Determinants, Areas and Volumes

Theodore Voronov

December 4, 2005

The aim of these notes is to relate the algebraic notion of determinant with the geometric notions of area and volume (and their multidimensional generalizations) in order to make it possible to use these concepts in further geometry and calculus/analysis courses.

Contents

§1 Determinants

You might have met determinants of small order (two by two and three by three) in high school algebra and you will definitely study determinants in detail in the linear algebra course. Here we shall remind the definition and basic facts about determinants in the form most convenient for seeing their geometric meaning.

1.1 Matrices and vectors

We shall consider a set denoted $\mathbb{R}^n$. By definition, its elements are arrays of real numbers of length n:

$$\mathbb{R}^n = \{\mathbf{x} = (x_1, \ldots, x_n) \mid x_i \in \mathbb{R} \text{ for all } i = 1, \ldots, n\}\,.$$

We shall call elements of $\mathbb{R}^n$, *vectors*. The set $\mathbb{R}^n$ is called the n-dimensional arithmetic space. For a given vector $\mathbf{x}$ the numbers x_i are called its *coordinates* or *components*. Vectors can be added and multiplied by numbers componentwise:

Example 1.1. $(-2, 5, 1, 0) + (0, 1, 7, -1) = (-2, 6, 8, -1)$,
$3(-2, 5, 1, 0) = (-6, 15, 3, 0)$

Clearly, we can add vectors only of the same size, i.e., for a fixed n. Later on you will study the general notion of a 'vector space' of dimension n and will learn that $\mathbb{R}^n$ is the first example of such a space.

Vectors in $\mathbb{R}^n$ can be written as rows: $\mathbf{x} = (x_1, \ldots, x_n)$, or as columns:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}.$$

If we need to make a distinction, then we refer to these two ways of representing vectors as to *row-vectors* or *column-vectors*. One can consider rectangular arrays of more general kind, with n rows and m columns:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}$$

Such arrays are called *matrices* or $n \times m$ *matrices* when it is necessary to indicate the dimensions. The numbers ij in the notation a_{ij} serve as labels specifying the element in the ith row and jth column. They are pronounced, for example "two, three" for a_{23}, not "twenty three". The elements of a matrix are also called *matrix entries.* Row-vectors and column-vectors are particular cases of matrices (with only one row or only one columns, respectively). An important role is played by matrices with $n = m$, called *square matrices.*

Example 1.2. $A = \begin{pmatrix} 2 & 0 \\ -1 & 8 \end{pmatrix}$ is a 2×2 matrix (a square matrix). Here $a_{11} = 2$, $a_{12} = 0$, $a_{21} = -1$, $a_{22} = 8$.

Example 1.3. There are two special matrices. The *zero matrix*

$$\mathbf{0} = \begin{pmatrix} 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{pmatrix}$$

(all entries are zeros) is defined for any n and m. In particular, there is the zero vector in $\mathbb{R}^n$, all coordinates of which are zero. When $n = m$, i.e., among square matrices, there is another important matrix called the *identity matrix*:

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Its entries on the diagonal are 1 and all others are 0. An alternative notation is I. If we need to stress the dimension, the we can write E_n or I_n.

In the same way as for vectors, matrices of a fixed size can be multiplied by numbers and added or subtracted. This is done entry by entry:

$$\begin{pmatrix} 1 & 3 & 0 \\ -2 & 0 & 5 \end{pmatrix} + \begin{pmatrix} 0 & 2 & 1 \\ 1 & 4 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 1 \\ -1 & 4 & 3 \end{pmatrix}$$

$$7 \begin{pmatrix} 2 & 2 \\ -3 & 0 \end{pmatrix} = \begin{pmatrix} 14 & 14 \\ -21 & 0 \end{pmatrix}$$

A non-trivial operation in the matrix algebra is the **matrix multiplication**. It is introduced as follows. First we define the *product* of a row-vector $\mathbf{a} = (a_1, \dots, a_n)$ with a column-vector $\mathbf{b} = \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix}$ as the number

$\mathbf{ab} = a_1b_1 + \cdots + a_nb_n$. This extends to more general matrices. If we are given a matric A with n rows and p columns, and a matrix B with p rows and m columns, then each row of A can be multiplied with each column of B according to the above rule, giving, by the definition, a matrix element of the *matrix product* of A and B:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & \dots & b_{1m} \\ b_{21} & \dots & b_{2m} \\ \dots & \dots & \dots \\ b_{p1} & \dots & b_{pm} \end{pmatrix},$$

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1p}b_{p1} & \dots & a_{11}b_{1m} + a_{12}b_{2m} + \cdots + a_{1p}b_{pm} \\ \dots & \dots & \dots \\ a_{n1}b_{11} + a_{n2}b_{21} + \cdots + a_{np}b_{p1} & \dots & a_{n1}b_{1m} + a_{n2}b_{2m} + \cdots + a_{np}b_{pm} \end{pmatrix}.$$

In other words, we take the products of the first row of A with all the columns of B and thus obtain the first row of AB; then take the products of the second row of A with the columns of B to obtain the second row of AB, and so on. The product of an $n \times p$ matrix with a $p \times m$ matrix will be an $n \times m$ matrix. This contains the product of a row-vector and a column-vector, giving a number (a '1×1 matrix'), as a particular case.

The matrix multiplication is associative, i.e., satisfies $A(BC) = (AB)C$, and distributive w.r.t. the addition of matrices. However, in general it is not true that AB equals BA even if the dimensions are matching. One can check that $A\mathbf{0} = \mathbf{0}A = \mathbf{0}$ and $AE = EA = A$ for all matrices A.

1.2 Determinant as a multilinear alternating function

The determinant $\det A$ is a particular polynomial function of the matrix entries of a square matrix A. It has the appearance of the sum of products of elements each belonging to a different row of the matrix A (and also to a different column) taken with certain signs.

Example 1.4. For a 1×1 matrix, which is just a number, $A = a$, the determinant coincides with this number: $\det A = a$.

Example 1.5. For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we have

$$\det A = ad - bc$$

(you can memorize this formula).

Example 1.6. For a 3×3 matrix $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ we have

$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31}\,.$$

(you do not have to memorize this formula).

Determinants appear in innumerable applications and are, without exaggeration, the most important algebraic notion.

The easiest way to actually define the determinant, so that it will be clear how to calculate it for a particular matrix and to see which properties does it have in general, is to do so by using the axiomatic method.

Consider a square $n \times n$ matrix A as consisting of n column-vectors (we treat each column of A as a vector). So we can write $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ where $\mathbf{a}_i$ stands for the ith column of A.

Theorem 1.1. *There is exists a unique polynomial function $D(A)$ of the matrix entries of A such that when considered as a function $D(\mathbf{a}_1, \dots, \mathbf{a}_n)$ of the columns $\mathbf{a}_1, \dots, \mathbf{a}_n$ it has the following properties:*

(1) if any column is multiplied by a number, it can be taken out:

$$D(\mathbf{a}_1, \dots, c\mathbf{a}_k, \dots, \mathbf{a}_n) = c\, D(\mathbf{a}_1, \dots, \mathbf{a}_k, \dots, \mathbf{a}_n)$$

(for all $k = 1, 2, \dots, n$);

(2) if any column is substituted by the sum of two arbitrary column-vectors, the function will become the sum:

$$D(\mathbf{a}_1, \dots, \underbrace{\mathbf{a}' + \mathbf{a}''}_{k\text{th place}}, \dots, \mathbf{a}_n) = D(\mathbf{a}_1, \dots, \mathbf{a}', \dots, \mathbf{a}_n) + D(\mathbf{a}_1, \dots, \mathbf{a}'', \dots, \mathbf{a}_n)$$

(for all $k = 1, 2, \dots, n$);

(3) it any two columns are swapped, then the function changes sign:

$$D(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n) = -D(\mathbf{a}_2, \mathbf{a}_1, \mathbf{a}_3, \dots, \mathbf{a}_n)$$

(and the same for any pair of columns $\mathbf{a}_i$ and $\mathbf{a}_j$, where $i, j = 1, \dots, n$);

(4) for the identity matrix, $D(E) = 1$.

