Chapter 7 Cross product

We are getting ready to study *integration* in several variables. Until now we have been doing only *differential* calculus. One outcome of this study will be our ability to compute volumes of interesting regions of $\mathbb{R}^n$. As preparation for this we shall learn in this chapter how to compute volumes of parallelepipeds in $\mathbb{R}^3$. In this material there is a close connection with the vector product, which we now discuss.

A. **Definition of the cross product**

We begin with a simple but interesting problem. Let x, y be given vectors in $\mathbb{R}^3$: $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$. Assume that x and y are linearly independent; in other words, 0, x, y determine a unique plane. We then want to determine a nonzero vector z which is orthogonal to this plane. That is, we want to solve the equations $x \bullet z = 0$ and $y \bullet z = 0$. We are certain in advance that z will be uniquely determined up to a scalar factor. The equations in terms of the coordinates of z are

$$\begin{aligned} x_1 z_1 + x_2 z_2 + x_3 z_3 &= 0, \\ y_1 z_1 + y_2 z_2 + y_3 z_3 &= 0. \end{aligned}$$

It is no surprise that we have two equations but three "unknowns," as we know z is not going to be unique. Since x and y are linearly independent, the matrix

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

has row rank equal to 2, and thus also has column rank 2. Thus it has two linearly independent columns. To be definite, suppose that the first two columns are independent; in other words,

$$x_1 y_2 - x_2 y_1 \neq 0.$$

(This is all a special case of the general discussion in Section 6B.) Then we can solve the following two equations for the "unknowns" z_1 and z_2:

$$\begin{aligned} x_1 z_1 + x_2 z_2 &= -x_3 z_3, \\ y_1 z_1 + y_2 z_2 &= -y_3 z_3. \end{aligned}$$

The result is

$$\begin{aligned} z_1 &= \frac{x_2y_3 - x_3y_2}{x_1y_2 - x_2y_1}\, z_3, \\ z_2 &= \frac{x_3y_1 - x_1y_3}{x_1y_2 - x_2y_1}\, z_3. \end{aligned}$$

Notice of course we have an undetermined scalar factor z_3.

Now we simply make the choice $z_3 = x_1y_2 - x_2y_1 (\neq 0)$. Then the vector z can be written

$$\begin{cases} z_1 = x_2y_3 - x_3y_2, \\ z_2 = x_3y_1 - x_1y_3, \\ z_3 = x_1y_2 - x_2y_1. \end{cases}$$

This is precisely what we were trying to find, and we now simply make this a definition:

DEFINITION. The *cross product* (or *vector product*) of two vectors x, y in $\mathbb{R}^3$ is the vector

$$x \times y = (x_2y_3 - x_3y_2,\ x_3y_1 - x_1y_3,\ x_1y_2 - x_2y_1).$$

DISCUSSION. 1. Our development was based on the assumption that x and y are linearly independent. But the *definition* still holds in the case of linear dependence, and produces $x \times y = 0$. Thus we can say immediately that

$$x \text{ and } y \text{ are linearly dependent} \iff x \times y = 0.$$

2. We also made the working assumption that $x_1y_2 - x_2y_1 \neq 0$. Either of the other two choices of independent columns produces the same sort of result. This is clearly seen in the nice symmetry of the definition.

3. The definition is actually quite easily memorized. Just realize that the first component of $z = x \times y$ is

$$z_1 = x_2y_3 - x_3y_2$$

and then a cyclic permutation $1 \to 2 \to 3 \to 1$ of the indices automatically produces the other two components.

4. A convenient mnemonic for the definition is given by the formal determinant expression,

$$x \times y = \det \begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}.$$

In this expression the entries in the first row are the standard unit coordinate vectors, and the "determinant" is to be calculated by expansion by the minors along the first row.

5. We are left with one true ambiguity in the definition, and that is which sign to take. In our development we chose $z_3 = x_1y_2 - x_2y_1$, but we could have of course chosen $z_3 = x_2y_1 - x_1y_2$. In this case, the entire mathematical community agrees with the choice we have made.

6. Nice special cases:

$$\begin{aligned} \hat{\imath} \times \hat{\jmath} &= \hat{k}, \\ \hat{\jmath} \times \hat{k} &= \hat{\imath}, \\ \hat{k} \times \hat{\imath} &= \hat{\jmath}. \end{aligned}$$

7. By the way, two vectors in $\mathbb{R}^3$ have a *dot* product (a scalar) and a *cross* product (a vector). The words "dot" and "cross" are somehow weaker than "scalar" and "vector," but they have stuck.

ALGEBRAIC PROPERTIES. The cross product is *linear* in each factor, so we have for example for vectors x, y, u, v,

$$(ax + by) \times (cu + dv) = acx \times u + adx \times v + bcy \times u + bdy \times v.$$

It is *anticommutative*:

$$y \times x = -x \times y.$$

It is not associative: for instance,

$$\begin{aligned} \hat{\imath} \times (\hat{\imath} \times \hat{\jmath}) &= \hat{\imath} \times \hat{k} = -\hat{\jmath}; \\ (\hat{\imath} \times \hat{\imath}) \times \hat{\jmath} &= 0 \times \hat{\jmath} = 0. \end{aligned}$$

PROBLEM 7–1. Let $x \in \mathbb{R}^3$ be thought of as fixed. Then $x \times y$ is a linear function from $\mathbb{R}^3$ to $\mathbb{R}^3$ and thus can be represented in a unique way as a matrix times the column vector y. Show that in fact

$$x \times y = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix} y.$$

PROBLEM 7–2. Assuming $x \neq 0$ in the preceding problem, find the characteristic polynomial of the 3×3 matrix given there. What are its eigenvalues?

