

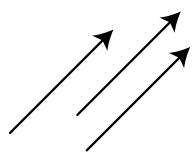
Vectors

CONTENTS

A.1	Scaling a vector	321
A.2	Unit or Direction vectors	321
A.3	Vector addition	322
A.4	Vector subtraction	322
A.5	Points and vectors	322
A.6	Parametric definition of lines and rays	323
A.7	Dot or inner product	323
A.7.1	Trigonometric interpretation of dot product	324
A.7.2	Geometric interpretation of dot product	324
A.7.3	Dot product example: The distance from a point to a line	325
A.7.4	Dot product example: Mirror reflection	325
A.8	Cross Product	326
A.8.1	Trigonometric interpretation of cross product	326
A.8.2	Cross product example: Finding surface normals	327
A.8.3	Cross product example: Computing the area of a triangle	327

320 ■ Foundations of Physically Based Modeling and Animation

To a mathematician, a vector is the fundamental element of what is known as a vector space, supporting the operations of scaling, by elements known as scalars, and also supporting addition between vectors. When using vectors to describe physical quantities, like velocity, acceleration, and force, we can move away from this abstract definition, and stick with a more concrete notion. We can view them as arrows in space, of a particular length and denoting a particular direction, and we can think of the corresponding scalars as simply the real numbers. Practically speaking, a vector is simply a way of simultaneously storing and handling two pieces of information: a direction in space, and a magnitude or length.



An arrow is a convenient way to draw a vector; since both length and direction are clearly indicated. A real number is a convenient way to represent a scalar, which when multiplied by a vector changes its length. To the left are three visual representations of identical vectors. They are identical, since they are all of the same length and the same direction, i.e. they are parallel to each other. Their location within the space is irrelevant.

In the study of physically based animation, we will initially be interested in vectors in two-dimensional (2D) and in three-dimensional (3D) space, whose elements are real numbers. But, we will see later that vectors can be defined in a space of any number of dimensions, with elements that may themselves be multidimensional.

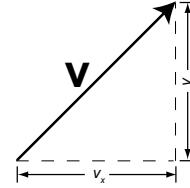
Notationally, a vector is usually denoted by a lower-case letter, which has a line over it, like \bar{v} , or is printed in bold type, like \mathbf{v} . For hand written notes, the line is most convenient, but in printed form the bold form is more usual. Throughout these notes the form \mathbf{v} is used.

A vector in 2D Euclidean space is defined by a pair of scalars arranged in a column, like

$$\mathbf{v} = \begin{bmatrix} v_x \\ v_y \end{bmatrix}.$$

Examining the diagram to the right, we see that v_x denotes the horizontal extent or *component* of the vector, and v_y its vertical component. Note, that in a computer program this structure can be easily represented as a two-element array of floating point numbers, or a struct containing two floats. When working in 2D, the direction of the vector can be given by the slope $m = v_y/v_x$. Its magnitude, also called its *norm*, is written $\|\mathbf{v}\|$. By the Pythagorean Theorem,

$$\|\mathbf{v}\| = \sqrt{v_x^2 + v_y^2}.$$



A vector in 3D space is defined by three scalars arranged in a column,

$$\mathbf{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix},$$

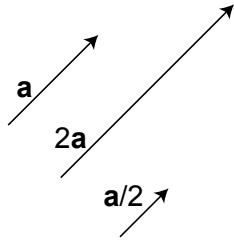
where v_x is the horizontal component, v_y the vertical component, and v_z the depth

component. The norm of a 3D vector \mathbf{v} is

$$\|\mathbf{v}\| = \sqrt{v_x^2 + v_y^2 + v_z^2}.$$

In 3D there is no simple equivalent to the slope. The direction of a 3D vector is often given in terms of its azimuth and elevation. But, for our purposes it will be best understood by its corresponding unit vector, which we will describe after first defining some key algebraic vector operations.