Proof. We shall give a proof for the case $n = 2$, and it will be clear how the same logic works for an arbitrary n. Assume that the properties (1) to (4) hold for a function $D(A)$. As we shall see, this will allow to obtain a unique formula for $D(A)$. (This establishes the uniqueness part of the theorem.) Taking it as the definition of $D(A)$, we shall notice that the properties will be indeed satisfied. (This establishes the existence part.) So consider a 2×2 matrix A. Assuming the existence of $D(A)$ with the above properties we obtain:

$$D(A) = D(\mathbf{a}_1, \mathbf{a}_2) = D\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = D\big(a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2, a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2\big) =$$
$$a_{11}a_{12}\, D\big(\mathbf{e}_1, \mathbf{e}_1\big) + a_{11}a_{22}\, D\big(\mathbf{e}_1, \mathbf{e}_2\big) + a_{21}a_{12}\, D\big(\mathbf{e}_2, \mathbf{e}_1\big) + a_{21}a_{22}\, D\big(\mathbf{e}_2, \mathbf{e}_2\big)\,.$$

Here $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. We used the properties (1) and (2) to "open the brackets". Now we shall use the property (3). (Note that it implies vanishing of $D(A)$ if two columns of the matrix coincide.) We have, continuing the calculation,

$$D(A) = a_{11}a_{22}\, D\big(\mathbf{e}_1, \mathbf{e}_2\big) + a_{21}a_{12}\, D\big(\mathbf{e}_2, \mathbf{e}_1\big) =$$
$$a_{11}a_{22}\, D\big(\mathbf{e}_1, \mathbf{e}_2\big) - a_{21}a_{12}\, D\big(\mathbf{e}_1, \mathbf{e}_2\big) = \Big(a_{11}a_{22} - a_{21}a_{12}\Big)\, D\big(\mathbf{e}_1, \mathbf{e}_2\big)\,,$$

where we used $D(\mathbf{e}_1, \mathbf{e}_1) = D(\mathbf{e}_2, \mathbf{e}_2) = 0$ and $D(\mathbf{e}_1, \mathbf{e}_2) = -D(\mathbf{e}_2, \mathbf{e}_1)$. Finally, noticing that $D(\mathbf{e}_1, \mathbf{e}_2) = D(E) = 1$, we arrive at the formula

$$D(A) = a_{11}a_{22} - a_{21}a_{12}.$$

It can be immediately seen that the function defined by this explicit formula will enjoy all the properties (1) to (4). The theorem is proved for $n = 2$. Similar argument works for all n. □

Functions satisfying (3) are called *alternating* or *skew-symmetric*. Properties (1) and (2) are referred to as *linearity*. 'Multilinearity' means linearity w.r.t. each of the arguments.

Definition 1.1. The function $D(A)$ defined by the properties (1)–(4) of Theorem 1.1 is called the *determinant* of a matrix A and denoted either $\det A$ or $|A|$ or

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

(we use round brackets for a matrix and vertical straight lines, for its determinant).

We shall shortly see how such a definition by a set of properties or axioms leads to practical ways of calculating determinants. Let us make some remarks.

Remark 1.1. From the proof of Theorem 1.1 follows that an arbitrary function of $D(A)$ possessing the properties (1)–(3) but not necessarily (4), is unique up to a factor:

$$D(A) = \det A \cdot D(E)\,.$$

Remark 1.2. We have already noticed that if two columns of a matrix A coincide, then $\det A = 0$. *If for a matrix A, a column is replaced by the sum with a multiple of another column, then the determinant will not change.* (That is, if a column $\mathbf{a}_i$ is replaced by $\mathbf{a}_i + c\,\mathbf{a}_j$, for an arbitrary j, then the determinant of the new matrix will be the same $\det A$. Indeed, the new determinant will be the sum of $\det A$ and the one with two repeating columns $\mathbf{a}_j$, at the ith and jth places, which is zero.)

Remark 1.3. The property (1) in Theorem 1.1, in fact, implies the property (2), as long as we consider polynomial functions. Indeed, a polynomial function of a vector with the property $f(c\,\mathbf{a}) = c\,f(\mathbf{a})$ can only be of degree one, i.e., a linear function.

1.3 Properties of determinants

In this subsection we shall give the main properties of determinants. They all follow from the fundamental properties of skew-symmetricity, multilinearity and $\det E = 1$ that we have used for the definition of determinant. Proofs can be found at the end of the subsection.

The most important property of determinants is contained in the following statement.

Theorem 1.2. *Determinant is a multiplicative function of a matrix: for arbitrary $n \times n$ matrices A and B,*

$$\det(AB) = \det A \cdot \det B\,.$$

Remark 1.4. The determinant of the product is the product of determinants. You should be warned that it is not true that the determinant of the sum is the sum of determinants! In fact, there is a formula for $\det(A + B)$, but it is quite complicated.

We have considered determinant as a function of the matrix columns. What about the properties of $\det A$ as a function of rows?

Theorem 1.3. *The determinant* $\det A$ *of a square matrix* A *is a multilinear alternating function of the rows of* A*. (That is,* $\det A$ *possesses the same properties (1)–(3) w.r.t. the rows as it has w.r.t. the columns. It is uniquely defined by these conditions and the condition* $\det E = 1$*.)*

For an $n \times m$ matrix A with the entries a_{ij}, the *transpose* of A, notation: A^T, is defined as the $m \times n$ matrix of the form

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix}.$$

In other words, the columns of A^T are the rows of A written as column-vectors, and vice versa. The (ij)-th matrix element of A^T is a_{ji}, the (ji)-th element of A.

Example 1.7.

$$\begin{pmatrix} 2 & 3 \\ -1 & 0 \end{pmatrix}^T = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$$

Theorem 1.4. *For any square matrix* A*,*

$$\det A^T = \det A\,.$$

Determinants have several other important properties, which we shall not discuss here. For example, it is possible to give a closed "general formula" for the determinant of an $n \times n$ matrix, generalizing the formulas for determinants of orders 2 and 3 given above in Examples 1.5 and 1.6. However, it is more important to develop practical methods of calculating determinants, which is done in the next subsection.

Proofs of Theorems 1.2, 1.3 and 1.4 follow. Notice that they are all based on the axioms (multilinearity and skew-symmetricity with respect to columns) by which we defined the determinant.