B. **The norm of the cross product**

The approach I want to take here goes back to the Schwarz inequality on p. 1–15, for which we are now going to give an entirely different proof. Suppose then that $x,\, y \in \mathbb{R}^n$. We are going to prove $|x \bullet y| \leq \|x\|\ \|y\|$ by *calculating* the difference of the squares of the two sides, as follows:

$$
\begin{aligned}
\|x\|^2\ \|y\|^2 - (x \bullet y)^2 &= \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n x_i y_i\right)^2 \\
&= \sum_{i=1}^n x_i^2 \sum_{j=1}^n y_j^2 - \sum_{i=1}^n x_i y_i \sum_{j=1}^n x_j y_j \\
&= \sum_{i,j=1}^n x_i^2 y_j^2 - \sum_{i,j=1}^n x_i y_i x_j y_j \\
&= \sum_{i<j} x_i^2 y_j^2 + \sum_{i=1}^n x_i^2 y_i^2 + \sum_{i>j} x_i^2 y_j^2 \\
&\quad - \sum_{i<j} x_i y_i x_j y_j - \sum_{i=1}^n x_i^2 y_i^2 - \sum_{i>j} x_i y_i x_j y_j \\
&= \sum_{i<j} x_i^2 y_j^2 + \sum_{j>i} x_j^2 y_i^2 - 2 \sum_{i<j} x_i y_i x_j y_j \\
&= \sum_{i<j} \left(x_i^2 y_j^2 - 2 x_i y_i x_j y_j + x_j^2 y_i^2\right) \\
&= \sum_{i<j} \left(x_i y_j - x_j y_i\right)^2 .
\end{aligned}
$$

This then proves that $\|x\|^2\ \|y\|^2 - (x \bullet y)^2 \geq 0$.

If we specialize to $\mathbb{R}^3$, we have

$$\|x\|^2\ \|y\|^2 - (x \bullet y)^2 = (x_1 y_2 - x_2 y_1)^2 + (x_2 y_3 - x_3 y_2)^2 + (x_1 y_3 - x_3 y_1)^2.$$

But the right side is precisely $z_3^2 + z_1^2 + z_2^2$ in the above notation. Thus we have proved the wonderful *Lagrange's identity*

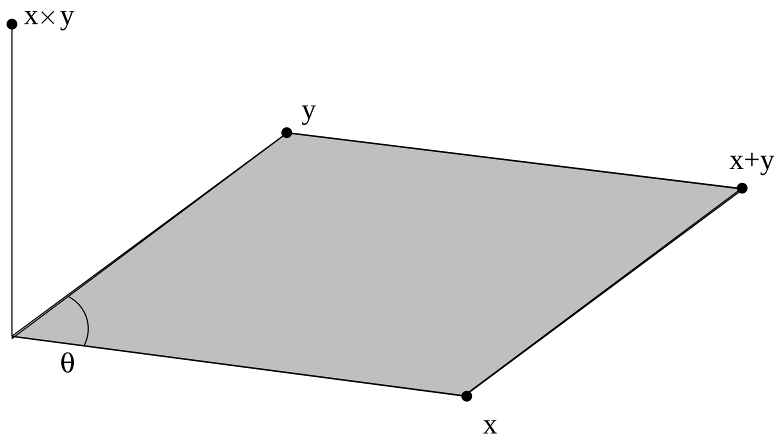
$$\boxed{\|x\|^2\ \|y\|^2\ =\ (x \bullet y)^2 + \|x \times y\|^2.}$$

This identity relates norms, dot products, and cross products.

In terms of the angle θ between x and y, we have from p. 1–17 the formula $x \bullet y = \|x\|\ \|y\| \cos\theta$. Thus,

$$\boxed{\|x \times y\|\ =\ \|x\|\ \|y\| \sin\theta.}$$

This result completes the *geometric* description of the cross product, up to sign. The vector $x \times y$ is orthogonal to the plane determined by 0, x and y, and its norm is given by the formula we have just derived. An even more geometric way of saying this is that the norm of $x \times y$ is the *area* of the parallelogram with vertices 0, x, y, $x + y$:



PROBLEM 7–3. The Lagrange identity can easily be generalized. Using the following outline, prove that for x, y, u, $v \in \mathbb{R}^3$,

$$(x \times y) \bullet (u \times v) = (x \bullet u)(y \bullet v) - (x \bullet v)(y \bullet u).$$

Outline: the right side equals

$$\begin{aligned}
\left(\sum_i x_i u_i\right)\left(\sum_j y_j v_j\right) - \left(\sum_i x_i v_i\right)\left(\sum_j y_j u_j\right) \\
&= \sum_{i,j} x_i y_j (u_i v_j - u_j v_i) && \text{(why?)} \\
&= \sum_{i,j} x_j y_i (u_j v_i - u_i v_j) && \text{(why?)} \\
&= \frac{1}{2} \sum_{i,j} (x_i y_j - x_j y_i)(u_i v_j - u_j v_i) && \text{(why?)} \\
&= \sum_{i<j} (x_i y_j - x_j y_i)(u_i v_j - u_j v_i) && \text{(why?)} \\
&= (x \times y) \bullet (u \times v).
\end{aligned}$$

C. **Triple products**

The formal representation of $x \times y$ as a determinant enables us to calculate inner products readily: for any $u \in \mathbb{R}^3$

$$\begin{aligned}
(x \times y) \bullet u &= \det \begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} \bullet \left(u_1 \hat{\imath} + u_2 \hat{\jmath} + u_3 \hat{k}\right) \\
&= \det \begin{pmatrix} u_1 & u_2 & u_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}.
\end{aligned}$$

(Notice how clearly this formula displays the fact that $(x \times y) \bullet x = (x \times y) \bullet y = 0$.)

