Chapter 8 Volumes of parallelograms

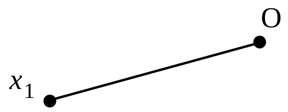
In the present short chapter we are going to discuss the elementary geometrical objects which we call parallelograms. These are going to be contained in some ambient Euclidean space, but each will have its own innate dimension. We shall usually denote the ambient space as $\mathbb{R}^N$, and we use its usual inner product $\bullet$ and norm. In discussing these parallelograms, we shall always for simplicity assume one vertex is the origin $0 \in \mathbb{R}^N$. This restriction is of course easily overcome in applications by a simple translation of $\mathbb{R}^N$.

DEFINITION. Let $x_1, \dots, x_n$ be arbitrary points in $\mathbb{R}^N$ (notice that these subscripts do not here indicate coordinates). We then form the set of all linear combinations of these points of a certain kind:

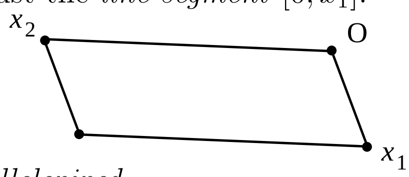
$$P = \left\{ \sum_{i=1}^{n} t_i x_i \mid 0 \le t_i \le 1 \right\}.$$

We call this set an n-dimensional *parallelogram* (with one "vertex" 0). We also refer to the vectors $x_1, \dots, x_n$ as the *edges* of P.

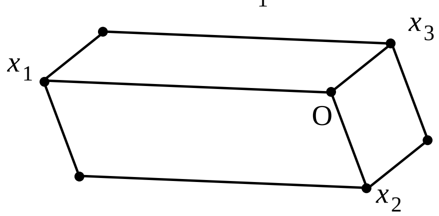
- For $n = 1$ this "parallelogram" is of course just the *line segment* $[0, x_1]$.



- For $n = 2$ we obtain a "true" parallelogram.



- For $n = 3$ the word usually employed is *parallelepiped*. However, it seems to be a good idea to pick one word to be used for all n, so we actually call this a 3-dimensional "parallelogram."



We are not assuming that the edge vectors $x_1, \dots, x_n$ are linearly independent, so it could happen that a 3-dimensional parallelogram could lie in a plane and itself be a 2-dimensional parallelogram, for instance.

PROBLEM 8–1. Consider the "parallelogram" in $\mathbb{R}^3$ with "edges" equal to the **three** points $\hat{\imath}$, $\hat{\jmath}$, $\hat{\imath} - 2\hat{\jmath}$. Draw a sketch of it and conclude that it is actually a six-sided figure in the $x - y$ plane.

A. **Volumes in dimensions 1, 2, 3**

We are going to work out a definition of the n-dimensional volume of an n-dimensional parallelogram P, and we shall denote this as vol(P); if we wish to designate the dimension, we write $\text{vol}_n(P)$. In this section we are going to recall some formulas we have already obtained, and cast them in a somewhat different format.

• **Dimension 1**. We have the line segment $[0, x_1] \subset \mathbb{R}^N$, and we simply say its 1-dimensional volume is its *length*, i.e. the norm of x_1:

$$\text{vol}([0, x_1]) = \|x_1\|.$$

• **Dimension 3**. We skip ahead to this dimension because of our already having obtained the pertinent formula in Section 7C. Namely, suppose $P \subset \mathbb{R}^3$, so we are dealing at first with the dimension of P equal to the dimension of the ambient space. Then if we write x_1, x_2, x_3 as column vectors,

$$\text{vol}(P) = |\det(x_1\ x_2\ x_3)|,$$

the absolute value of the determinant of the 3×3 matrix formed from the column vectors.

This formula has the advantage of being easy to calculate. But it suffers from two related disadvantages: (1) it depends on the *coordinates* of the edge vectors and thus is not very geometric, and (2) it requires the ambient space to be $\mathbb{R}^3$.