Proof of Theorem 1.2. Consider $\det(AB)$ as a function of the columns of B. Denote them $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Notice that the columns of the matrix AB are the column-vectors $A\mathbf{b}_1, \ldots, A\mathbf{b}_n$ (check the definition of the matrix product). It follows that if a column of B is multiplied by a number c, the corresponding column of AB will be multiplied by c. Similarly, if two columns of B are interchanged, then the corresponding columns of AB will be interchanged. Therefore, by the properties of the determinant applied to $\det(AB)$, it follows that $\det(AB)$ as a function of $\mathbf{b}_1, \ldots, \mathbf{b}_n$ possesses the properties (1)–(3) of Theorem 1.1. $\square$

Proof of Theorem 1.3. Notice that for a matrix A, the multiplication from the left by a matrix B acts as a transformation of rows of A. In particular, for

$$B = \begin{pmatrix} c & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

the map $A \mapsto BA$ is the multiplication of the first row of A by the number c. Similarly, by putting c on the kth place on the diagonal (and keeping other diagonal elements equal to 1 and all off-diagonal, to 0) we obtain the matrix such that $A \mapsto BA$ is the multiplication of the kth row by c. Notice that the determinant of such matrix B equals c, for any $k = 1, \ldots, n$ (as it obtained from E by multiplying the kth column by c). Hence, by Theorem 1.2, if a row of A is multiplied by c, the determinant of the resulting matrix will be $\det(BA) = \det B \det A = c \det A$. This implies linearity (see Remark 1.3). In the same way we can perform the interchange of two rows of A by multiplying A from the left by a certain matrix B. For example, $A \mapsto BA$ with

$$B = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

acts as the interchange of the first and the second rows of A. Here B is obtained from the identity matrix E by swapping the first and the second column. Hence $\det B = -1$. For given i and j, one can similarly construct B such that $A \mapsto BA$ is the interchange of the ith row and the jth row. By Theorem 1.2 we have $\det(BA) = \det B \det A = -\det A$ for the result of the interchange. That means that $\det A$ is also an alternating function of rows, and the theorem is proved. $\square$

Proof of Theorem 1.4. Consider $f(A) = \det A^T$ as a function of the columns of A. Since the columns of A are exactly the rows of A^T, and by Theorem 1.3,

$\det A^T$ is a multilinear alternating function of the rows of A^T, it follows that $f(A) = \det A^T$ is a multilinear alternating function of the columns of A. Notice also that $f(E) = \det E^T = 1$, since $E^T = E$. Hence $f(A)$ satisfies all the properties (1)–(4) of Theorem 1.1 and must coincide with $\det A$. $\square$

1.4 Calculations

Example 1.8. If a matrix A has a zero row or a zero column, then $\det A = 0$. Indeed, it follows from linearity. A linear function is zero for a zero argument: if $f(c\,\mathbf{a}) = c\,f(\mathbf{a})$, then $f(\mathbf{0}) = f(c\,\mathbf{0}) = c\,f(\mathbf{0})$ for any c, so $f(\mathbf{0}) = 0$.

Example 1.9. For a diagonal 2×2 matrix we have $\begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} = ad - 0 = ad$. In general, for a diagonal $n \times n$ matrix, its determinant is the product of the diagonal entries:

$$\begin{vmatrix} \lambda_1 & \dots & 0 \\ 0 & \dots & 0 \\ 0 & \dots & \lambda_n \end{vmatrix} = \lambda_1 \lambda_2 \dots \lambda_n$$

Indeed, λ_i can be taken out successively from each row (using linearity) until we obtain the identity matrix.

Example 1.10. If a matrix has all zeros below the diagonal, its determinant is again the product of the diagonal entries.

Consider, for example, the case $n = 3$. Suppose

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}.$$

We have

$$\det A = a_{33} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & 1 \end{vmatrix} = a_{33} \begin{vmatrix} a_{11} & a_{12} & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & 1 \end{vmatrix} = a_{33}a_{22} \begin{vmatrix} a_{11} & a_{12} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} =$$

$$a_{33}a_{22} \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = a_{33}a_{22}a_{11} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = a_{33}a_{22}a_{11}\,,$$

and the claim holds. Here we repeatedly used Remark 1.2 — more precisely, its analog for rows: adding to any row a multiple of another row will not change the determinant. In the same way we can treat the case of general n.

The same method of 'row operations' (i.e., simplifying the matrix by multiplying/dividing a row by a number and adding a multiple of one row to another) can be effectively applied for calculating an arbitrary determinant.

Example 1.11. Calculate $\begin{vmatrix} 1 & -2 & 0 \\ 2 & 0 & 4 \\ 3 & -5 & 5 \end{vmatrix}$. We have

$$\begin{vmatrix} 1 & -2 & 0 \\ 2 & 0 & 4 \\ 3 & -5 & 5 \end{vmatrix} = \begin{vmatrix} 1 & -2 & 0 \\ 0 & 4 & 4 \\ 0 & 1 & 5 \end{vmatrix} = 4 \begin{vmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 5 \end{vmatrix} = 4 \begin{vmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{vmatrix} = 4 \cdot 4 = 16\,.$$

There is another general method of calculating determinants of arbitrary order n, known as 'row expansion' (or, a variant called 'column expansion').

Let us first consider one example.

Example 1.12. Suppose that a matrix A in a certain row has all entries equal to zero except for one equal to 1. (A variant: the same for a column.) What can be said about its determinant? Consider a particular case of this situation. Let the first row of A has 1 as the first element and 0 at all other places:

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Then we can repeatedly apply row operations, subtracting the first row multiplied by a_{21}, a_{22}, etc., respectively, from the second, the third, etc., and the last row. We arrive at

$$\det A = \begin{vmatrix} 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

What is the value of the resulting determinant? Notice that is a function $f(\mathbf{r}_1, \dots, \mathbf{r}_{n-1})$ of the $n-1$ row-vectors $\mathbf{r}_1 = (a_{22}, \dots, a_{2n})$, $\dots$, $\mathbf{r}_{n-1} = (a_{n2}, \dots, a_{nn})$ belonging to $\mathbb{R}^{n-1}$. Clearly, from the properties of $\det A$ follows that this function is linear in each row-vector and alternating. Hence it is the $(n-1)\times(n-1)$ determinant $\begin{vmatrix} a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{n2} & \dots & a_{nn} \end{vmatrix}$ up to a factor equal to the value of $f(\mathbf{r}_1, \dots, \mathbf{r}_{n-1})$ at the identity matrix. We have

$$f(E) = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{vmatrix} = \det E = 1$$

(here at the LHS the identity matrix E is $(n-1) \times (n-1)$, while at the RHS the identity matrix is $n \times n$). Hence, finally,

$$\det A = \begin{vmatrix} 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

What will change if 1 appears as the second element in the first row rather than the first? Similarly to the above we will have

$$\begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ a_{21} & 0 & a_{23} & \dots & a_{2n} \\ a_{31} & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & 0 & a_{n3} & \dots & a_{nn} \end{vmatrix} = C \cdot \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix},$$

where C is the value of the original $n \times n$ determinant when the row-vectors $\mathbf{r}_1 = (a_{21}, a_{23}, \dots, a_{2n})$, $\dots$, $\mathbf{r}_{n-1} = (a_{n1}, a_{n3}, \dots, a_{nn})$ make the identity matrix $(n-1) \times (n-1)$. We have

$$C = \begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} = -1$$

because this is just the determinant of the $n \times n$ identity matrix with the first and second rows swapped (equivalently, the first and second column swapped). Hence

$$\begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix}.$$

One can in the same way see that for 1 at the kth position in the first row, the answer will include the factor $(-1)^{k-1}$:

$$\begin{vmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{21} & \dots & a_{2,k-1} & a_{2k} & a_{2,k+1} & \dots & a_{2n} \\ a_{31} & \dots & a_{3,k-1} & a_{3k} & a_{3,k+1} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,k-1} & a_{nk} & a_{n,k+1} & \dots & a_{nn} \end{vmatrix} = (-1)^{k-1} \begin{vmatrix} a_{21} & \dots & a_{2,k-1} & a_{2,k+1} & \dots & a_{2n} \\ a_{31} & \dots & a_{3,k-1} & a_{3,k+1} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,k-1} & a_{n,k+1} & \dots & a_{nn} \end{vmatrix}.$$

The main idea in this example is that, for matrices of a special appearance, the determinant of order n reduces to a determinant of order $n-1$. This can be used for a general practical algorithm of calculating determinants.

