple (x, y, z), called the homogeneous coordinates or projective coordinates of the point, where x, y and z are not all 0.
2. The point represented by a given set of homogeneous coordinates is unchanged if the coordinates are multiplied by a common factor.
3. Two sets of homogeneous coordinates represent the same point if and only if one is obtained from the other by multiplying all the coordinates by the same non-zero constant.
4. When z is not 0 the point represented is the point $(x/z, y/z)$ in the Euclidean plane.
5. When z is 0 the point represented is a point at infinity.

Note that

1. For a given point (x, y) on the Euclidean plane, and for any non-zero real number z, the triple (xz, yz, z) is called a set of homogeneous coordinates for the point.
2. Generaly speaking, any point with n coordinates in Cartesian coordinate system, has $n + 1$ coordinates in homogeneous coordinate system.
3. The triple $(0, 0, 0)$ is omitted and does not represent any point. The origin is represented by $(0, 0, 1)$.

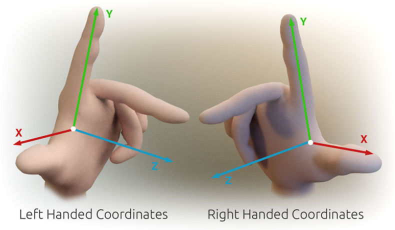
We will be back to this concept when *viewing frustum* will be discussed.

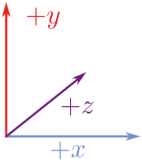
Left- and right-handed Cartesian coordinate systems

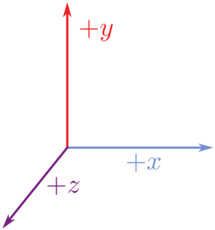
In three dimensions the three axes of 3-dimensional space have two possible orientations. Once the x- and y-axes are specified, they determine a plane and the perpendicular line to it along which the z-axis should lie. The only problem is that there are two possible directions on this line. The two possible coordinate systems which result are called *left-handed* and *right-handed*. The name derives from the left- or right-hand rule we use to distinguish them.

A common mnemonic for understanding orientation conventions for vectors in three dimensions are the *left-* and *right-hand* rule. If the thumb of the left (right) hand is pointed along positive x axis, the index finger at a right angle to it is pointed along positive y axis, and the middle finger bent inward at a right angle to both, the three fingers indicate the relative directions of the x-, y-, and z-axes in a left (right)-handed system. The thumb indicates the positive x-axis, the index finger the positive y-axis and the middle finger the positive z-axis.

Left- and right-handed Cartesian coordinate systems

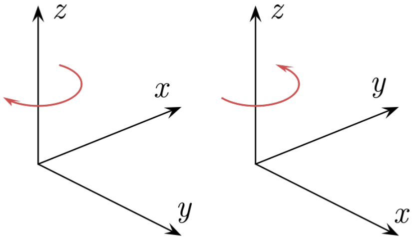






Left- and right-handed Cartesian coordinate systems

For left (right) handed coordinates when the left (right) thumb points along the z axis in a positive z-direction, the curled fingers of the left (right) hand represent a motion from x axis to y axis. When viewed from z axis the system is clockwise (counter-clockwise).



Because, as we can see, there is no one way to express a 3D system, whenever we are working with a new framework or a set of frameworks it is important to check which coordinate system is used by which framework. It ultimately doesn't matter which coordinate system our game uses, as long as it's consistent across the code and frameworks.

Vector is an object that has *magnitude* (or *length*) and *direction*. A vector is what is needed to "carry" (move) the point A to the point B as the Latin word *vector* means *carrier*.
In n-dimensional Euclidean space, vectors are identified with n-tuple of scalar (number) components:

$$\mathbf{a} = [a_1, a_2, \ldots, a_n].$$

Vector

A 3D vector can be represented by a triple of scalars (x, y, z), just as a point can be. The distinction between points and vectors is actually quite subtle. Technically, a vector is just an offset relative to some known point. A vector can be moved anywhere in 3D space – as long as its magnitude and direction don't change, it is the same vector.
As a cosequence we can say that a vector has no concept of position. This means that two vectors are identical as long as they have the same magnitude (or length) and point in the same direction.
A vector can be used to represent a point, provided that we fix the tail of the vector to the origin of the chosen coordinate system. In such conditions, its head pokrywa sie with this point. Such a vector is sometimes called a *position vector* or *radius vector*. We can interpret any triple of scalars as either a point or a vector. One might say that points are **absolute**, while vectors are **relative**.