Both of these objections can be overcome by resorting to the by now familiar device of computing the matrix product

$$\begin{aligned}(x_1\ x_2\ x_3)^t(x_1\ x_2\ x_3) &= \begin{pmatrix} x_1^t \\ x_2^t \\ x_3^t \end{pmatrix} (x_1\ x_2\ x_3) \\ &= \begin{pmatrix} x_1 \bullet x_1 & x_1 \bullet x_2 & x_1 \bullet x_3 \\ x_2 \bullet x_1 & x_2 \bullet x_2 & x_2 \bullet x_3 \\ x_3 \bullet x_1 & x_3 \bullet x_2 & x_3 \bullet x_3 \end{pmatrix}.\end{aligned}$$

Notice that whereas both matrices on the left side of this equation depend on the coordinates of x_1, x_2, and x_3, the matrix $(x_i \bullet x_j)$ on the right side depends only on the (geometric) inner products of the vectors! The properties of determinant produce the formula

$$(\det(x_1\ x_2\ x_3))^2 = \det(x_i \bullet x_j).$$

Thus we have obtained the formula

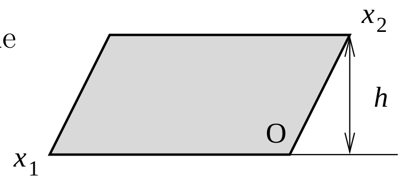
$$\text{vol}(P) = \sqrt{\det(x_i \bullet x_j)}.$$

This is our desired formula. It displays vol(P) in such a way that we no longer need the assumption $P \subset \mathbb{R}^3$. For if the ambient space is $\mathbb{R}^N$, we can simply regard x_1, x_2, x_3 as lying in a 3-dimensional subspace of $\mathbb{R}^N$ and use the formula we have just derived. The point is that the inner products $x_i \bullet x_j$ from $\mathbb{R}^N$ are the same inner products as in the subspace of $\mathbb{R}^N$, since the inner product has a completely geometric description.

• **Dimension 2**. Here we are going to present the procedure for obtaining the formula analogous to the above, but without using a coordinate representation at all. We shall see in Section C that this procedure works for all dimensions.

We suppose P to be the 2-dimensional parallelogram in $\mathbb{R}^N$ with vertices 0, x_1, x_2, and $x_1 + x_2$. We choose the "base" of the parallelogram to be $[0, x_1]$ and then use the classical formula for the area of P:

$$\begin{aligned} \operatorname{vol}(P) &= \text{base times altitude} \\ &= \|x_1\| h \end{aligned}$$



(see the figure). The altitude h is the distance from x_2 to the "foot" tx_1 of the line segment from x_2 orthogonal to the base. The definition of t is then that

$$(tx_1 - x_2) \bullet x_1 = 0.$$

(How many times have we used this technique in this book?!) The definition of h then can be rewritten

$$\begin{aligned} h^2 &= \|x_2 - tx_1\|^2 = (x_2 - tx_1) \bullet (x_2 - tx_1) \\ &= (x_2 - tx_1) \bullet x_2. \end{aligned}$$

We now regard what we have obtained as two linear equations for the two "unknowns" t and h^2:

$$\begin{cases} tx_1 \bullet x_1 + 0h^2 &= x_1 \bullet x_2, \\ tx_1 \bullet x_2 + h^2 &= x_2 \bullet x_2. \end{cases}$$

We don't care what t is in this situation, so we use Cramer's rule to compute only the desired h^2:

$$h^2 = \frac{\det\begin{pmatrix} x_1 \bullet x_1 & x_1 \bullet x_2 \\ x_1 \bullet x_2 & x_2 \bullet x_2 \end{pmatrix}}{\det\begin{pmatrix} x_1 \bullet x_1 & 0 \\ x_1 \bullet x_2 & 1 \end{pmatrix}} = \frac{\det(x_i \bullet x_j)}{\|x_1\|^2}.$$

Thus

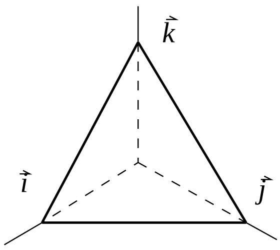
$$\mathrm{vol}(P) = \sqrt{\det(x_i \bullet x_j)}.$$

This is the desired analog to what we found for $n = 3$, but we have obtained it without using coordinates and without even thinking about $\mathbb{R}^2$!