A.1 SCALING A VECTOR



Multiplication of a vector by a real number scalar leaves the vector's direction unchanged, but multiplies its magnitude by the scalar. Algebraically, we multiply each term of the vector by the scalar. For example

$$2\mathbf{a} = 2 \begin{bmatrix} a_x \\ a_y \end{bmatrix} = \begin{bmatrix} 2a_x \\ 2a_y \end{bmatrix}.$$

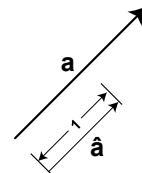
Division by a scalar is the same as multiplication by the reciprocal of the scalar:

$$\mathbf{a}/2 = \begin{bmatrix} a_x/2 \\ a_y/2 \end{bmatrix}.$$

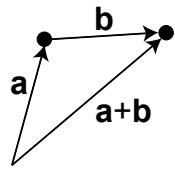
A.2 UNIT OR DIRECTION VECTORS

The direction of a vector is most easily described by a *unit vector*, also called a *direction vector*. A unit vector, for a particular vector, is parallel to that vector but of unit length. Therefore, it retains the direction, but not the norm of the parent vector. Throughout these notes the notation $\hat{\mathbf{v}}$ will be used to indicate a unit vector in the direction of parent vector \mathbf{v} . For example, the unit or direction vector corresponding with the 2D vector \mathbf{a} would be

$$\hat{\mathbf{a}} = \begin{bmatrix} a_x/\|\mathbf{a}\| \\ a_y/\|\mathbf{a}\| \end{bmatrix} = \begin{bmatrix} \hat{a}_x \\ \hat{a}_y \end{bmatrix}.$$



A.3 VECTOR ADDITION

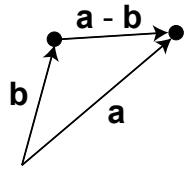


Addition of vectors can be expressed by a diagram. Placing the vectors end to end, the vector from the start of the first vector to the end of the second vector is the sum of the vectors. One way to think of this is that we start at the beginning of the first vector, travel along that vector to its end, and then travel from the start of the second vector to its end. An arrow constructed between the starting and ending points defines a new vector, which is the sum of the original vectors. Algebraically, this is equivalent to adding corresponding terms of the two vectors:

$$\mathbf{a} + \mathbf{b} = \begin{bmatrix} a_x \\ a_y \end{bmatrix} + \begin{bmatrix} b_x \\ b_y \end{bmatrix} = \begin{bmatrix} a_x + b_x \\ a_y + b_y \end{bmatrix}.$$

We can think of this as again making a trip from the start of the first vector to the end of the second vector, but this time traveling first horizontally the distance $a_x + b_x$ and then vertically the distance $a_y + b_y$.

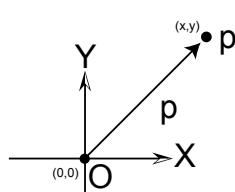
A.4 VECTOR SUBTRACTION



Subtraction of vectors can be shown in diagram form by placing the starting points of the two vectors together, and then constructing an arrow from the head of the second vector in the subtraction to the head of the first vector. Algebraically, we subtract corresponding terms:

$$\mathbf{a} - \mathbf{b} = \begin{bmatrix} a_x \\ a_y \end{bmatrix} - \begin{bmatrix} b_x \\ b_y \end{bmatrix} = \begin{bmatrix} a_x - b_x \\ a_y - b_y \end{bmatrix}.$$

A.5 POINTS AND VECTORS



This leads us to the idea that points and vectors can be interchanged — almost. While vectors can exist anywhere in space, a point is always defined relative to the origin, O . Thus, we can say that a point, $p = (x, y)$, is defined by the origin, $O = (0, 0)$ and a vector, $\mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix}$, i.e.

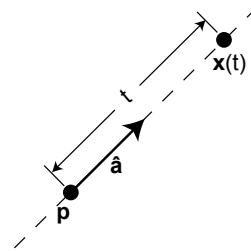
$$p = O + \mathbf{p}.$$

Because the origin is assumed to be the point $(0, 0)$, points and vectors can be represented the same way, e.g. the point $(2, 3)$ can be represented as the vector $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$. This interchangeability can be very convenient in many cases, but can also lead

to confusion. It is a good idea to make sure that when storing data, you clearly indicate which values are points, and which are vectors. As will be seen below, the homogeneous coordinates used to define transformations can help with this.