Let $\mathbf{a} = [a_1, a_2, \ldots, a_n]$ be a vector in n-dimensional Euclidean space. Then

- The *length* (or *magnitude* or *norm*) of the vector a is denoted by $||\mathbf{a}||$ and can be computed with the Euclidean norm

$$||\mathbf{a}|| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

  We can think about this formula as a distance from the origin (tail) to the position at which the vector is pointing (head).

Let $\mathbf{a} = [a_1, a_2, \ldots, a_n]$ be a vector in n-dimensional Euclidean space. Then

- A *unit vector* is any vector with a length of one. A unit vector is often indicated with a "hat" above the vector's symbol, as in $\hat{\mathbf{a}}$. Any vector of arbitrary length greater than 0 can be divided by its length to create a unit vector. This is known as *normalization*.

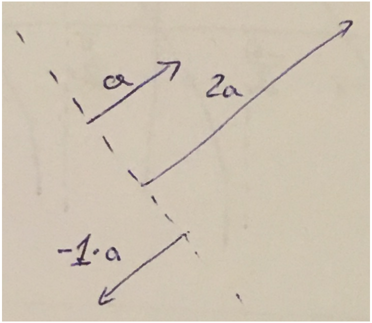
$$\hat{\mathbf{a}} = \left[\frac{a_1}{||\mathbf{a}||}, \frac{a_2}{||\mathbf{a}||}, \ldots, \frac{a_n}{||\mathbf{a}||} \right].$$

  Because, when a vector is normalized, it will lose any magnitude information, we have to take care if we can do this or not, not to normalize the wrong vectors. A good rule of thumb to follow is if we only care about the direction of the vector, we should normalize it.

This is the simples vector opration. To multiply a vector, or re-scaled it, by a real number s (often called *scalar* from scale), we simply multiply each component by this number

$$s \cdot \mathbf{a} = [s \cdot a_1, s \cdot a_2, \ldots, s \cdot a_n]$$

Intuitively, multiplying by a scalar s stretches a vector out by a factor of s. If s is negative, then the vector changes direction: it flips around by an angle of 180 degree.



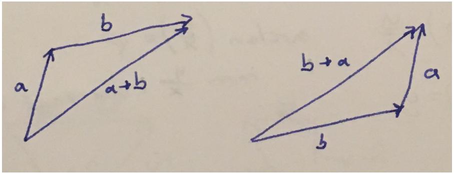
Let $\mathbf{a} = [a_1, a_2, \ldots, a_n]$ and $\mathbf{b} = [b_1, b_2, \ldots, b_n]$ are any vectors, maybe with different magnitudes and directions.
The sum of **a** and **b** is

$$\mathbf{a} + \mathbf{b} = [a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n].$$

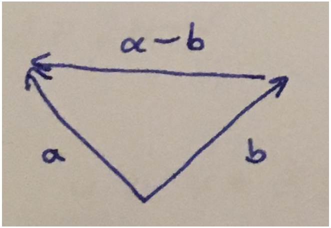
The difference of **a** and **b** is

$$\mathbf{a} - \mathbf{b} = [a_1 - b_1, a_2 - b_2, \ldots, a_n - b_n].$$

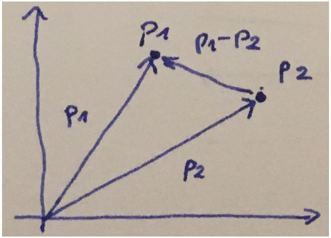
More than definition we should take care about graphical interpretation of addition and subtraction.



Notice that with addition order of vectors doesn't matter.



Subtraction is important, because it enables us to construct a vector between two points.

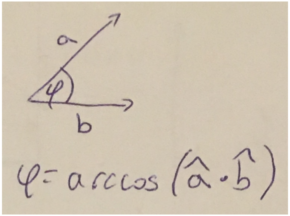


The *scalar product* – since its result is a scalar (number) – (or *dot product* or *inner product* or sometimes *projection product* for emphasizing the geometric significance), is an algebraic operation that takes two equal-length sequences of numbers (usually coordinate vectors) and returns a single number.