EXAMPLE. We find the area of the triangle in $\mathbb{R}^3$ with vertices $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$. We handle this by finding the area of an associated parallelogram and then dividing by 2 (really, 2!). A complication appears because we do not have the origin of $\mathbb{R}^3$ as a vertex, but we get around this by thinking of vectors emanating from one of the vertices (say $\vec{\jmath}$) thought of as the origin. Thus we take

$$x_1 = \vec{\imath} - \vec{\jmath} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad x_2 = \vec{k} - \vec{\jmath} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Then the square of the area of the parallelogram is



$$\det(x_i \bullet x_j) = \det\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = 3.$$

Thus,

$$\text{area of triangle} = \frac{1}{2}\sqrt{3}.$$

PROBLEM 8–2. Compute the area of the triangle in $\mathbb{R}^3$ whose vertices are $a\hat{\imath}$, $b\hat{\jmath}$, and $c\hat{k}$. Express the result in a form that is a symmetric function of a, b, c.

B. **The Gram determinant**

Our work thus far indicates that the matrix of dot products $(x_i \bullet x_j)$ will play a major role in these sorts of calculations. We therefore pause to consider it as an object of tremendous interest in its own right. It is named after the Danish mathematician Jorgen Pedersen Gram.

DEFINITION. Suppose x_1, $x_2, \dots, x_n$ are points in the Euclidean space $\mathbb{R}^N$. Their *Gram matrix* is the $n \times n$ symmetric matrix of their inner products

$$G = (x_i \bullet x_j) \; = \begin{pmatrix} x_1 \bullet x_1 & x_1 \bullet x_2 & \dots & x_1 \bullet x_n \\ \vdots & \vdots & & \vdots \\ x_n \bullet x_1 & x_n \bullet x_2 & \dots & x_n \bullet x_n \end{pmatrix}.$$

The *Gram determinant* of the same points is

$$\begin{aligned} \operatorname{Gram}(x_1, x_2, \dots, x_n) &= \det G \\ &= \det(x_i \bullet x_j). \end{aligned}$$

Here are the first three cases:

$$\operatorname{Gram}(x) = x \bullet x = \|x\|^2;$$

$$\begin{aligned} \operatorname{Gram}(x_1, x_2) &= \det \begin{pmatrix} \|x_1\|^2 & x_1 \bullet x_2 \\ x_2 \bullet x_1 & \|x_2\|^2 \end{pmatrix} \\ &= \|x_1\|^2 \|x_2\|^2 - (x_1 \bullet x_2)^2; \end{aligned}$$

$$\operatorname{Gram}(x_1, x_2, x_3) = \det \begin{pmatrix} \|x_1\|^2 & x_1 \bullet x_2 & x_1 \bullet x_3 \\ x_1 \bullet x_2 & \|x_2\|^2 & x_2 \bullet x_3 \\ x_1 \bullet x_3 & x_2 \bullet x_3 & \|x_3\|^2 \end{pmatrix}.$$

Notice that for $n = 1$, $\operatorname{Gram}(x) \geq 0$ and $\operatorname{Gram}(x) > 0$ unless $x = 0$. For $n = 2$, $\operatorname{Gram}(x_1, x_2) \geq 0$ and $\operatorname{Gram}(x_1, x_2) > 0$ unless x_1 and x_2 are linearly dependent; this is the Schwarz inequality. We didn't encounter $\operatorname{Gram}(x_1, x_2, x_3)$ until the preceding section, but we realize that the analogous property is valid. In fact, we shall now give an independent verification of this property in general.

THEOREM. *If $x_1, \dots, x_n \in \mathbb{R}^N$, then their Gram matrix G is positive semidefinite. In particular,*

$$\operatorname{Gram}(x_1, \dots, x_n) \geq 0.$$

Moreover, G is positive definite $\iff$

$$\operatorname{Gram}(x_1, \dots, x_n) \;>\; 0$$

$\iff x_1, \dots, x_n$ *are linearly independent.*

PROOF. We first notice the formula for the norm of $\sum_{i=1}^n t_i x_i$:

$$\begin{aligned} \Big\| \sum_{i=1}^n t_i x_i \Big\|^2 &= \sum_{i=1}^n t_i x_i \bullet \sum_{j=1}^n t_j x_j \\ &= \sum_{i,j=1}^n (x_i \bullet x_j) t_i t_j. \end{aligned}$$

In terms of the Gram matrix, the calculation we have just done can thus be written in the form

$$\Big\| \sum_{i=1}^n t_i x_i \Big\|^2 = \; Gt \bullet t,$$

where t is the column vector with components $t_1, \dots, t_n$.

We see therefore that $Gt \bullet t \geq 0$ for all $t \in \mathbb{R}^n$. That is, the matrix G is positive semidefinite (p. 4–26). Thus $\det G \geq 0$, thanks to the easy analysis on p. 4–26. (Remember: the eigenvalues of G are all nonnegative and $\det G$ is the product of the eigenvalues.) This proves the first statement of the theorem.