This number $(x \times y) \bullet u$ is called the *scalar triple product* of the vectors x, y, and u. The formula we have just given makes it clear that

$$\begin{aligned}(x \times y) \bullet u &= (u \times x) \bullet y = (y \times u) \bullet x \\ &= -(y \times x) \bullet u = -(x \times u) \bullet y = -(u \times y) \bullet x.\end{aligned}$$

Thus we have the interesting phenomenon that writing x, y, u in order gives

$$(x \times y) \bullet u = x \bullet (y \times u).$$

So the triple product doesn't care which you call $\times$ and which you call $\bullet$. For this reason, it's frequently written

$$[x, y, u] = (x \times y) \bullet u.$$

PROBLEM 7–4. The *vector triple product* is $(x \times y) \times u$. It can be related to dot products by the identity

$$(x \times y) \times u = (x \bullet u)y - (y \bullet u)x.$$

Prove this by using Problem 7–3 to calculate the dot product of each side of the proposed formula with an arbitrary $v \in \mathbb{R}^3$.

PROBLEM 7–5. Prove quickly that the other vector triple product satisfies

$$x \times (y \times u) = (x \bullet u)y - (x \bullet y)u.$$

The identities in Problems 7–4 and 7–5 can be remembered by expressing the right side of each as

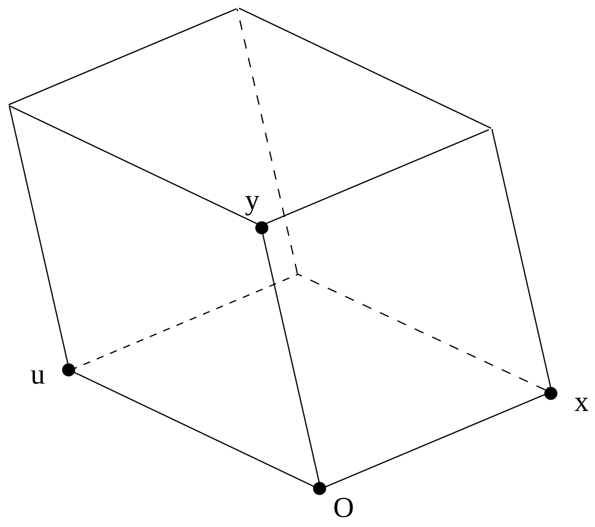
$$\text{vector triple product} = (\text{outer}\bullet\text{far})\ \text{near} - (\text{outer}\bullet\text{near})\ \text{far}.$$

The scalar triple product is exactly the device that enables us to compute volumes of parallelepipeds in $\mathbb{R}^3$. Consider three vectors x, y, u in $\mathbb{R}^3$. They generate a region D called a *parallelepiped* as follows:

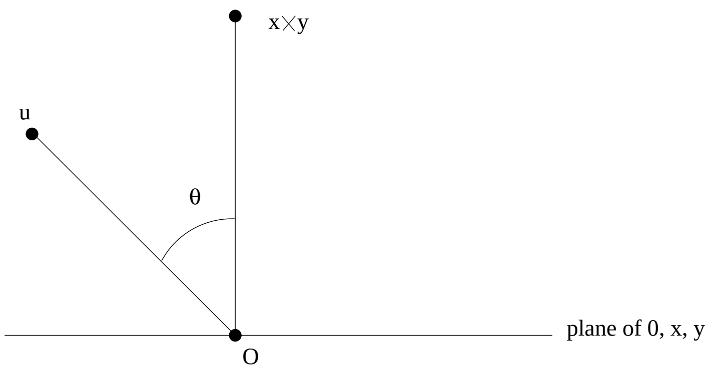
$$D = \{ax + by + cu \mid 0 \le a \le 1,\ 0 \le b \le 1,\ 0 \le c \le 1\}.$$

PROBLEM 7–6. Prove that

$$[x \times y, y \times u, u \times x] = [x, y, u]^2.$$



We use the usual intuitive definition of volume as the area of the base times the altitude. If we regard x and y as spanning the parallelogram which is the base, its area is of course $\|x \times y\|$. The altitude is then calculated as the norm of the projection of the side u onto the direction orthogonal to the base; of course, $x \times y$ has the required direction:



In terms of the angle θ between $x \times y$ and u, the altitude is

$$\begin{aligned}\|u\|\ |\cos\theta| &= \|u\|\ \frac{|(x\times y)\bullet u|}{\|x\times y\|\ \|u\|}\\ &= \frac{|(x\times y)\bullet u|}{\|x\times y\|}.\end{aligned}$$

Thus the volume of D is

$$\mathrm{vol}(D) = |(x\times y)\bullet u|.$$

It is therefore computed as the absolute value of the scalar triple product of the three spanning edges of the parallelepiped.

In terms of the determinant,

$$\mathrm{vol}(D) = \left|\ \det\begin{pmatrix} u_1 & u_2 & u_3\\ x_1 & x_2 & x_3\\ y_1 & y_2 & y_3\end{pmatrix}\ \right|.$$

Notice in particular that $\mathrm{vol}(D) = 0 \iff$ the vectors x, y, u are linearly dependent $\iff$ the parallelepiped D lies in a two dimensional subspace of $\mathbb{R}^n$.

D. **Orientation**

We now have the cross product characterized geometrically, up to a sign. Given linearly independent x, $y \in \mathbb{R}^3$, the cross product has *direction* orthogonal to both x and y, and *norm* equal to $\|x\|\ \|y\|\sin\theta$, where θ is the angle between the vectors x and y. There are exactly two vectors with these properties, and we now want to discuss a geometric way to single out the cross product.