Equivalent to the above, we can write, $\mathbf{p} = p - O$, i.e. a vector defines the measure from the origin to a particular point in space. More generally, a vector can always be defined by the difference between any two points, p and q . The vector $\mathbf{v} = p - q$ represents the direction and distance from point q to point p . Conversely, the point q and the vector \mathbf{v} define the point, $p = q + \mathbf{v}$, which is translated from q by the components of \mathbf{v} .

A.6 PARAMETRIC DEFINITION OF LINES AND RAYS



This leads us to a compact definition of a line in space, written in terms of a unit vector and a point. Let \mathbf{p} be a known point (expressed in vector form) on the line being defined, and let $\hat{\mathbf{a}}$ be a unit vector whose direction is parallel to the desired line. Then, the locus of points on the line is the set of all points \mathbf{x} , satisfying

$$\mathbf{x}(t) = \mathbf{p} + t\hat{\mathbf{a}}.$$

The variable t is a real number, and is known as the line parameter. It measures the distance from the point \mathbf{p} to the point $\mathbf{x}(t)$. If t is positive, the point \mathbf{x} lies in the direction of the unit vector from point \mathbf{p} , and if t is negative, the point lies in the direction opposite to the unit vector.

The definition of a ray is identical to the definition of a line, except that the parameter t of a ray is limited to the positive real numbers. Thus, a ray can be interpreted as starting from the point \mathbf{p} , and traveling in the direction of $\hat{\mathbf{a}}$ a distance corresponding to t , as t goes from 0 to increasingly large positive values. On a ray, the point \mathbf{p} is called the ray origin, $\hat{\mathbf{a}}$ the ray direction, and t the distance along the ray.

A.7 DOT OR INNER PRODUCT

Vector-vector multiplication is not as easily defined as addition, subtraction and scalar multiplication. There are actually several vector products that can be defined. First, we will look at the *dot product* of two vectors, which is often called their *inner product*.

Defined algebraically, the dot product of two vectors is given by

$$\mathbf{a} \cdot \mathbf{b} = \begin{bmatrix} a_x \\ a_y \end{bmatrix} \cdot \begin{bmatrix} b_x \\ b_y \end{bmatrix} = a_x b_x + a_y b_y.$$

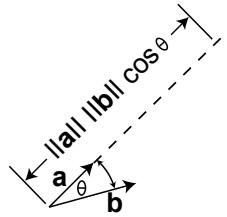
We multiply corresponding terms and add the result. The result is not a vector, but is

in fact a scalar. This turns out to have many ramifications. The dot product is a *mighty* operation and has many uses in graphics!

A.7.1 Trigonometric interpretation of dot product

The dot product can be written in trigonometric form as

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



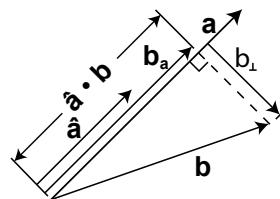
where θ is the smallest angle between the two vectors. Note, that this definition of θ applies in both 2D and 3D. Two nonparallel vectors always define a plane, and the angle θ is the angle between the vectors measured in that plane. Note that if both \mathbf{a} and \mathbf{b} are unit vectors, then $\|\mathbf{a}\| \|\mathbf{b}\| = 1$, and $\mathbf{a} \cdot \mathbf{b} = \cos \theta$. So, in general if you want to find the cosine of the angle between two vectors \mathbf{a} and \mathbf{b} , first compute the unit vectors $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ in the directions of \mathbf{a} and \mathbf{b} then

$$\cos \theta = \hat{\mathbf{a}} \cdot \hat{\mathbf{b}}.$$

Other things to note about the trigonometric representation of dot product that follow directly from the cosine relationship are that

1. the dot product of *orthogonal* (perpendicular) vectors is zero, so if $\mathbf{a} \cdot \mathbf{b} = 0$, for vectors \mathbf{a} and \mathbf{b} with non-zero norms, we know that the vectors must be orthogonal,
2. the dot product of two vectors is positive if the magnitude of the smallest angle between the vectors is less than 90° , and negative if the magnitude of this angle exceeds 90° .