For the second we first note that of course G is positive definite $\iff \det G > 0$. If these conditions do not hold, then there exists a nonzero $t \in \mathbb{R}^n$ such that $Gt = 0$. Then of course also $Gt \bullet t = 0$. According to the formula above, $\sum_{i=1}^n t_i x_i$ has zero norm and is therefore the zero vector. We conclude that $x_1, \dots, x_n$ are linearly dependent. Conversely, assume $\sum_{i=1}^n t_i x_i = 0$ for some $t \neq 0$. Take the dot product of this equation with each x_j to conclude that $\sum_{i=1}^n t_i x_i \bullet x_j = 0$. This says that the j^{th} component of $Gt = 0$. As j is arbitrary, $Gt = 0$. Thus $\det G = 0$.

QED

PROBLEM 8–3. For the record, here is the general principle used above: prove that for a general real symmetric $n \times n$ matrix A which is positive semidefinite and for $t \in \mathbb{R}^n$, $At \bullet t = 0 \iff At = 0$.

The following problem strikingly shows that Gram matrices provide a concrete realization of all positive semidefinite matrices.

PROBLEM 8–4. Let A be a real symmetric positive semidefinite $n \times n$ matrix. Prove that A is a Gram matrix; that is, there exist $x_1, \ldots, x_n \in \mathbb{R}^n$ such that $A = (x_i \bullet x_j)$. (HINT: use $\sqrt{A}$.)

PROBLEM 8–5. Let A be a real symmetric positive definite $n \times n$ matrix. Suppose A is represented as a Gram matrix in two different ways:

$$A = (x_i \bullet x_j) = (y_i \bullet y_j).$$

Prove that

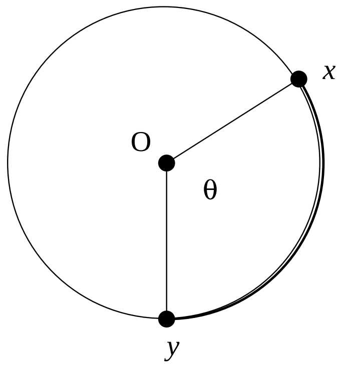
$$x_i = By_i, \qquad 1 \le i \le n,$$

where B is a matrix in $O(n)$. Prove the converse as well.

As we noted earlier, the inequality $\mathrm{Gram}(x_1, x_2) \ge 0$ is exactly the Schwarz inequality. We pause to demonstrate geometrical significance of the next case, $\mathrm{Gram}(x_1, x_2, x_3) \ge 0$. Expanding the determinant, this says that

$$\|x_1\|^2\|x_2\|^2\|x_3\|^2 + 2(x_1 \bullet x_2)(x_2 \bullet x_3)(x_3 \bullet x_1) \\ - \|x_1\|^2(x_2 \bullet x_3)^2 - \|x_2\|^2(x_3 \bullet x_1)^2 - \|x_3\|^2(x_1 \bullet x_2)^2 \ge 0.$$

Now consider points which lie on the unit sphere in $\mathbb{R}^N$. These are points with norm 1. For any two such points x and y, we can think of the distance between them, measured not from the viewpoint of the ambient $\mathbb{R}^N$, but rather from the viewpoint of the sphere itself. We shall later discuss this situation. The idea is to try to find the curve of shortest length that starts at x, ends at y, and that lies entirely in the sphere. We shall actually prove that this curve is the *great circle* passing through x and y. Here's a side view, where we have intersected $\mathbb{R}^N$ with the unique 2-dimensional plane containing 0, x and y (minor technicality: if $y = \pm x$ then many planes will do):