Theorem 1.5 ("Expansion in the first row"). *An arbitrary determinant of order n can be calculated via determinants of orders $n-1$ as follows:*

$$\det A = a_{11}M_{11} - a_{12}M_{12} + \dots + (-1)^{n-1}a_{1n}M_{1n}$$

where a_{11}, ..., a_{1n} are the entries of the first row and M_{11}, ..., M_{1n} are the determinants of the $(n-1)\times(n-1)$ matrices obtained from A by crossing out the first row and the first, the second, ..., and the last columns, respectively.

Proof. Consider the first row of A. It is the sum $a_{11}\mathbf{e}_1 + \dots + a_{1n}\mathbf{e}_n$ where $\mathbf{e}_k$, $k = 1, \dots, n$, is the row-vector with the single nonzero entry equal to 1 at the kth place. By the linearity of $\det A$, we have $\det A = a_{11} \det A_1 + \dots + a_{nn} \det A_n$ where A_k is the $n \times n$ matrix obtained from A by replacing its first row by $\mathbf{e}_k$. These are exactly the determinants calculated in Example 1.12, so $\det A_k = (-1)^{k-1}M_{1k}$, for all $k = 1, \dots, n$, which proves the theorem. $\square$

The determinants M_{11}, ..., M_{1n} (and similar ones) are called the minors of the matrix A. In general, a *minor* of A is the determinant of the square matrix obtained from A by crossing out some rows and columns (the same number of rows and columns). There are generalizations of the above expansion including other minors. In particular, there is an expansion in the second row (instead of the first row), and in any given row, and in the first column, as well as in any given column. The formulas are similar to the above, but the signs will depend both on a row and a column.

Example 1.13. Calculate the determinant of the 3×3 matrix $\begin{pmatrix} 2 & 1 & 1 \\ 0 & -3 & 5 \\ 1 & 0 & 3 \end{pmatrix}$

using the expansion in the first row. Solution: we have

$$\begin{vmatrix} 2 & 1 & 1 \\ 0 & -3 & 5 \\ 1 & 0 & 3 \end{vmatrix} = 2\begin{vmatrix} -3 & 5 \\ 0 & 3 \end{vmatrix} - \begin{vmatrix} 0 & 5 \\ 1 & 3 \end{vmatrix} + \begin{vmatrix} 0 & -3 \\ 1 & 0 \end{vmatrix} = 2(-3\cdot 3) - (-5) + 3 = -18+5+3 = -10\,.$$

Example 1.14. Calculate the determinant of the 4×4 matrix $\begin{pmatrix} 1 & 1 & 0 & 4 \\ -2 & 0 & 3 & 6 \\ 0 & 1 & 5 & -1 \\ 3 & -3 & 6 & 1 \end{pmatrix}$ using the expansion in the first row. Solution: calculate first the minors $M_{11}, \ldots, M_{14}$. We have

$$M_{11} = \begin{vmatrix} 0 & 3 & 6 \\ 1 & 5 & -1 \\ -3 & 6 & 1 \end{vmatrix} = -3\begin{vmatrix} 1 & -1 \\ -3 & 1 \end{vmatrix} + 6\begin{vmatrix} 1 & 5 \\ -3 & 6 \end{vmatrix} = -3\cdot 4 + 6\cdot 21 = 114$$

$$M_{12} = \begin{vmatrix} -2 & 3 & 6 \\ 0 & 5 & -1 \\ 3 & 6 & 1 \end{vmatrix} = -2\begin{vmatrix} 5 & -1 \\ 6 & 1 \end{vmatrix} - 3\begin{vmatrix} 0 & -1 \\ 3 & 1 \end{vmatrix} + 6\begin{vmatrix} 0 & 5 \\ 3 & 6 \end{vmatrix} = -2\cdot 11 - 3\cdot 3 + 6(-15) = -121$$

$$M_{13} = \begin{vmatrix} -2 & 0 & 6 \\ 0 & 1 & -1 \\ 3 & -3 & 1 \end{vmatrix} = -2\begin{vmatrix} 1 & -1 \\ -3 & 1 \end{vmatrix} + 6\begin{vmatrix} 0 & 1 \\ 3 & -3 \end{vmatrix} = -2(-2) + 6(-3) = -14$$

$$M_{14} = \begin{vmatrix} -2 & 0 & 3 \\ 0 & 1 & 5 \\ 3 & -3 & 6 \end{vmatrix} = -2\begin{vmatrix} 1 & 5 \\ -3 & 6 \end{vmatrix} + 3\begin{vmatrix} 0 & 1 \\ 3 & -3 \end{vmatrix} = -2\cdot 21 + 3(-3) = -51$$

Now we have

$$\det A = 1\cdot M_{11} - 1\cdot M_{12} + 0\cdot M_{13} - 4\cdot M_{14} = 114 - (-121) - 4(-51) = 439\,.$$

(We might have noticed earlier that it was not necessary to calculate M_{13} !)

1.5 Problems

Problem 1.1. Carry out the following matrix operations:

(a) AB and BA if $A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}$;

(b) $AB - BA$ where $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 1 & -2 \\ 3 & -2 & 4 \\ -3 & 5 & -1 \end{pmatrix}$.

(*Ans.:* **(a)** $\begin{pmatrix} 6 & 2 \\ 3 & 2 \end{pmatrix}$ and $\begin{pmatrix} 7 & 1 \\ 1 & 1 \end{pmatrix}$; **(b)** $\mathbf{0}$.)

Problem 1.2. Evaluate the determinants:

(a) $\begin{vmatrix} 5 & 2 \\ 7 & 3 \end{vmatrix}$, **(b)** $\begin{vmatrix} 3 & 2 \\ 8 & 5 \end{vmatrix}$, **(c)** $\begin{vmatrix} 2 & 1 & 3 \\ 5 & 3 & 2 \\ 1 & 4 & 3 \end{vmatrix}$, **(d)** $\begin{vmatrix} 4 & -3 & 5 \\ 3 & -2 & 8 \\ 1 & -7 & -5 \end{vmatrix}$, **(e)** $\begin{vmatrix} 5 & 1 & 2 & 7 \\ 3 & 0 & 0 & 2 \\ 1 & 3 & 4 & 5 \\ 2 & 0 & 0 & 3 \end{vmatrix}$.

For the third and fourth order determinants you should use row operations or the expansion in the first row.

(*Ans.:* **(a)** 1; **(b)** -1; **(c)** 40; **(d)** 100; **(e)** 10.)

Problem 1.3. Suppose $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$. Verify directly that $\det AB = \det A \cdot \det B$.

Problem 1.4. Consider a system of linear equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \, , \\ a_{21}x_1 + a_{22}x_2 &= b_2 \, . \end{aligned}$$

Solve it and show that the solution is given by the formulae

$$x_1 = \frac{1}{\Delta} \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} , \quad x_2 = \frac{1}{\Delta} \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

where $\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$. (For example, you can express x_1 from the first equation, substitute it into the second equation, solve the resulting equation for x_2 and substitute the answer into the expression for x_1; you may assume that you can divide by any expression whenever you need it.)

Problem 1.5. For vectors in $\mathbb{R}^n$ there is the notion of a 'scalar product'. For row-vectors $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$ their *scalar product* $(\mathbf{a}, \mathbf{b})$ is the number $\mathbf{a}\mathbf{b}^T = a_1b_1 + \dots + a_nb_n$. (For column-vectors the formula will be $(\mathbf{a}, \mathbf{b}) = \mathbf{a}^T\mathbf{b}$.) Vectors are said to be *orthogonal* or *perpendicular* if their scalar product vanishes. Now, in $\mathbb{R}^3$ there is another notion of a 'vector product': for $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ their *vector product* $\mathbf{a} \times \mathbf{b}$ is a vector defined as the symbolic determinant

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

where the first row consists of the vectors $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, $\mathbf{e}_3 = (0,0,1)$ and the determinant (whose value is a vector in $\mathbb{R}^3$) is understood via its expansion in the first row.