The concept we need is important in general dimensions, so we shall deal with $\mathbb{R}^n$ for any n.

DEFINITION. A *frame* for $\mathbb{R}^n$ is an ordered basis. That is, we have n linearly independent vectors $\varphi_1, \dots, \varphi_n$, and they are written in a particular order. For example,

$$\{\hat\imath, \hat\jmath, \hat k\},\quad \{\hat\imath + 2\hat\jmath, \hat\jmath, -\hat k\},\quad \{\hat\imath, \hat k, \hat\jmath\}$$

are three distinct frames for $\mathbb{R}^3$.

DEFINITION. Let $\{\varphi_1, \dots, \varphi_n\}$ and $\{\psi_1, \dots, \psi_n\}$ be frames for $\mathbb{R}^n$. Then there exists a unique $n \times n$ matrix A such that

$$\varphi_i = A\psi_i, \qquad 1 \le i \le n.$$

This matrix is necessarily invertible. We say the two frames have the *same orientation* if $\det A > 0$; and *opposite orientation* if $\det A < 0$.

PROBLEM 7–7. In the above definition we have used the column vector representation of vectors. We may then write the corresponding $n \times n$ matrices

$$\begin{aligned} \Phi &= (\varphi_1 \dots \varphi_n), \\ \Psi &= (\psi_1 \dots \psi_n). \end{aligned}$$

Prove that the frames have the same orientation $\iff$ $\det \Phi$ and $\det \Psi$ have the same sign.

PROBLEM 7–8. Prove that "having the same orientation" is an equivalence relation for frames. (Thus, as far as orientation is concerned, there are just two types of frames.)

PROBLEM 7–9. Prove that if $\{\varphi_1, \dots, \varphi_n\}$ and $\{\psi_1, \dots, \psi_n\}$ happen to be *orthonormal* frames, the matrix A in the definition is in $O(n)$; and the frames have the same orientation $\iff A \in \mathrm{SO}(n)$.

Let us say that the coordinate frame $\{\hat{e}_1, \dots, \hat{e}_n\}$ (also written $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ in the case of $\mathbb{R}^3$) has the *standard* orientation. The corresponding matrix is the identity matrix I, and $\det I = 1$. Thus we say that an arbitrary frame $\{\varphi_1, \dots, \varphi_n\}$ has the *standard* orientation if the corresponding matrix $\Phi = (\varphi_1 \dots \varphi_n)$ has $\det \Phi > 0$.

PROBLEM 7–10. For which dimensions do both $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$ and $\{\varphi_n, \varphi_1, \varphi_2, \dots, \varphi_{n-1}\}$ have the same orientation?

PROBLEM 7–11. For which dimensions do $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$ and its "reversal" $\{\varphi_n, \dots, \varphi_2, \varphi_1\}$ have the same orientation?

Now we consider $\mathbb{R}^3$ and two vectors x, y whose cross product $x \times y \neq 0$. Writing everything as column vectors, we then have from Section C that

$$\det(x \; y \; x \times y) \;=\; \|x \times y\|^2 \;>\; 0.$$

Thus the frame $\{x, y, \; x \times y\}$ has the standard orientation. This fact is what settles the choice of direction of the cross product.

SUMMARY. The cross product $x \times y$ is uniquely determined by the geometric description:

(1) $x \times y$ is orthogonal to x and y,

(2) $\|x \times y\| =$ area of the parallelogram with edges x and y,

(3) in case $x \times y \neq 0$, the frame $\{x,\ y,\ x \times y\}$ has the standard orientation.

Since the cross product is characterized geometrically, we would expect the action of elements of $O(3)$ to have a predictable outcome. Indeed this is so:

THEOREM. Let $A \in O(3)$. Then

$$A(x \times y) \;=\; (\det A) Ax \times Ay.$$

Thus in case $A \in \mathrm{SO}(3)$, then $A(x \times y) = Ax \times Ay$.

PROOF. We use the characterization in terms of column vectors

$$(x \times y) \bullet u \;=\; \det(x\ y\ u), \qquad \text{all}\ \ u \in \mathbb{R}^3.$$

Then for any 3×3 real matrix A,

$$\begin{aligned} (Ax \times Ay) \bullet Au &= \det(Ax\ Ay\ Au) \\ &= \det A(x\ y\ u) \\ &= \det A \det(x\ y\ u) \\ &= \det A (x \times y) \bullet u. \end{aligned}$$

But

$$(Ax \times Ay) \bullet Au \;=\; A^t(Ax \times Ay) \bullet u.$$

Since we have two vectors with the same inner product with all $u \in \mathbb{R}^3$, we thus conclude

$$A^t(Ax \times Ay) \;=\; \det A(x \times y).$$

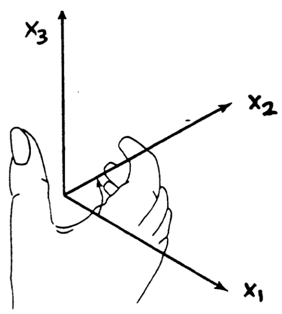
If A is also invertible, then

$$Ax \times Ay \;=\; \det A (A^t)^{-1}(x \times y).$$

Finally, if $A \in O(3)$, then $(A^t)^{-1} = A$ and $\det A = \pm 1$.

QED

The outcome of our discussion is that the cross product enjoys the double-edged properties of being characterized completely geometrically and of being easily computed in coordinates. In Chapter 1 we experienced the same happy situation with dot products on $\mathbb{R}^n$. The dot product $x \bullet y$ can be completely characterized as $\|x\| \, \|y\| \cos\theta$, where θ is the angle between x and y, but also it is easily computed in coordinates.