The angle θ between x and y is itself the length of the great circle arc joining x and y. This is of course calculated by the dot product: $\cos\theta = x \bullet y$. As $-1 \le x \bullet y \le 1$, we then define the *intrinsic distance* from x to y by the formula

$$d(x, y) = \arccos x \bullet y.$$

(Remember that arccos has values only in the interval $[0, \pi]$.) The following properties are immediate:

$$0 \le d(x, y) \le \pi;$$
$$d(x, y) = 0 \iff x = y;$$
$$d(x, y) = d(y, x).$$

It certainly seems reasonable that the triangle inequality for this distance is true, based on the fact that the curve of shortest length joining two points on the sphere should be the great circle arc and therefore should have length $d(x, y)$. However, we want to show that the desired inequality is intimately related to the theorem. The inequality states that

$$d(x, y) \le d(x, z) + d(z, y).$$

We now prove this inequality. It states that

$$\arccos x \bullet y \le \arccos x \bullet z + \arccos z \bullet y.$$

If the right side of this equation is greater than π, then we have nothing to prove. We may thus assume it is between 0 and π. On this range cosine is a strictly decreasing function, so we must prove

$$x \bullet y \ge \cos(\arccos x \bullet z + \arccos z \bullet y).$$

The addition formula for cosine enables a computation of the right side; notice that

$$\sin(\arccos x \bullet z) \;=\; \sqrt{1-(x \bullet z)^2}.$$

Thus we must prove

$$x \bullet y \geq (x \bullet z)(z \bullet y) - \sqrt{1-(x \bullet z)^2}\sqrt{1-(z \bullet y)^2}.$$

That is,

$$\sqrt{1-(x \bullet z)^2}\sqrt{1-(z \bullet y)^2} \geq (x \bullet z)(z \bullet y) - x \bullet y.$$

If the right side is negative, we have nothing to prove. Thus it suffices to square both sides and prove that

$$(1-(x \bullet z)^2)(1-(z \bullet y)^2) \geq \left((x \bullet z)(z \bullet y) - x \bullet y\right)^2.$$

A little calculation shows this reduces to

$$1 + 2(x \bullet z)(z \bullet y)(x \bullet y) - (x \bullet z)^2 - (z \bullet y)^2 - (x \bullet y)^2 \geq 0.$$

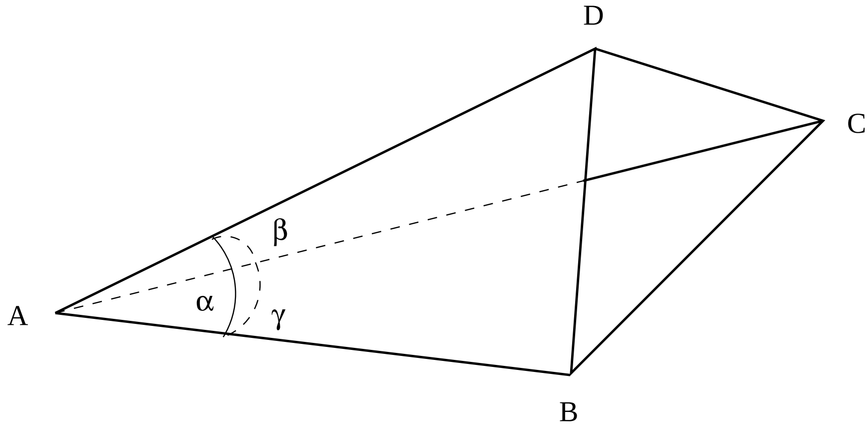
Aha! This is exactly the inequality $\mathrm{Gram}(x, y, z) \geq 0$.

PROBLEM 8–6. In the above situation, show that equality holds in

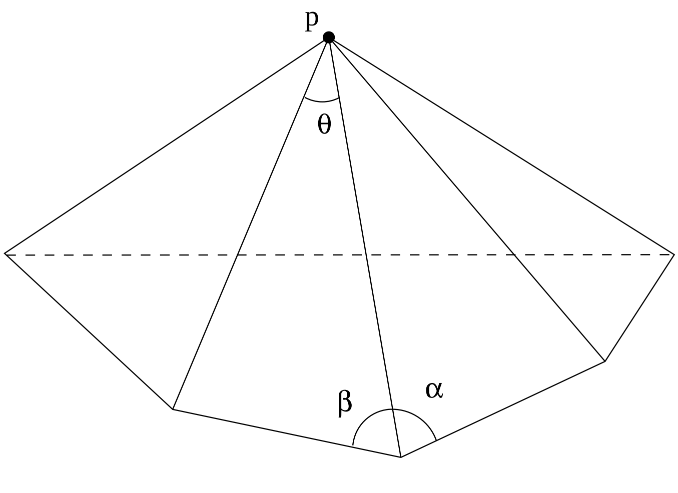
$$\mathrm{dist}(x, y) \;=\; \mathrm{dist}(x, z) + \mathrm{dist}(z, y)$$

if and only if z lies on the "short" great circle arc joining x and y.