(a) Evaluate $\mathbf{e}_1 \times \mathbf{e}_2$, $\mathbf{e}_2 \times \mathbf{e}_3$, $\mathbf{e}_3 \times \mathbf{e}_1$.

(b) Show that for any vectors $\mathbf{a}$ and $\mathbf{b}$, the vector product $\mathbf{a} \times \mathbf{b}$ is perpendicular to each of $\mathbf{a}$ and $\mathbf{b}$. (*Hint:* check that for an arbitrary vector $\mathbf{c} \in \mathbb{R}^3$,

$$(\mathbf{c}, \mathbf{a} \times \mathbf{b}) = (\mathbf{b}, \mathbf{c} \times \mathbf{a}) = (\mathbf{a}, \mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

and use the properties of determinants.)

(b) Calculate $\mathbf{n} = \mathbf{a} \times \mathbf{b}$ for $\mathbf{a} = (0,1,-2)$ and $\mathbf{b} = (3,-4,7)$ and directly verify that $(\mathbf{n}, \mathbf{a}) = (\mathbf{n}, \mathbf{b}) = 0$.

(c) Think how the notion of the vector product can be extended to $\mathbb{R}^n$ for arbitrary n. (*Hint:* it cannot be a product of <u>two</u> vectors except for $n = 3$.)

Problem 1.6. Consider an arbitrary 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(a) Check that $f(\lambda) = \det(A - \lambda E)$, where λ is a parameter, equals

$$\lambda^2 - (a+d)\lambda + ad - bc.$$

(a) Show that, for an arbitrary matrix A, it satisfies the matrix identity

$$A^2 - (a+d)A + (ad - bc)E = \mathbf{0}\,.$$

(A similar statement holds for $n \times n$ matrices, for all n; it is called the *Hamilton–Cayley theorem.* It expresses a fundamental fact that a square matrix of dimension n always satisfies a polynomial equation of degree n with "universal" coefficients, i.e., which are expressions in the matrix entries of the same form for all matrices.)

§2 Areas and Volumes

The area of a two-dimensional object such as a region of the plane and the volume of a three-dimensional object such as a solid body in space, as well the length of an interval of the real line, are all particular cases of a very general notion of *measure*. General measure theory is a part of analysis. Here we shall focus on the geometrical side of the idea of measure and its relation with the algebraic notion of determinant.

2.1 Why the area of a parallelogram is represented by a determinant

Although from practice we know very well what is the area of simple geometrical figures such as, for example, a rectangle or a disk, it is not easy to give a rigorous general definition of area. However, some properties of area should be clear; consider subsets of the plane $\mathbb{R}^2$:

(1) Area is always non-negative: $\operatorname{area} S \geq 0$ for all subsets $S \subset \mathbb{R}^2$ such that it makes sense to speak about their area;

(2) Area is additive: $\operatorname{area}(S_1 \cup S_2) = \operatorname{area} S_1 + \operatorname{area} S_2$, if the intersection $S_1 \cap S_2$ is empty.

We have not specified an exact class of subsets of $\mathbb{R}^2$ for which area makes sense. We want, at least, that all polygons such as triangles, rectangles, etc., belong to this class. For them we agree that their boundaries, consisting of several straight line segments, must have area zero. Thus it makes no difference whether we consider, say, a closed rectangle (including the boundary points) or an open one (without the boundary); for both the area will be the same. Also, it follows that if two polygons intersect by a segment or a finite number of segments only, then the area of their union will be the sum of areas. We follow the natural idea that area is a "two-dimensional measure", so every one-dimensional object, such as a segment should have this (two-dimensional) measure zero.

Restriction by polygons in the plane, of course, is too strong; it excludes familiar examples such as disks and more general domains with "curvilinear" boundaries. It also excludes surfaces such as, e.g., regions of a sphere in $\mathbb{R}^3$. For them, too, we want to have the notion of area. This will become possible after we master the basic case of polygons in $\mathbb{R}^2$.

Properties (1) and (2) above are very general. They are applicable to arbitrary abstract sets and as such they are turned into axioms in abstract measure theory (where the condition of additivity is typically extended to certain infinite unions). We want to add to them some properties peculiar for the plane $\mathbb{R}^2$.

A *translation* of the space $\mathbb{R}^n$ is the map $T_{\mathbf{a}}\colon \mathbb{R}^n \to \mathbb{R}^n$ that takes every point $\mathbf{x} \in \mathbb{R}^n$ to $\mathbf{x} + \mathbf{a}$, where $\mathbf{a} \in \mathbb{R}^n$ is a fixed vector. For a subset $S \subset \mathbb{R}^n$, a translation $T_{\mathbf{a}}$ "shifts" all points of S along $\mathbf{a}$, i.e., each point $\mathbf{x} \in S$ is mapped to $\mathbf{x} + \mathbf{a}$, and S moves "rigidly" to its new location in $\mathbb{R}^n$.

Example 2.1. A disk of radius R and center $O = (0,0)$ in $\mathbb{R}^2$ under the translation $T_{\mathbf{a}}$ where $\mathbf{a} = (a_1, a_2)$ is mapped to the disk of the same radius R with center (a_1, a_2). The area of a disk, clearly, should not depend on a position of the center.

The following natural properties hold for areas in $\mathbb{R}^2$:

(3) Area is invariant under translations: $\operatorname{area} T_{\mathbf{a}}(S) = \operatorname{area} S$, for all vectors $\mathbf{a} \in \mathbb{R}^2$.

(4) For a one-dimensional object, such as a segment, the area should vanish.

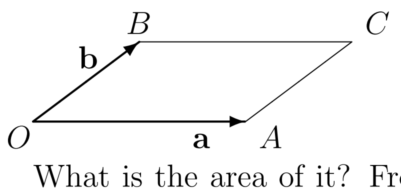
We do not define precisely what is a 'one-dimensional object'. However, examples such as segments and their unions will be sufficient for our purposes.

The plan now is as follows. Using conditions (1)–(4) we shall establish a deep link between the notion of area and the theory of determinants. To this end, we consider the area of a simple polygon, a parallelogram. Later our considerations will be generalized to $\mathbb{R}^n$.

Let $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$ be vectors in $\mathbb{R}^2$. The *parallelogram* on $\mathbf{a}, \mathbf{b}$ with basepoint $O \in \mathbb{R}^2$ is the set of points of the form

$$\mathbf{x} = O + t\mathbf{a} + s\mathbf{b} \quad \text{where } 0 \leq t, s \leq 1\,.$$

One can easily see that it is the plane region bounded by the two pairs of parallel straight line segments: OA, BC and OB, AC where $C = O + \mathbf{a} + \mathbf{b}$.



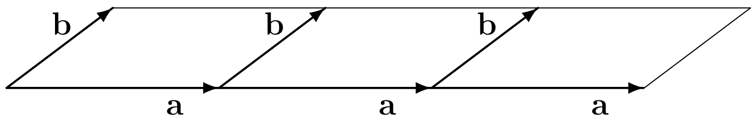
What is the area of it? From property (3) it follows that the area does not depend on the location of our parallelogram in $\mathbb{R}^2$: by a translation the basepoint O can be made an arbitrary point of the plane without changing the area. Let us assume that $O = \mathbf{0}$ is the point $(0,0)$. Denote the parallelogram by $\Pi(\mathbf{a}, \mathbf{b})$. Then $\operatorname{area} \Pi(\mathbf{a}, \mathbf{b})$ is a function of vectors $\mathbf{a}, \mathbf{b}$.