A.7.2 Geometric interpretation of dot product



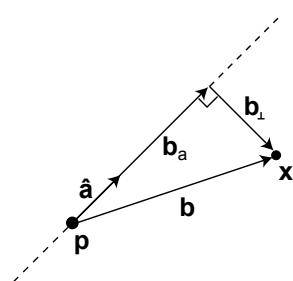
Another very useful interpretation of the dot product is that it can be used to compute the component of one vector in the direction parallel to another vector. For example, let $\hat{\mathbf{a}}$ be a unit vector in the direction of vector \mathbf{a} . Then the length of the projection of another vector \mathbf{b} in the direction of vector \mathbf{a} is $\hat{\mathbf{a}} \cdot \mathbf{b}$. You can think of this as the length of the shadow of vector \mathbf{b} on vector \mathbf{a} . Therefore, the vector component of \mathbf{b} in the direction of \mathbf{a} is

$$\mathbf{b}_a = (\hat{\mathbf{a}} \cdot \mathbf{b})\hat{\mathbf{a}}.$$

So, \mathbf{b}_a is parallel to \mathbf{a} and has length equal to the projection of \mathbf{b} onto \mathbf{a} . Note also that $\mathbf{b}_{\perp} = \mathbf{b} - \mathbf{b}_a$ will be the component of \mathbf{b} perpendicular to vector \mathbf{a} .

The dot product has many uses in graphics that the following two examples will serve to illustrate.

A.7.3 Dot product example: The distance from a point to a line



Let us look at how dot product can be used to compute an important geometric quantity: the distance from a point to a line. We will use the parametric definition of a line, described above, specified by point p and a direction vector \hat{a} . To compute the distance of an arbitrary point x from this line, first compute the vector $b = x - p$, from the point p on the line to the point x . The component of b in the direction of vector \hat{a} is

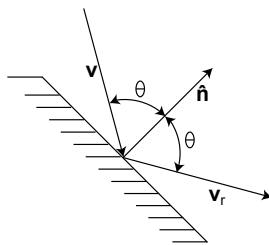
$$b_a = (\hat{a} \cdot b) \hat{a}.$$

The component of b perpendicular to \hat{a} is

$$b_{\perp} = b - b_a,$$

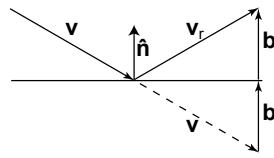
and the distance of point x from the line is simply $\|b_{\perp}\|$.

A.7.4 Dot product example: Mirror reflection



Another very useful example of the use of dot product in geometric calculations is the computation of the mirror reflection from a surface. Assume that we have a flat mirror surface, whose *surface normal* is the unit vector \hat{n} . The surface normal is defined to be a direction vector perpendicular to the surface. Since there are two such vectors at any point on a surface, the convention is to take the direction of the surface normal to be pointing in the “up” direction of the surface. For example, on a sphere it would point out of the sphere, and on a plane it would point in the direction considered to be the top of the plane. Now,

we shine a light ray with direction v at the surface. The direction of the reflected ray will be given by v_r . What must be true is that the angle θ between the normal \hat{n} and the light ray v should be the same as the angle between the reflected ray and the normal, and all three vectors v , \hat{n} , and v_r must lie in the same plane. Given these constraints, below is one way to calculate the light reflection ray v_r .



To make the figure to the left, we first rotated the scene so everything is in a convenient orientation, with the surface normal \hat{n} pointing vertically, and the surface horizontal. Now, move vector v so that its tail is at the reflection point, as shown by the vector drawn with a dashed line in the figure. If b is the vector parallel to \hat{n} from the head of v to the surface, then by vector addition we have

$$v_r = v + 2b.$$

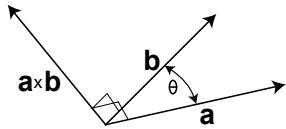
Now the vector b is just the negative of the component of v in the direction of \hat{n} . So,

$$b = -(\hat{n} \cdot v)\hat{n}.$$

Thus,

$$\mathbf{v}_r = \mathbf{v} - 2(\hat{\mathbf{n}} \cdot \mathbf{v})\hat{\mathbf{n}}.$$

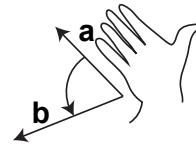
A.8 CROSS PRODUCT



The cross product $\mathbf{a} \times \mathbf{b}$ between two vectors \mathbf{a} and \mathbf{b} is a new vector perpendicular to the plane defined by the original two vectors. In other words, the cross product of two vectors is a vector that is perpendicular to both of the original vectors. The figure to the left illustrates the construction.