PROBLEM 8–7. (Euclid) Show that when three rays in $\mathbb{R}^3$ make a solid angle in space, the sum of any two of the angles between them exceeds the third:

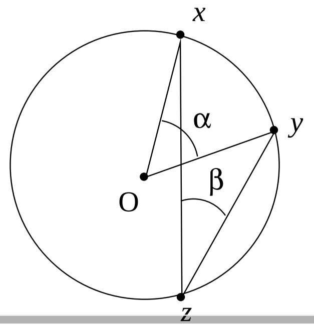


PROBLEM 8–8. Suppose m rays meet in convex fashion at a point p in $\mathbb{R}^3$. Show that the angles at p sum to less than 2π.
(HINT: sum all the angles α, β, θ over all the triangles. Use the preceding problem.)



PROBLEM 8–9*. Here is a familiar and wonderful theorem of plane Euclidean geometry. Consider the angles shown in a circle with center 0. Both angles subtend the same arc of the circle. We say α is a *central* angle, β an *inscribed* angle. The theorem asserts that $\alpha = 2\beta$.

Now label the three points on the circle as shown and show that the equation $\alpha = 2\beta$ is exactly the equation $\mathrm{Gram}(x, y, z) = 0$.

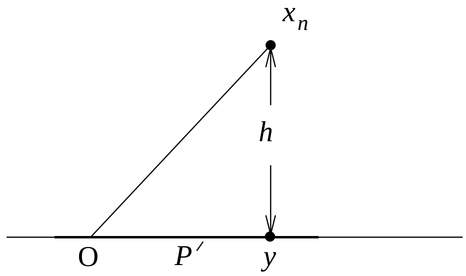


C. **Volumes in all dimensions**

We are now ready to derive the general formula for volumes of parallelograms. We shall proceed inductively, and shall take an intuitive viewpoint: the n-dimensional parallelogram P with edges $x_1, \dots, x_n$ has as its "base" the $(n-1)$-dimensional parallelogram P' with edges $x_1, \dots, x_{n-1}$. We then compute the altitude h as the distance to the base, and *define*

$$\mathrm{vol}_n(P) \;=\; \mathrm{vol}_{n-1}(P')h.$$

"Side view":



This "definition" is not really adequate, but the calculations are too interesting to miss. Moreover, *Fubini's theorem* in the next chapter provides the simple theoretical basis for the validity of the current naive approach.

To complete this analysis we simply follow the outline given in Section A for the case $n = 2$. That is, we want to investigate the "foot" $y = \sum_{j=1}^{n-1} t_j x_j$ of the altitude. We have

$$(y - x_n) \bullet x_i \;=\; 0 \qquad \text{for} \;\; 1 \le i \le n-1.$$

These equations should determine the t_i's. Then the altitude can be computed by the equation

$$\begin{aligned} h^2 &= \|x_n - y\|^2 \\ &= (x_n - y) \bullet (x_n - y) \\ &= (x_n - y) \bullet x_n. \end{aligned}$$

We now display these n equations for our "unknowns" $t_1, \dots, t_{n-1}$, h^2:

$$\begin{cases} \sum_{j=1}^{n-1} t_j x_i \bullet x_j + 0h^2 &= x_i \bullet x_n, \qquad 1 \le i \le n-1; \\ \sum_{j=1}^{n-1} t_j x_n \bullet x_j + h^2 &= x_n \bullet x_n. \end{cases}$$

Cramer's rule should then produce the desired formula for h^2:

$$h^2 = \frac{\det \begin{pmatrix} x_1 \bullet x_1 & \dots & x_1 \bullet x_{n-1} & x_1 \bullet x_n \\ x_2 \bullet x_1 & \dots & x_2 \bullet x_{n-1} & x_2 \bullet x_n \\ \vdots & & & \\ x_n \bullet x_1 & \dots & x_n \bullet x_{n-1} & x_n \bullet x_n \end{pmatrix}}{\det \begin{pmatrix} x_1 \bullet x_1 & \dots & x_1 \bullet x_{n-1} & 0 \\ x_2 \bullet x_1 & \dots & x_2 \bullet x_{n-1} & 0 \\ \vdots & & & \\ x_n \bullet x_1 & \dots & x_n \bullet x_{n-1} & 1 \end{pmatrix}}.$$

Aha! This is precisely

$$h^2 = \frac{\mathrm{Gram}(x_1, \dots, x_n)}{\mathrm{Gram}(x_1, \dots, x_{n-1})}.$$

Of course, this makes sense only if the determinant of the coefficients in the above linear system is nonzero. That is, only if $x_1, \dots, x_{n-1}$ are linearly independent. Notice then that $h^2 = 0 \iff x_1, \dots, x_n$ are linearly dependent; that is, $\iff x_n$ lies in the $(n-1)$-dimensional subspace containing the "base" of the parallelogram.