Proposition 2.1. *The function* $\operatorname{area}\Pi(\mathbf{a},\mathbf{b})$ *has the following properties:*

(1) $\operatorname{area}\Pi(n\mathbf{a},\mathbf{b}) = \operatorname{area}\Pi(\mathbf{a},n\mathbf{b}) = |n|\cdot\operatorname{area}\Pi(\mathbf{a},\mathbf{b})$ *for any* $n\in\mathbb{Z}$*;*

(2) $\operatorname{area}\Pi(\mathbf{a},\mathbf{b}+k\mathbf{a}) = \operatorname{area}\Pi(\mathbf{a}+k\mathbf{b},\mathbf{b}) = \operatorname{area}\Pi(\mathbf{a},\mathbf{b})$ *for any* $k\in\mathbb{R}$.

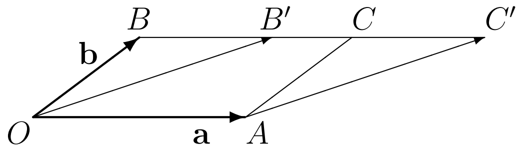
Proof. Suppose we replace $\mathbf{a}$ by $n\mathbf{a}$ for a positive integer n. Then $\Pi(n\mathbf{a},\mathbf{b})$ is the union of n copies of the parallelogram $\Pi(\mathbf{a},\mathbf{b})$:



From the additivity of area it follows that

$$\operatorname{area}\Pi(n\mathbf{a},\mathbf{b}) = n\,\Pi(\mathbf{a},\mathbf{b}).$$

The same is true if we replace $\mathbf{b}$ by $n\mathbf{b}$, for positive n. For $n=0$, $\Pi(\mathbf{0},\mathbf{b})$ or $\Pi(\mathbf{a},\mathbf{0})$ are just segments, therefore have zero area, by (4). Notice also that $\Pi(-\mathbf{a},\mathbf{b})$ and $\Pi(\mathbf{a},\mathbf{b})$ differ by a shift, so have the same area. Hence, $\operatorname{area}\Pi(n\mathbf{a},\mathbf{b}) = \operatorname{area}\Pi(\mathbf{a},n\mathbf{b}) = |n|\cdot\operatorname{area}\Pi(\mathbf{a},\mathbf{b})$ holds in general. To prove the second relation, we again use the additivity of area: it is clear that to obtain $\Pi(\mathbf{a},\mathbf{b}+k\mathbf{a})$, one has to cut from $\Pi(\mathbf{a},\mathbf{b})$ the triangle OBB', shift it by the vector $\mathbf{a}$ and attach it back as the triangle ACC':



Clearly, due to additivity and invariance under translations, the area will not change, and we arrive at $\operatorname{area}\Pi(\mathbf{a},\mathbf{b}+k\mathbf{a}) = \operatorname{area}\Pi(\mathbf{a},\mathbf{b})$ as claimed. $\square$

In fact, the first assertion in Proposition 2.1 is valid in a stronger form.

Proposition 2.2. *The area of a parallelogram satisfies*

$$\operatorname{area}\Pi(k\mathbf{a},\mathbf{b}) = \operatorname{area}\Pi(\mathbf{a},k\mathbf{b}) = |k|\cdot\operatorname{area}\Pi(\mathbf{a},\mathbf{b})$$

for any real number $k\in\mathbb{R}$.

Proof. For rational numbers $r=n/m$, where $n,m\in\mathbb{Z}$ this directly follows from part (1) of Proposition 2.1. Now, every real number can be approximated by rational numbers as closely as we wish, and the statement follows by continuity. (We assume that the area of $\Pi(\mathbf{a},\mathbf{b})$ continuously depend on $\mathbf{a}$, $\mathbf{b}$, which is natural to expect.) $\square$

Theorem 2.1. *Suppose* $\mathbf{a} = (a_1, a_2)$*,* $\mathbf{b} = (b_1, b_2)$*. Then*

$$\operatorname{area}\Pi(\mathbf{a},\mathbf{b}) = C \cdot |\Delta|\,, \tag{1}$$

with some constant C*, where*

$$\Delta = \det(\mathbf{a},\mathbf{b}) = \det\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}\,. \tag{2}$$

Proof. Indeed, by Propositions 2.1 and 2.2, $\operatorname{area}\Pi(\mathbf{a},\mathbf{b})$ satisfies almost the same properties as the determinant $\det(\mathbf{a},\mathbf{b})$. They can be used to calculate $\operatorname{area}\Pi(\mathbf{a},\mathbf{b})$; it is similar to using row operations for calculating determinants. We have

$$\operatorname{area}\Pi(\mathbf{a},\mathbf{b}) = \operatorname{area}\Pi\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} = \operatorname{area}\Pi\begin{pmatrix} a_1 & a_2 \\ 0 & b_2 - \frac{b_1}{a_1}\,a_2 \end{pmatrix} =$$
$$\operatorname{area}\Pi\begin{pmatrix} a_1 & 0 \\ b_1 & b_2 - \frac{b_1}{a_1}\,a_2 \end{pmatrix} = |a_1|\,|b_2 - \frac{b_1}{a_1}\,a_2|\ \operatorname{area}\Pi\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =$$
$$|a_1b_2 - b_1a_2| \cdot \operatorname{area}\Pi(\mathbf{e}_1,\mathbf{e}_2)\,,$$

where $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$, as desired. □

We see that the constant C in the theorem is the just area of the parallelogram built on the 'basis vectors' $\mathbf{e}_1$, $\mathbf{e}_2$. Knowing this number allows to calculate the area of any parallelogram by using formulas (1), (2). This number can be set arbitrarily, which is equivalent to a choice of the *unit of area.*

The meaning of Theorem 2.1 is very deep: it tells that the natural properties of areas, such as additivity and invariance under translations, imply that the area of a parallelogram $\Pi(\mathbf{a},\mathbf{b})$ has characteristic properties the same as those of (the absolute value of) the determinant $\det(\mathbf{a},\mathbf{b})$. Therefore they should essentially coincide.

In the following examples let us assume that a unit of area is chosen so that $\operatorname{area}\Pi(\mathbf{e}_1,\mathbf{e}_2) = 1$.

Example 2.2. Find the area of the parallelogram built on vectors $\mathbf{a} = (-2,5)$ and $\mathbf{b} = (1,1)$. Solution: we have

$$\Delta = \begin{vmatrix} -2 & 5 \\ 1 & 1 \end{vmatrix} = -7\,.$$

Hence $\operatorname{area}\Pi(\mathbf{a},\mathbf{b}) = |-7| = 7$.

Example 2.3. Find the area of the triangle ABC if $A=(3,2)$, $B=(4,2)$, $C=(1,0)$. Solution: it is half of the area of the parallelogram built on $\overrightarrow{CA}=A-C=(2,2)$, $\overrightarrow{CB}=B-C=(3,2)$. Hence

$$\operatorname{area}(ABC)=\frac{1}{2}\begin{vmatrix}3 & 2\\ 2 & 2\end{vmatrix}=1\,.$$

How we can make sense of the determinant $\det(\mathbf{a},\mathbf{b})$ as such, not its absolute value? It corresponds to the notion of *signed*, or *oriented area.* Denote it $\operatorname{Area}\Pi(\mathbf{a},\mathbf{b})$ with capital "a". By definition, *signed area* satisfies

$$\begin{gathered}\operatorname{Area}\Pi(k\mathbf{a},\mathbf{b})=\operatorname{Area}\Pi(\mathbf{a},k\mathbf{b})=k\ \operatorname{Area}\Pi(\mathbf{a},\mathbf{b})\\ \operatorname{Area}\Pi(\mathbf{a},\mathbf{b}+k\mathbf{a})=\operatorname{Area}\Pi(\mathbf{a}+k\mathbf{b},\mathbf{b})=\operatorname{Area}\Pi(\mathbf{a},\mathbf{b})\end{gathered}$$

for all $k\in\mathbb{R}$. We have

$$\operatorname{Area}(\mathbf{a},\mathbf{b})=\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}\tag{3}$$

if we assume that the signed area of $\Pi(\mathbf{e}_1,\mathbf{e}_2)$ equals 1.