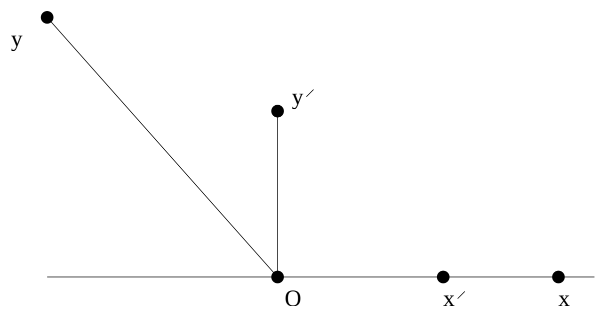
E. **Right hand rule**

This section just gives another way of presenting orientation. Given a Euclidean three-dimensional space, we may think of first choosing an origin and then placing an orthogonal coordinate system x_1, x_2, x_3. We say this is a *right-handed* coordinate system if when you place your right hand so that your fingers curl from the positive x_1-axis through 90 degrees to the positive x_2-axis, your thumb points in the direction of the positive x_3-axis.

Another way of saying this is to place your right hand so that your fingers curl from $\hat{\imath}$ to $\hat{\jmath}$. Then your thumb points in the direction $\hat{k}$ if the coordinate system is right-handed.

Notice that $\hat{\imath} \times \hat{\jmath} = \hat{k}$. This then gives the *right-hand rule* for cross products, supposing we have a right-handed coordinate system. Supposing that $x \times y \neq 0$, place your right hand so that your fingers curl from x toward y through an acute angle. Then your thumb points in the direction of $x \times y$.

This is easily proved by the device of using SO(3) to move your right hand from the $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ frame to an orthonormal frame associated with the frame $\{x, y, x \times y\}$. Here's the picture, looking "down" at the plane determined by 0, x, y:



Let $x' = x/\|x\|$, and let y' be the unit vector shown:

$$y' \;=\; \frac{y - x' \bullet y\; x'}{\|y - x' \bullet y\; x'\|}\;.$$

Thus $y = c_1 x' + c_2 y'$, where $c_2 > 0$. Then

$$x \times y \;=\; c_2 x \times y' \;=\; c_2 \|x\|\; x' \times y'.$$

Thus $x \times y$ is a positive scalar multiple of $x' \times y'$. Now define the matrix A in terms of its columns,

$$A \;=\; (x'\; y'\; x' \times y').$$

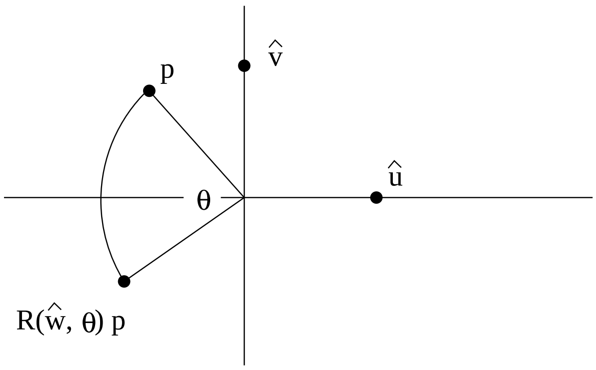
Then $A\hat{\imath} = x'$, $A\hat{\jmath} = y'$, $A\hat{k} = x' \times y'$. And we know $A \in \mathrm{SO}(3)$. Thus A takes the right-handed frame $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ to the right-handed frame $\{x', y', x' \times y'\}$. This concludes the proof of the right-hand rule.

F. **What *is* SO(3)?**

This seems to be a good place to discuss the fine details of SO(3), the special orthogonal group on $\mathbb{R}^3$. The members of this group are of course 3×3 orthogonal matrices with determinants equal to 1. Let $A \in \mathrm{SO}(3)$. This matrix of course "acts on" $\mathbb{R}^3$ by matrix multiplication, sending column vectors x to the column vectors Ax. We are going to prove that this "action" can be described as an ordinary *rotation* of $\mathbb{R}^3$ about an axis.

We first pause to describe rotations. Suppose that $\hat{w}$ is a fixed unit vector in $\mathbb{R}^3$, and that θ is a fixed angle. We use the standard right-handed orientation for $\mathbb{R}^3$. Now we take a view of $\mathbb{R}^3$ looking from $\hat{w}$ towards 0, and rotate the entire space counter-clockwise about the axis $\hat{w}$ through the angle θ. Let us call the resulting matrix $R(\hat{w}, \theta)$. A couple of things should be noticed. One is that a counter-clockwise rotation with a negative angle θ is actually a clockwise rotation with the positive angle $-\theta$. The other is that adding any integer multiple of 2π to θ results in the same rotation, according to our convention here.