Thanks to this formula, we now see inductively that

$$\boxed{\mathrm{vol}_n(P) = \sqrt{\mathrm{Gram}(x_1, \dots, x_n)}.}$$

PROBLEM 8–10. In the special case of an n-dimensional parallelogram P with edges $x_1, \dots, x_n \in \mathbb{R}^n$, show that

$$\mathrm{vol}_n(P) = |\det(x_1 \; x_2 \; \dots \; x_n)|.$$

D. **Generalization of the cross product**

This is a good place to give a brief discussion of "cross products" on $\mathbb{R}^n$. We argue by analogy from the case $n = 3$. Suppose then that $x_1, \dots, x_{n-1}$ are linearly independent vectors in $\mathbb{R}^n$. Then we want to find a vector z which is orthogonal to all of them. We definitely anticipate that z will exist and will be uniquely determined up to a scalar factor.

The formula we came up with for the case $n = 3$ can be presented as a formal 3×3 determinant as

$$z = \det \begin{pmatrix} & & \hat{\imath} \\ x_1 & x_2 & \hat{\jmath} \\ & & \hat{k} \end{pmatrix},$$

where x_1 and x_2 are column vectors and we expand the "determinant" along the third column. This is a slight rearrangement of the formula we derived in Section 7A.

We now take the straightforward generalization for any dimension $n \geq 2$:

$$z = \det \begin{pmatrix} & & & \hat{e}_1 \\ & & & \hat{e}_2 \\ x_1 & x_2 \dots & x_{n-1} & \vdots \\ & & & \vdots \\ & & & \hat{e}_n \end{pmatrix}.$$

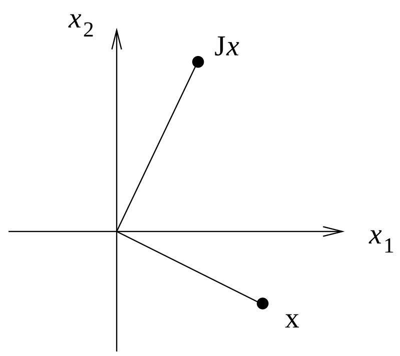
This is a formal $n \times n$ determinant, in which the last column contains the unit coordinate vectors in order. The other columns are of course the usual column vectors.

Notice the special case $n = 2$: we have just one vector, say

$$x = \begin{pmatrix} a \\ b \end{pmatrix},$$

and the corresponding vector z is given by

$$\begin{aligned} z &= \det \begin{pmatrix} a & \hat{e}_1 \\ b & \hat{e}_2 \end{pmatrix} \\ &= a\hat{e}_2 - b\hat{e}_1 \\ &= \begin{pmatrix} -b \\ a \end{pmatrix}. \end{aligned}$$



NOTATION. For $n \geq 3$ we denote

$$z = x_1 \times x_2 \times \cdots \times x_{n-1},$$

and we call z the *cross product* of the vectors $x_1, \dots, x_{n-1}$. For $n = 2$ this terminology doesn't work, so we write

$$z = Jx,$$

where J can actually be regarded as the 2×2 matrix

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

(Another way to think of this is to use complex notation $x = a + ib$, and then $Jx = ix$ $(= -b + ia)$.) If we depict $\mathbb{R}^2$ as in the above figure, then J is a 90° counterclockwise rotation. (Thus J is in SO(2).)

We now list the properties of the cross product. We shall not explicitly state the corresponding easy properties for the case $n = 2$.

PROPERTIES.

1. $x_1 \times \cdots \times x_{n-1}$ is a linear function of each factor.

2. Interchanging two factors in the cross product results in multiplying the product by -1.

3. The cross product is uniquely determined by the formula

$$(x_1 \times \cdots \times x_{n-1}) \bullet y = \det(x_1\ x_2 \cdots x_{n-1}\ y), \quad \text{all } y \in \mathbb{R}^n.$$

4. $(x_1 \times \cdots \times x_{n-1}) \bullet x_i = 0$.