2.2 Volumes and determinants

All the above results can be generalized to higher dimensions. Consider vectors $\mathbf{a}_1,\dots,\mathbf{a}_n$ in $\mathbb{R}^n$. The *parallelipiped* on $\mathbf{a}_1,\dots,\mathbf{a}_n$ with basepoint $O\in\mathbb{R}^n$ is the set of points

$$\mathbf{x}=O+t_1\mathbf{a}_1+\ldots+t_n\mathbf{a}_n\quad\text{where } 0\leq t_1\leq 1 \text{ for all } i=1,\dots,n\,.$$

Denote it $\Pi(\mathbf{a}_1,\dots,\mathbf{a}_n)$. In the sequel nothing depends on the basepoint, so we shall suppress any mentioning of it. Instead of deducing a formula for the volume of a parallelipiped similar to (1) from general properties of volumes such as additivity, as we did above for area, it is convenient to set by definition

$$\operatorname{Vol}\Pi(\mathbf{a}_1,\dots,\mathbf{a}_n)=\begin{vmatrix}a_{11} & \dots & a_{1n}\\ \dots & \dots & \dots\\ a_{n1} & \dots & a_{nn}\end{vmatrix}\tag{4}$$

if $\mathbf{a}_1=(a_{11},\dots,a_{1n})$, ..., $\mathbf{a}_n=(a_{n1},\dots,a_{nn})$. This is the *oriented* or *signed volume.* The 'usual' volume vol is the absolute value of Vol. Note that this definition implies that the "unit of volume" is such that the oriented volume of $\Pi(\mathbf{e}_1,\dots,\mathbf{e}_n)$ is set to 1.

Example 2.4. Find the oriented volume of the parallelipiped built on $\mathbf{a}_1 = (2, 1, 0)$, $\mathbf{a}_2 = (0, 3, 11)$ and $\mathbf{a}_3 = (1, 2, 7)$ in $\mathbb{R}^3$. Solution:

$$\operatorname{Vol}\Pi(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3) = \begin{vmatrix} 2 & 1 & 0 \\ 0 & 3 & 11 \\ 1 & 2 & 7 \end{vmatrix} = 2 \cdot 1 - 1 \cdot 11 = -9\,.$$

(The volume 'without sign' equals 9.)

2.3 Areas and volumes in Euclidean space

In the above analysis of areas and volumes, there was an arbitrary choice of "unit of volume", i.e., a choice of the constant before the determinant in formulas such as (1), (3), (4). However, there is a natural choice linking measurement of volumes and areas with measurement of lengths and angles. It is given by the rule: *the volume of a unit cube equals* 1. A unit cube in $\mathbb{R}^n$ is a parallelipiped has unit edges that are perpendicular to each other. In $\mathbb{R}^2$ it is a unit square.

Recall that on $\mathbb{R}^n$ one can define the *scalar product* of vectors by the formula

$$(\mathbf{a}, \mathbf{b}) = a_1 b_1 + \ldots + a_n b_n$$

(see Problem 1.5). It follows that the 'standard basis vectors' $\mathbf{e}_1, \ldots, \mathbf{e}_n$ satisfy $(\mathbf{e}_i, \mathbf{e}_j) = 0$ if $i \neq j$ and $(\mathbf{e}_i, \mathbf{e}_i) = 1$. The length of a vector is defined as

$$|\mathbf{a}| = \sqrt{(\mathbf{a}, \mathbf{a})}$$

and the angle between two vectors is defined by the equality

$$(\mathbf{a}, \mathbf{b}) = |\mathbf{a}|\,|\mathbf{b}|\,\cos\alpha$$

from where we can find $\cos\alpha$ if $(\mathbf{a}, \mathbf{b})$, $(\mathbf{a}, \mathbf{a})$, $(\mathbf{b}, \mathbf{b})$ are known. Hence the relations for $\mathbf{e}_i$ mean that they all have unit length and are mutually perpendicular. Therefore $\Pi(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ is a unit cube. We set

$$\operatorname{Vol}\Pi(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1\,.$$

Then for any n vectors $\mathbf{a}_i = (a_{i1}, \ldots, a_{in})$, where $i = 1, \ldots, n$, we obtain

$$\operatorname{Vol}\Pi(\mathbf{a}_1, \ldots, \mathbf{a}_n) = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}\,. \tag{5}$$

Remark 2.1. The space $\mathbb{R}^n$ together with the scalar product is called the n-dimensional *Euclidean space*. The adjective 'Euclidean' points to the fact that the scalar product allows to define lengths and angles, i.e., the main notions of the classical Euclidean geometry.

Remark 2.2. The scalar product is alternatively denoted $\mathbf{a} \cdot \mathbf{b}$ and hence is often referred to as the 'dot product'.

It seems that the unit cube $\Pi(\mathbf{e}_1, \dots, \mathbf{e}_n)$ plays a distinguished role. Later we shall show that **any** unit cube in $\mathbb{R}^n$ has unit volume. Consider an example (for $n = 2$ we continue to use Area instead of Vol).

Example 2.5. Let $\mathbf{g}_1 = (\cos\alpha, \sin\alpha)$, $\mathbf{g}_2 = (-\sin\alpha, \cos\alpha)$ in $\mathbb{R}^2$. We can immediately see that $|\mathbf{g}_1| = |\mathbf{g}_2| = 1$ and $\mathbf{g}_1 \cdot \mathbf{g}_2 = 0$, so $\Pi(\mathbf{g}_1, \mathbf{g}_2)$ is a unit square. We have

$$\operatorname{Area}\Pi(\mathbf{g}_1, \mathbf{g}_2) = \begin{vmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{vmatrix} = \cos^2\alpha + \sin^2\alpha = 1.$$

There is a way of expressing the volume of a parallelipiped entirely in terms of 'intrinsic' geometric information: lengths of vectors and angles between them, rather than their coordinates as in the previous formulas. Consider the matrix

$$G(\mathbf{a}_1, \dots, \mathbf{a}_k) = \begin{pmatrix} (\mathbf{a}_1, \mathbf{a}_1) & \dots & (\mathbf{a}_1, \mathbf{a}_k) \\ \dots & \dots & \dots \\ (\mathbf{a}_k, \mathbf{a}_1) & \dots & (\mathbf{a}_k, \mathbf{a}_k) \end{pmatrix}. \tag{6}$$

Here $k \leq n$ may be less than n.

Definition 2.1. The matrix $G(\mathbf{a}_1, \dots, \mathbf{a}_n)$ is called the *Gram matrix* of the system of vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ and its determinant, the *Gram determinant.*

Theorem 2.2. *The Gram determinant of* $\mathbf{a}_1, \dots, \mathbf{a}_n$ *is the square of the volume of the parallelipiped* $\Pi(\mathbf{a}_1, \dots, \mathbf{a}_n)$.