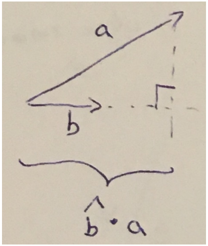
$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \cos\varphi$$

where φ is the measure of the angle between $\mathbf{a}$ and $\mathbf{b}$.
We can use dot product to find the angle betwen two unit vectors

$$\varphi = \arccos\left(\hat{\mathbf{a}} \cdot \hat{\mathbf{b}}\right).$$



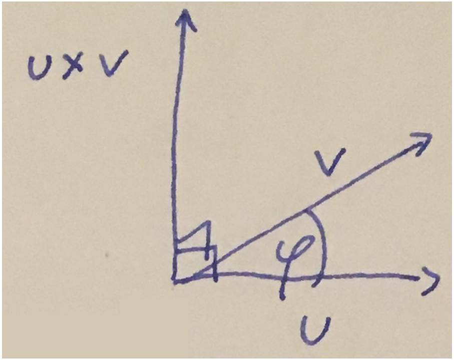
- If the dot product between two vectors results in 0, it means they are perpendicular to each other (because $\cos(90) = 0$).
- If the dot product of two unit wectors results in 1, it means the vectors are parallel and facing in the same direction.
- If the dot product of two unit wectors results in -1 means they are parallel and face in the opposite direction.
- If $\hat{\mathbf{u}}$ is a unit vector, then the dot product $\mathbf{v} \cdot \hat{\mathbf{u}}$ represents the length of the **projection** of a vector $\mathbf{v}$ onto the infinie line defined by the direction of $\mathbf{u}$.



See `http://www.falstad.com/dotproduct/`

The cross product between two 3D vectors **u** and **v** results in a third vector. Given two vectors, there is only a single plane that contains both vectors. The cross product finds a vector that is perpendicular to this plane, which is known as a normal to the plane.

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = [(u_y v_z - u_z v_y), (u_z v_x - u_x v_z), (u_x v_y - u_y v_x)]$$



$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = [(u_y v_z - u_z v_y), (u_z v_x - u_x v_z), (u_x v_y - u_y v_x)]$$

To help memmorize this formula it's good to notice that

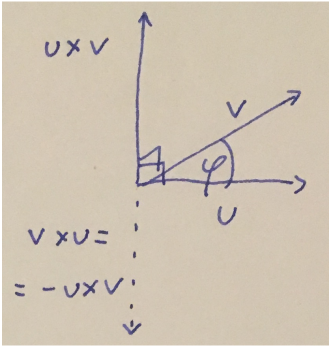
- Each component of $\mathbf{w}$ is of the form $u_{s_1} v_{s_2} - u_{s_2} v_{s_1}$, with subscripts s_1, s_2.
- Subscripts s_1, s_2 for the first component takes values from string $\mathtt{yz}$.
- Subscripts for the next component takes values with letter substituded in the following manner

$$yz \rightarrow zx \rightarrow xy$$

The magnitude of the cross product $\mathbf{u} \times \mathbf{v}$ is equal to the area of the parallelogram whose sides are $\mathbf{u}$ and $\mathbf{v}$ and is equal to

$$|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}||\mathbf{v}| \sin(\varphi).$$

An important thing to note is that there is a second vector that is perpendicular to the plane: the vector that points in the **opposite** direction of $\mathbf{w}$.



To find which one is correct we have to use the aftermentioned concept of handedness of the coordinate system. If we have a left-handed (right-handed) coordinate system then take left (right) hand and

1. Line up a thumb with the first vector ($\mathbf{u}$).
2. Line up an index finger with a second vector ($\mathbf{v}$).
3. The direction a middle finger points in is the direction the cross product will face.

Note that to get a second vector it's enough to flip fingers so thumb lines up with $\mathbf{v}$ and index finger with $\mathbf{u}$.