Now we show how to calculate $R(\hat{w}, \theta)$. To do this use a right-handed orthonormal frame $\{\hat{u}, \hat{v}, \hat{w}\}$. This can be computed by choosing *any* unit vector $\hat{u}$ which is orthogonal to $\hat{w}$, and then $\hat{v}$ is uniquely determined by the required formula $\hat{v} = \hat{w} \times \hat{u}$. Now think of projecting $\mathbb{R}^3$ orthogonally onto the $\hat{u}$, $\hat{v}$ plane, and then rotating through the angle θ. Here's the view looking down from $\hat{w}$:



Thus we have

$$\begin{aligned} R(\hat{w}, \theta)\hat{u} &= \hat{u}\cos\theta + \hat{v}\sin\theta, \\ R(\hat{w}, \theta)\hat{v} &= -\hat{u}\sin\theta + \hat{v}\cos\theta. \end{aligned}$$

Of course, since $\hat{w}$ is the axis of rotation,

$$R(\hat{w}, \theta)\hat{w} = \hat{w}.$$

The summary of this information is expressed in terms of matrices as

$$\begin{aligned} R(\hat{w}, \theta)(\hat{u}\ \hat{v}\ \hat{w}) &= (\hat{u}\cos\theta + \hat{v}\sin\theta \quad -\hat{u}\sin\theta + \hat{v}\cos\theta \quad \hat{w}) \\ &= (\hat{u}\ \hat{v}\ \hat{w}) \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Therefore

$$R(\hat{w}, \theta) = (\hat{u}\ \hat{v}\ \hat{w}) \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} (\hat{u}\ \hat{v}\ \hat{w})^{-1}. \qquad (*)$$

Remember: the computation of the inverse of the orthogonal matrix $(\hat{u}\ \hat{v}\ \hat{w})$ is just a matter of writing down its transpose

$$\begin{pmatrix} \hat{u}^t \\ \hat{v}^t \\ \hat{w}^t \end{pmatrix}.$$

Based on our understanding of what $R(\hat{w}, \theta)$ is, we are certain that this formula does not really depend on the choice of $\hat{u}$ and $\hat{v}$. See Problem 7–16 for the explicit verification of this fact.

You should notice that the inner matrix in the above formula is precisely $R(\hat{k}, \theta)$. So the formula displays how we are able to move from a rotation about the x_3-axis to a rotation about the axis determined by $\hat{w}$.

PROBLEM 7–12. Prove that $R(\hat{w}, \theta) \in \mathrm{SO}(3)$.

PROBLEM 7–13. Using the above formula, show directly that

$$R(\hat{w}, 0) = I.$$

Then give a heuristic explanation of this equation.

PROBLEM 7–14. Using the above formula, show directly that

$$R(-\hat{w}, -\theta) = R(\hat{w}, \theta).$$

Then give a heuristic explanation of this equation. Also explain why we may thus assume without loss of generality that $0 \le \theta \le \pi$.

PROBLEM 7–15. Using the above formula, calculate directly that

$$R(\hat{w}, \theta_1) R(\hat{w}, \theta_2) = R(\hat{w}, \theta_1 + \theta_2).$$

Then give a heuristic explanation of this equation.

PROBLEM 7–16. Calculate the product and thus obtain from (∗)

$$R(\hat{w},\theta) \;=\; (\cos\theta)I + (1-\cos\theta)(w_iw_j) + \sin\theta \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix}.$$

(This displays $R(\hat{w},\theta)$ in a form that shows its independence of the choice of $\hat{u}$ and $\hat{v}$.)

PROBLEM 7–17. Show that the action of $R(\hat{w},\theta)$ can also be expressed in the form

$$R(\hat{w},\theta)x \;=\; (\cos\theta)x + (1-\cos\theta)(\hat{w}\bullet x)\hat{w} + (\sin\theta)\hat{w}\times x,$$

for all $x \in \mathbb{R}^3$. Do this two ways:

a. by simply using Problem 7–16;

b. by describing the action of $R(\hat{w},\theta)$ geometrically.

PROBLEM 7–18. Prove that for all $x \bullet \hat{w} = 0$,

$$R(\hat{w},\pi)x \;=\; -x.$$

Give a geometric interpretation of this result. Conversely, suppose that for a given rotation there exists a nonzero $x \in \mathbb{R}^3$ such that

$$R(\hat{w},\theta)x \;=\; -x.$$

Prove that $\theta = \pi$ and $x \bullet \hat{w} = 0$. (Of course, θ could also be $\pi + 2m\pi$ for any integer m.)

PROBLEM 7–19. Prove that $R(\hat{w},\theta)$ is a symmetric matrix $\Longleftrightarrow$ $\theta = 0$ or π.

PROBLEM 7–20. Backward Andy all his life has insisted upon using a left-handed coordinate system for $\mathbb{R}^3$. What formula must he use for $R(\hat{w}, \theta)$ that is the analog of ours found in Problem 7–16?

We now turn to the perhaps surprising converse of the above material, which states that the *only* matrices in SO(3) are rotations! The crucial background for this result is contained in the following

THEOREM. Let $A \in \mathrm{SO}(n)$, where n is odd. Then 1 is an eigenvalue of A.

PROOF. Since $A^{-1} = A^t$, we have

$$\begin{aligned} \det(A - I) &= \det(A - AA^t) \\ &= \det(A(I - A^t)) \\ &= \det A \det(I - A^t) \\ &= \det(I - A^t) \\ &= \det(I - A) \\ &= (-1)^n \det(A - I). \end{aligned}$$

Since n is odd, $\det(A - I) = 0$.

QED

PROBLEM 7–21. Give a different proof of this result for the case $n = 3$, based on the characteristic polynomial

$$\det(A - \lambda I) = -\lambda^3 + \alpha\lambda^2 + \beta\lambda + 1.$$

PROBLEM 7–22. Let $A \in O(n)$, but $A \notin \mathrm{SO}(n)$. Prove that -1 is an eigenvalue of A.

THEOREM. Let $A \in \mathrm{SO}(3)$. Then there exists a unit vector $\hat{w} \in \mathbb{R}^3$ and $0 \leq \theta \leq \pi$ such that $A = R(\hat{w}, \theta)$.

PROOF. Since $A \in \mathrm{SO}(3)$, we know from the preceding theorem that 1 is an eigenvalue of A. Therefore, there exists a unit vector $\hat{w}$ such that $A\hat{w} = \hat{w}$.