This notion of cross product does not make sense in 2D space, since it is not possible for a third 2D vector to be perpendicular to two (non parallel) 2D vectors. Thus, in graphics, the notion of cross product is reserved for working in 3D space.

Since there are two directions perpendicular to the plane formed by two vectors, we must have a convention to determine which of these two directions to use. In graphics, it is most common to use the *right hand rule*, and we use this convention throughout this text. The right-hand rule works as follows. Hold your right hand out flat, with the thumb out, aligning the fingers so they point in the direction of \mathbf{a} .



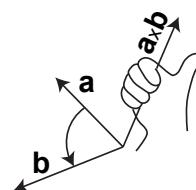
Now, rotate your hand so you can curl your fingers in the direction from vector \mathbf{a} to vector \mathbf{b} . Your thumb will point in the direction of $\mathbf{a} \times \mathbf{b}$. If you reverse this, and first align your fingers with \mathbf{b} and then curl them towards \mathbf{a} you will see that you have to turn your hand upside down, reversing the direction in which your thumb is pointing. From this it should be apparent that $\mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b})$. In other words, the order of the operands in the cross product changes the polarity of the resulting cross product vector. The result is still perpendicular to both of the original vectors, but the direction is flipped.

A.8.1 Trigonometric interpretation of cross product

The magnitude of the cross product is given by

$$\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| \sin \theta,$$

where θ is the small angle between vectors \mathbf{a} and \mathbf{b} . Thus, if \mathbf{a} and \mathbf{b} are unit vectors, the magnitude of the cross product is the magnitude of $\sin \theta$.



Note, that the cross product of two parallel vectors will be the

zero vector $\mathbf{0}$. This is consistent with the geometric notion that the cross product produces a vector orthogonal to the original two vectors. If the original vectors are parallel, then there is no unique direction perpendicular to both vectors (i.e. there are infinitely many orthogonal vectors, all parallel to any plane perpendicular to either vector).

Algebraically, the cross product is defined as follows. If two vectors are defined

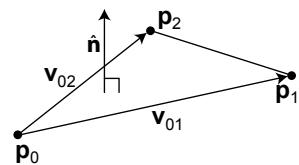
$$\mathbf{a} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}, \text{ and } \mathbf{b} = \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix},$$

then

$$\mathbf{a} \times \mathbf{b} = \begin{bmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{bmatrix}.$$

The cross product has many uses in graphics, which the following two examples will serve to illustrate.

A.8.2 Cross product example: Finding surface normals



Suppose we have triangle $(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2)$, and we want to find the triangle's surface normal. We can do this easily by use of a cross product operation. First, define vectors along two of the triangle edges: $\mathbf{v}_{01} = \mathbf{p}_1 - \mathbf{p}_0$, and $\mathbf{v}_{02} = \mathbf{p}_2 - \mathbf{p}_0$. Then the cross product $\mathbf{v}_{01} \times \mathbf{v}_{02}$ is a vector perpendicular to both \mathbf{v}_{01} and \mathbf{v}_{02} , and therefore perpendicular to the plane of the triangle. Scaling this vector to a unit vector yields the surface normal vector

$$\hat{\mathbf{n}} = (\mathbf{v}_{01} \times \mathbf{v}_{02}) / \|\mathbf{v}_{01} \times \mathbf{v}_{02}\|.$$

A.8.3 Cross product example: Computing the area of a triangle

Another application of cross product to triangles uses the trigonometric definition of the magnitude of the cross product. Suppose we have a triangle, like the one shown to the right. If we know the lengths of sides a and b , and we know the angle θ between these sides, the area computation is straightforward. Relative to side a , the height of the triangle is given by $h = b \sin \theta$, and we know that the area of the triangle is $A = 1/2ah$, so we have $A = 1/2ab \sin \theta$. If we represent the sides of the triangle by vectors \mathbf{a} and \mathbf{b} , $a = \|\mathbf{a}\|$ and $b = \|\mathbf{b}\|$. Since the magnitude of the cross product $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| |\sin \theta|$, it follows that

$$A = 1/2 \|\mathbf{a} \times \mathbf{b}\|.$$