5. $\|x_1 \times \cdots \times x_{n-1}\|^2 = \det(x_1 \dots x_{n-1}\ x_1 \times \cdots \times x_{n-1})$.

6. $x_1 \times \cdots \times x_{n-1} = 0 \Longleftrightarrow x_1, \dots, x_{n-1}$ are linearly dependent.
(Proof: $\Leftarrow$ follows from 5, as the right side of 5 is zero. To prove the converse, assume $x_1, \dots, x_{n-1}$ are linearly independent. Choose any $y \in \mathbb{R}^n$ such that $x_1, \dots, x_{n-1}, y$ are linearly independent. Then 3 implies $(x_1 \times \dots \times x_{n-1}) \bullet y \neq 0$. Thus $x_1 \times \dots \times x_{n-1} \neq 0$.)

7. $\|x_1 \times \cdots \times x_{n-1}\| = \sqrt{\text{Gram}(x_1, \dots, x_{n-1})}$.
(Proof: let $z = x_1 \times \cdots \times x_{n-1}$. Then 5 implies

$$\begin{aligned} \|z\|^4 &= (\det(x_1 \dots x_{n-1}\ z))^2 \\ &= \text{Gram}(x_1, \dots, x_{n-1}, z). \end{aligned}$$

Now in the computation of this Gram determinant all the terms $x_i \bullet z = 0$. Thus

$$\begin{aligned} \|z\|^4 &= \det \begin{pmatrix} x_1 \bullet x_1 & \dots & x_1 \bullet x_{n-1} & 0 \\ \vdots & & & \\ x_{n-1} \bullet x_1 & \dots & x_{n-1} \bullet x_{n-1} & 0 \\ 0 & \dots & 0 & z \bullet z \end{pmatrix} \\ &= \|z\|^2 \text{Gram}(x_1, \dots, x_{n-1}).) \end{aligned}$$

8. If $x_1, \dots, x_{n-1}$ are linearly independent, then the frame $\{x_1, \dots, x_{n-1}, x_1 \times \dots \times x_{n-1}\}$ has the standard orientation.
(Proof: the relevant determinant equals $\|x_1 \times \dots \times x_{n-1}\|^2 > 0$, by 5.)

This completes the *geometric* description of the cross product. Exactly as on p. 7–9, we have the following

SUMMARY. The cross product $x_1 \times \dots \times x_{n-1}$ of $n-1$ vectors in $\mathbb{R}^n$ is uniquely determined by the geometric description:

(1) $x_1 \times \dots \times x_{n-1}$ is orthogonal to $x_1, \dots, x_{n-1}$,

(2) $\|x_1 \times \dots \times x_{n-1}\|$ is the volume of the parallelogram with edges $x_1, \dots, x_{n-1}$,

(3) in case $x_1 \times \dots \times x_{n-1} \neq 0$, the frame $\{x_1, \dots, x_{n-1}, x_1 \times \dots \times x_{n-1}\}$ has the standard orientation.

PROBLEM 8–11. Conclude immediately that

$$\hat{e}_1 \times \dots \times \hat{e}_n \ \ (\text{with } \hat{e}_i \text{ omitted}) \ = (-1)^{n+i} \hat{e}_i.$$

PROBLEM 8–12. Now prove that

$$\begin{gathered} (x_1 \times \dots \times x_{n-1}) \bullet \hat{e}_1 \times \dots \times \hat{e}_n \ \ (\text{with } \hat{e}_i \text{ omitted}) \\ = \det((x_1 \dots x_{n-1}) \ (\text{with row } i \text{ omitted})). \end{gathered}$$

PROBLEM 8–13. Now prove that

$$(x_1 \times \ldots \times x_{n-1}) \bullet (y_1 \times \ldots \times y_{n-1}) = \det(x_i \bullet y_j).$$

(HINT: each side of this equation is multilinear and alternating in its dependence on the y_j's. Therefore it suffices to check the result when the y_j's are themselves unit coordinate vectors.)

The result of the last problem is a generalization of the seventh property above, and is also a generalization of the Lagrange identity given in Problem 7–3.

PROBLEM 8–14. In the preceding problem we must be assuming $n \geq 3$. State and prove the correct statement corresponding to $\mathbb{R}^2$.