- $\mathbf{v} \times \mathbf{w} \neq \mathbf{w} \times \mathbf{v}$
- $\mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v}$
- $\mathbf{v} \times (\mathbf{w} + \mathbf{y}) = (\mathbf{v} \times \mathbf{w}) + (\mathbf{v} \times \mathbf{y})$
- the Cartesian basis vectors $\mathbf{e_1} = [1, 0, 0]$, $\mathbf{e_2} = [0, 1, 0]$, $\mathbf{e_3} = [0, 0, 1]$ are related by cross products as follows

$$\begin{aligned} \mathbf{e_1} \times \mathbf{e_2} &= -(\mathbf{e_2} \times \mathbf{e_1}) &= \mathbf{e_3} \\ \mathbf{e_2} \times \mathbf{e_3} &= -(\mathbf{e_3} \times \mathbf{e_2}) &= \mathbf{e_1} \\ \mathbf{e_3} \times \mathbf{e_1} &= -(\mathbf{e_1} \times \mathbf{e_3}) &= \mathbf{e_2} \end{aligned}$$

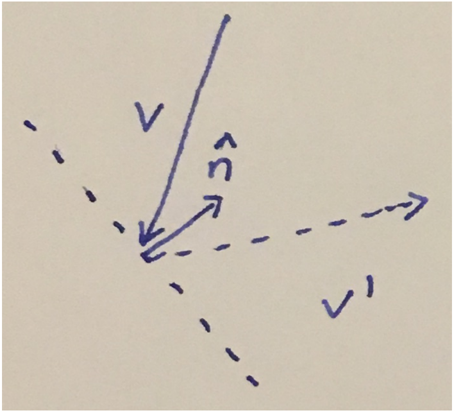
Remarks

- If the cross product returns a vector where all three components are 0, this means that the two input vectors are collinear (they lie on the same line).

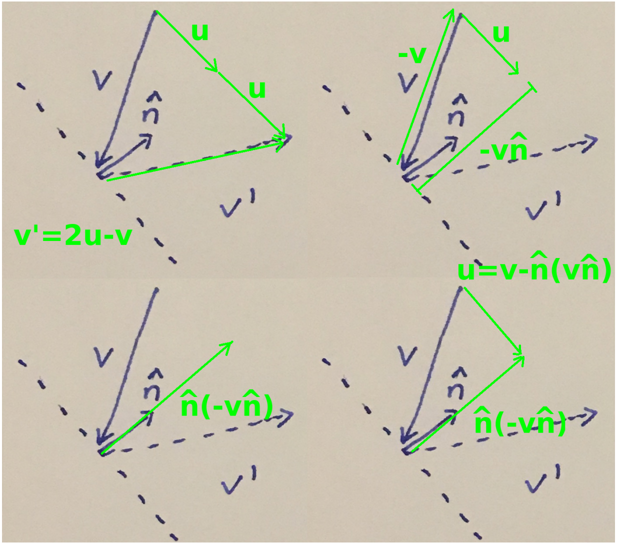
Very often 3D games use 4D vectors. As we know from the previous part, when 4D coordinates are used for a 3D space, we have homogenous coordinates. Let's denote the fourth component as the w-component. In most instances, the w-component will be either 0 or 1.

- If $w = 0$, this means that the homogenous coordinate represents a 3D vector.
- If $w = 1$, this means that the homogenous coordinate represents a 3D point.

There are two known values: the vector $\mathbf{v}$, which is the velocity of an object (for example a ball) prior to reflection, and the normal $\hat{\mathbf{n}}$, which is a unit vector that is perpendicular to the surface of reflection. We need to solve for the vector $\mathbf{v}'$, which represents the velocity after reflection.



v
n̂
v'


u
u
v
n̂
v'
v'=2u-v
-v
u
v
n̂
-vn̂
v'
u=v-n̂(vn̂)
v
n̂
n̂(-vn̂)
v'
v
n̂
n̂(-vn̂)
v'

Matrix

In mathematics, a matrix (plural matrices) is a rectangular array of objects (for example numbers, symbols, or expressions), arranged in rows and columns.

There are a number of basic operations that can be applied to modify matrices, but we will focus on

- matrix addition,
- scalar multiplication,
- transposition,
- matrix multiplication.

The sum **C** of two m-by-n matrices **A** and **B** is calculated entrywise

$$(c)_{i,j} = a_{i,j} + b_{i,j},$$

where $1 \leq i \leq m$ and $1 \leq j \leq n$ and $c_{i,j}$ is an element from matrix **C** at position (i, j); similarly for matrix **A** and **B**.
Example

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 7 & 8 & 9 \\ 0 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 1+7 & 2+8 & 3+9 \\ 4+0 & 5+(-1) & 6+(-2) \end{bmatrix} = \begin{bmatrix} 8 & 10 & 12 \\ 4 & 4 & 4 \end{bmatrix}$$

The product c of a scalar t (number) and a m-by-n matrix $\mathbf{A}$ is computed by multiplying every entry of $\mathbf{A}$ by c:

$$(c)_{i,j} = c \cdot a_{i,j},$$

where $1 \leq i \leq m$ and $1 \leq j \leq n$.
Example

$$3 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot 3 \\ 3 \cdot 4 & 3 \cdot 5 & 3 \cdot 6 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{bmatrix}$$

Please note that although this operation is called *scalar multiplication*, its result is not named *scalar product* to avoid confusion, since *scalar product* is sometimes used as a synonym for *inner product* what we have mentioned earlier.