Choose an orthonormal basis $\hat{u}$, $\hat{v}$ of the subspace of $\mathbb{R}^3$ which is orthogonal to $\hat{w}$; arrange them so that the frame $\{\hat{u}, \hat{v}, \hat{w}\}$ has the standard orientation:

$$\begin{aligned}\hat{w} &= \hat{u} \times \hat{v}, \\ \hat{v} &= \hat{w} \times \hat{u}, \\ \hat{u} &= \hat{v} \times \hat{w}.\end{aligned}$$

Then

$$\begin{aligned}A\hat{u} \bullet \hat{w} &= A\hat{u} \bullet A\hat{w} \\ &= \hat{u} \bullet \hat{w} \qquad (A \text{ is in } O(3)) \\ &= 0,\end{aligned}$$

so $A\hat{u}$ is also orthogonal to $\hat{w}$. Thus $A\hat{u}$ is a linear combination of $\hat{u}$ and $\hat{v}$:

$$A\hat{u} = c_1\hat{u} + c_2\hat{v}.$$

Since $\|A\hat{u}\| = \|\hat{u}\| = 1$, $c_1^2 + c_2^2 = 1$. Thus there exists an angle θ such that $c_1 = \cos\theta$, $c_2 = \sin\theta$. That is,

$$A\hat{u} = \hat{u}\cos\theta + \hat{v}\sin\theta.$$

We conclude that

$$\begin{aligned}A\hat{v} &= A(\hat{w} \times \hat{u}) \\ &= A\hat{w} \times A\hat{u} \qquad \text{(Section D)} \\ &= \hat{w} \times (\hat{u}\cos\theta + \hat{v}\sin\theta) \\ &= \hat{v}\cos\theta - \hat{u}\sin\theta.\end{aligned}$$

We now see that A and the rotation $R(\hat{w}, \theta)$ produce the same result on the frame $\{\hat{u}, \hat{v}, \hat{w}\}$! Since every $x \in \mathbb{R}^3$ is a linear combination of $\hat{u}$, $\hat{v}$, $\hat{w}$, we conclude that

$$Ax = R(\hat{w}, \theta)x \qquad \text{for all } x \in \mathbb{R}^3.$$

Thus $A = R(\hat{w}, \theta)$.

The final item, that we may require $0 \le \theta \le \pi$, follows immediately from Problem 7–14.

QED

DISCUSSION. For a given $A \in \mathrm{SO}(3)$, the representation $A = R(\hat{w}, \theta)$, where $\hat{w}$ is a unit vector and $0 \le \theta \le \pi$, is virtually unique. To see this we start with the result of Problem 7–16, and compute the trace of each matrix:

$$\begin{aligned} \operatorname{trace} A &= 3\cos\theta + (1 - \cos\theta)(w_1^2 + w_2^2 + w_3^2) \\ &= 3\cos\theta + (1 - \cos\theta) \\ &= 2\cos\theta + 1. \end{aligned}$$

Thus

$$\cos\theta = \frac{\operatorname{trace} A - 1}{2}.$$

Therefore, as $0 \le \theta \le \pi$, the angle is completely determined by the knowledge of $\operatorname{trace} A$.

Next notice that Problem 7–16 implies

$$A - A^t = 2\sin\theta \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix}.$$

We conclude immediately that if $0 < \theta < \pi$, the vector w is also uniquely determined.

The only cases of nonuniqueness are thus when $\sin\theta = 0$, or, equivalently, $A = A^t$. If $\theta = 0$, then $A = R(\hat{w}, 0) = I$ and the "axis" $\hat{w}$ is irrelevant. If $\theta = \pi$, then we have

$$A = -I + 2(w_i w_j).$$

It is easy to check in this case that $\hat{w}$ is uniquely determined up to sign. This is because $\hat{w} \ne 0$ and so some coordinate $w_j \ne 0$ and we obtain from the above equation

$$2 w_j \hat{w} = j^{\text{th}} \text{ column of } (A + I).$$

This determines $\hat{w}$ up to a scalar factor, and the fact that $\|\hat{w}\| = 1$ shows that the only possibilities are $\hat{w}$ and $-\hat{w}$.

We now have a truly intuitive understanding of $\mathrm{SO}(3)$. In particular, the right-hand rule makes wonderful sense in this context, as we now explain. Suppose $\{\hat{\varphi}_1, \hat{\varphi}_2, \hat{\varphi}_3\}$ is an arbitrary orthonormal frame with standard orientation, and form the usual matrix Φ with those columns,

$$\Phi = (\hat{\varphi}_1 \; \hat{\varphi}_2 \; \hat{\varphi}_3).$$

Then $\Phi \in \mathrm{SO}(3)$ and Φ^{-1} maps the frame $\{\hat{\varphi}_1, \hat{\varphi}_2, \hat{\varphi}_3\}$ to $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$. Moreover, Φ is a rotation $R(\hat{w}, \theta)$. Thus if you place your right hand to lie in the correct $\hat{k} = \hat{\imath} \times \hat{\jmath}$ position, then rotating your hand with $R(\hat{w}, \theta)$ produces the $\hat{\varphi}_3 = \hat{\varphi}_1 \times \hat{\varphi}_2$ position.

I would say it is almost amazing that a single rotation can move $\{\hat{\varphi}_1, \hat{\varphi}_2, \hat{\varphi}_3\}$ to $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$! Here is a numerical example. Let

$$A = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Clearly, $A \in \mathrm{SO}(3)$. Then we compute $\mathrm{trace} A = 0$, so that $\cos\theta = -\frac{1}{2}$ and we have $\theta = 2\pi/3$. Then

$$A - A^t = 2\sin\theta \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix}$$

becomes the equation

$$\begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix} = \sqrt{3} \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix},$$

so we read off immediately that

$$\hat{w} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Thus

$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} = R(\hat{w}, 2\pi/3).$$

PROBLEM 7–23. Let

$$A = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix}.$$

Find $\hat{w}$ and θ such that $A = R(\hat{w}, \theta)$.

PROBLEM 7–24. Compute

$$R(\hat{w}, \theta)^{-1} = R(?, ?).$$

PROBLEM 7–25. Show that the given matrix is in SO(3) and calculate the unique $\hat{w} \in \mathbb{R}^3$ and $0 \le \theta \le \pi$ for it:

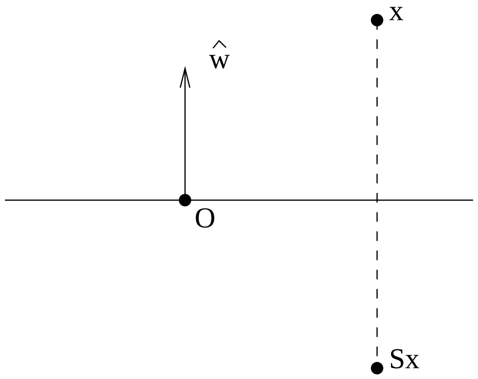
$$\frac{1}{45}\begin{pmatrix} 29 & -20 & 28 \\ 28 & 35 & -4 \\ -20 & 20 & 35 \end{pmatrix}.$$

G. **Orientation reversal**

It is a relatively easy task to understand those matrices in $O(n)$ which are not in SO(n). Suppose A is such a matrix. Of course, $\det A = -1$. We conclude from Problem 7–22 that -1 is an eigenvalue of A. Therefore, there exists a unit vector $\hat{w}$:

$$A\hat{w} = -\hat{w}.$$

Any unit vector $\hat{w}$ can be used to define a "reflection" matrix S in that direction. Here's a schematic description:



The algebra is governed by

$$\begin{aligned} (Sx + x) \bullet \hat{w} &= 0, \\ Sx &= x + t\hat{w} \quad \text{for some } \ t \in \mathbb{R}. \end{aligned}$$

Solving these equations gives $t = -2x \bullet \hat{w}$ so that

$$Sx = x - 2(x \bullet \hat{w})\hat{w}.$$

This matrix S is called "reflection in the direction $\hat{w}$," and we denote it $S(\hat{w})$.

PROBLEM 7–26. Prove that the matrix $S = S(\hat{w})$ satisfies

a. $S\hat{w} = -\hat{w}$,

b. $S^2 = I$,

c. $S \in O(n)$,

d. $\det S = -1$.

PROBLEM 7–27. Prove that

$$S\hat{w} \;=\; I - 2(w_i w_j).$$

Now we return to the situation of a matrix $A \in O(n)$ with $\det A = -1$. We know from Problem 7–22 that there is a unit vector $\hat{w}$ satisfying $A\hat{w} = -\hat{w}$. Now just define $B = S(\hat{w})A$. Clearly, $B \in \mathrm{SO}(n)$. Thus,

$$A \;=\; S(\hat{w})B.$$

Thus A is the product of a reflection $S(\hat{w})$ in a direction $\hat{w}$ and a matrix $B \in \mathrm{SO}(n)$ with $B\hat{w} = \hat{w}$.

PROBLEM 7–28. It made no difference whether we chose to write $B = SA$ or $B = AS$, for $SA = AS$. Prove this by noting that $SA\hat{w} = AS\hat{w}$ and then considering those vectors x which are orthogonal to $\hat{w}$.

In particular, for $n = 3$ the matrix A has the representation

$$A \;=\; S(\hat{w})R(\hat{w},\theta) \;=\; R(\hat{w},\theta)S(\hat{w}).$$

PROBLEM 7–29. Show that in this $n = 3$ situation

$$A \;=\; (\cos\theta)I - (1+\cos\theta)(w_i w_j) + \sin\theta \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix}.$$

PROBLEM 7–30. Show that in case $n = 2$ every matrix $A \in O(2)$ with $\det A = -1$ is equal to a reflection matrix $S(\hat{w})$.

PROBLEM 7–31. Refer back to inversion in the unit sphere in $\mathbb{R}^n$, p. 6–34:

$$f(x) = \frac{x}{\|x\|^2}.$$

Prove that

$$Df(x) = \|x\|^{-2} S\left(\frac{x}{\|x\|}\right).$$

Thus we say that inversion in the unit sphere is an *orientation—reversing* conformal mapping.

PROBLEM 7–32. Let u, v be linearly independent vectors in $\mathbb{R}^3$ and let u', v' also be vectors in $\mathbb{R}^3$.

a. Suppose that $A \in \mathrm{SO}(3)$ satisfies

$$Au = u',$$
$$Av = v'.$$

Prove that u' and v' are linearly independent and

$$\begin{cases} \|u\| = \|u'\|, \\ \|v\| = \|v'\|, \\ u \bullet v = u' \bullet v'. \end{cases} \qquad (*)$$

b. Conversely, suppose that $(*)$ is satisfied. Prove that u' and v' are linearly independent, and that there exists a unique $A \in \mathrm{SO}(3)$ such that

$$Au = u',$$
$$Av = v'.$$

(HINT: $A(u\ v\ u \times v) = (u'\ v'\ u' \times v')$.)

PROBLEM 7–33. Under the assumptions of the preceding problem prove that

$$(u\ v\ u \times v)^{-1} = \|u \times v\|^{-2}(v \times (u \times v) \quad -u \times (u \times v) \quad u \times v)^t.$$