The transpose C of an m-by-n matrix $\mathbf{A}$ (denoted as $\mathbf{A}^T$) is the n-by-m matrix $\mathbf{B}$ formed by turning rows into columns and vice versa

$$c_{i,j} = a_{j,i},$$

where $1 \leq i \leq m$ and $1 \leq j \leq n$.
Example

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Multiplication of two matrices is defined if and only if the number of columns of the left matrix is the same as the number of rows of the right matrix.
If **A** is an m-by-n matrix and **B** is an n-by-p matrix, then their matrix product **C** is the m-by-p matrix whose entries are given by dot product of the corresponding row of **A** and the corresponding column of **B**

$$c_{i,j} = \sum_{r=1}^{n} a_{i,r} b_{r,j} = a_{i,1} b_{1,j} + a_{i,2} b_{2,j} + \cdots + a_{i,m} b_{m,j},$$

where $1 \leq i \leq m$ and $1 \leq j \leq p$.
Example

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 7 & 8 \\ 9 & 0 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} (1 \cdot 7 + 2 \cdot 9 + 3 \cdot -1) & (1 \cdot 8 + 2 \cdot 0 + 3 \cdot -2) \\ (4 \cdot 7 + 5 \cdot 9 + 6 \cdot -1) & (4 \cdot 8 + 5 \cdot 0 + 6 \cdot -2) \end{bmatrix}$$

$$= \begin{bmatrix} (7 + 18 + -3) & (8 + 0 + -6) \\ (28 + 45 + -6) & (32 + 0 + -12) \end{bmatrix} = \begin{bmatrix} 22 & 2 \\ 67 & 20 \end{bmatrix}$$

There are two methods to represent a vector as a matrix: it could be

- a matrix with a single row (so called *row-major*, *row vector* or *row matrix*) or
- a matrix with a single column (so called *column-major*, *column vector* or *column matrix*).

If **A** is a row vector and **B** is a column vector, then their matrix products **C** is either

$$\mathbf{AB} = \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = ax + by + cz,$$

or

$$\mathbf{BA} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \begin{bmatrix} a & b & c \end{bmatrix} = \begin{bmatrix} xa & xb & xc \\ ya & yb & yc \\ za & zb & zc \end{bmatrix}.$$

Note **AB** and **BA** are two different matrices. The first is a 1×1 matrix (number) while the second is a 3×3 matrix.
The choice detemines the form of the matrix we will use in the future and the order of vector and matrix during multiplication.

$$
\begin{bmatrix} x & y & z \end{bmatrix}_{m,n=1,3} \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}_{n,p=3,3}
$$

$$
= \begin{bmatrix} xa + yb + zc & xd + ye + zf & xg + yh + zi \end{bmatrix}_{m,p=1,3}
$$

Matrix – multiplicaiton for column-major vector

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}_{m,n=3,3} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{n,p=3,1} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{bmatrix}_{m,p=3,1}$$

Matrix – multiplicaiton for row- or column-major vector

In some sense, both results are the same: we have a vector. The problem is that matrix for the first case is different from the second. In consequence, when we have diferent sources of information we have to check what vectors are used: row or column. Be aware of mixing matrices for row-major with matrices for column-major vectors if you don't want to get strange results.

- If matrices $\mathbf{A}$ and $\mathbf{B}$ are transformation matrices, then the product $\mathbf{P} = \mathbf{AB}$ is another transformation matrix that perform *both* of the original transformations. For example, if $\mathbf{A}$ is a scale matrix and $\mathbf{B}$ is a rotation, the matrix $\mathbf{P}$ would both scale and rotate the points or vectors to which it is applied.
- Matrix multiplication is often called *concatenation*.
- $\mathbf{AB} \neq \mathbf{BA}$
- $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
- $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$