Proof. Indeed, consider the $n \times n$ matrix A with rows $\mathbf{a}_i$. Consider

$$AA^T = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \dots & \dots & \dots \\ a_{1n} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1 \\ \dots \\ \mathbf{a}_n \end{pmatrix} \begin{pmatrix} \mathbf{a}_1^T & \dots & \mathbf{a}_n^T \end{pmatrix} =$$

$$\begin{pmatrix} \mathbf{a}_1\mathbf{a}_1^T & \dots & \mathbf{a}_1\mathbf{a}_n^T \\ \dots & \dots & \dots \\ \mathbf{a}_n\mathbf{a}_1^T & \dots & \mathbf{a}_n\mathbf{a}_n^T \end{pmatrix} = G(\mathbf{a}_1, \dots, \mathbf{a}_n).$$

Hence

$$\det G(\mathbf{a}_1, \dots, \mathbf{a}_n) = \det(AA^T) = \det A \, \det A^T = (\det A)^2 = \Big(\operatorname{vol} \Pi(\mathbf{a}_1, \dots, \mathbf{a}_n)\Big)^2.$$

$\square$

Corollary 2.1. *If $\mathbf{g}_i$, $i = 1, \dots, n$, are arbitrary mutually orthogonal unit vectors (i.e., $\Pi(\mathbf{g}_1, \dots, \mathbf{g}_n)$ is a unit cube), then* $\operatorname{vol} \Pi(\mathbf{g}_1, \dots, \mathbf{g}_n) = 1$.

Proof. Indeed, $\Big(\operatorname{vol} \Pi(\mathbf{g}_1, \dots, \mathbf{g}_n)\Big)^2 = \det G(\mathbf{g}_1, \dots, \mathbf{g}_n)) = \det E = 1$. $\square$

One of the advantages of expressing volumes via the Gram determinants is that it allows to consider easily parallelipipeds in $\mathbb{R}^n$ of dimensions less than n. More precisely, if we are given k vectors $\mathbf{a}_1, \dots, \mathbf{a}_k$, then we can consider a k-dimensional parallelipiped $\Pi(\mathbf{a}_1, \dots, \mathbf{a}_k)$ contained in the k-dimensional plane spanned by $\mathbf{a}_1, \dots, \mathbf{a}_k$. The formula

$$\Big(\operatorname{vol} \Pi(\mathbf{a}_1, \dots, \mathbf{a}_k)\Big)^2 = \det G(\mathbf{a}_1, \dots, \mathbf{a}_k)$$

is applicable, with the scalar products calculated in the ambient space $\mathbb{R}^n$.

Example 2.6. Find the area of the parallelogram built on $\mathbf{a} = (1, -1, 2)$ and $\mathbf{b} = (2, 0, 3)$. Solution: the Gram determinant is

$$\begin{vmatrix} 6 & 8 \\ 8 & 13 \end{vmatrix} = 14\,.$$

Hence the area is $\sqrt{14}$.

Remark 2.3. Differently from the case of an n-dimensional parallelipiped in $\mathbb{R}^n$, where one can consider 'signed' volume given by formula (4), the volume of a k-dimensional parallelipiped, for $k < n$, given by the Gram determinant is not signed. (For $k = n$, if we use the Gram determinant, to obtain a signed volume, we have to pick a sign independently.)

For $n = 2$, it is obvious that the area of a parallelogram $\Pi(\mathbf{a}, \mathbf{b})$ is the product of the 'base' $|\mathbf{a}|$ and 'height' h, which is the length of a vector $\mathbf{h} = \mathbf{b} + k\mathbf{a}$ such that it is perpendicular to $\mathbf{a}$ (the parameter k is defined by the condition $(\mathbf{a},\mathbf{h}) = 0$). It is a very basic fact following from the same sort of ideas that lead us to discovering the relation between areas and determinants. For $n = 3$, the similar fact about the volume of a 3-dimensional parallelipiped is also very familiar. This relation between the n-dimensional volume in the Euclidean space $\mathbb{R}^n$ and the $(n-1)$-dimensional volume in the Euclidean space $\mathbb{R}^{n-1}$ holds for any n. The easiest way to prove it is by using the Gram determinants.

Temporarily introduce notation vol_n and vol_{n-1} for distinguishing between volumes in $\mathbb{R}^n$ and $\mathbb{R}^{n-1}$. Let us consider the $(n-1)$-dimensional parallelipiped $\Pi(\mathbf{a}_1,\dots,\mathbf{a}_{n-1})$ as a *base* of a parallelipiped $\Pi(\mathbf{a}_1,\dots,\mathbf{a}_n)$. Then the *height* of $\Pi(\mathbf{a}_1,\dots,\mathbf{a}_n)$ is the length of a (unique) vector $\mathbf{h}$ defined by the condition $\mathbf{h} = \mathbf{a}_n + \mathbf{c}$ where $\mathbf{c}$ is a linear combination of $\mathbf{a}_1,\dots,\mathbf{a}_{n-1}$ and $\mathbf{h}$ is perpendicular to the plane of $\mathbf{a}_1,\dots,\mathbf{a}_{n-1}$ (i.e., to each of the vectors $\mathbf{a}_1,\dots,\mathbf{a}_{n-1}$).

Theorem 2.3. *The following formula holds:*

$$\mathrm{vol}_n\,\Pi(\mathbf{a}_1,\dots,\mathbf{a}_n) = \mathrm{vol}_{n-1}\,\Pi(\mathbf{a}_1,\dots,\mathbf{a}_{n-1})\cdot h\,,$$

where $h = |\mathbf{h}|$ *is the height of* $\Pi(\mathbf{a}_1,\dots,\mathbf{a}_n)$.

Proof. Using the properties of volumes we immediately conclude that

$$\mathrm{vol}_n\,\Pi(\mathbf{a}_1,\dots,\mathbf{a}_n) = \mathrm{vol}_n\,\Pi(\mathbf{a}_1,\dots,\mathbf{a}_{n-1},\mathbf{h}).$$

Now we can apply the Gram determinant:

$$\det G(\mathbf{a}_1,\dots,\mathbf{a}_{n-1},\mathbf{h}) = \begin{vmatrix} (\mathbf{a}_1,\mathbf{a}_1) & \dots & (\mathbf{a}_1,\mathbf{a}_{n-1}) & (\mathbf{a}_1,\mathbf{h}) \\ \dots & \dots & \dots & \dots \\ (\mathbf{a}_{n-1},\mathbf{a}_1) & \dots & (\mathbf{a}_{n-1},\mathbf{a}_{n-1}) & (\mathbf{a}_{n-1},\mathbf{h}) \\ (\mathbf{h},\mathbf{a}_1) & \dots & (\mathbf{h},\mathbf{a}_{n-1}) & (\mathbf{h},\mathbf{h}) \end{vmatrix} =$$

$$\begin{vmatrix} (\mathbf{a}_1,\mathbf{a}_1) & \dots & (\mathbf{a}_1,\mathbf{a}_{n-1}) & 0 \\ \dots & \dots & \dots & \dots \\ (\mathbf{a}_{n-1},\mathbf{a}_1) & \dots & (\mathbf{a}_{n-1},\mathbf{a}_{n-1}) & 0 \\ 0 & \dots & 0 & (\mathbf{h},\mathbf{h}) \end{vmatrix} = \det G(\mathbf{a}_1,\dots,\mathbf{a}_{n-1})\cdot|\mathbf{h}|^2\,.$$

and it remains to extract the square root. □

2.4 Examples and applications

Consider two applications of the methods introduced above.

2.4.1 The distance between a point and a plane

Consider a plane L in $\mathbb{R}^3$ and a point $\mathbf{x}$ not belonging to L. What is the distance between $\mathbf{x}$ and L? It is natural to define it as the minimum of the distances between $\mathbf{x}$ and points of the plane L. A practical calculation of it can be nicely related with areas and volumes. Indeed, let $\mathbf{y} \in L$ is an arbitrary point of the plane; the distance between $\mathbf{x}$ and $\mathbf{y}$ is $|\mathbf{x}-\mathbf{y}|$. We can write $\mathbf{x}-\mathbf{y} = \mathbf{a}_{\parallel} + \mathbf{a}_{\perp}$ where the vector $\mathbf{a}_{\parallel}$ is parallel to L and $\mathbf{a}_{\perp}$ is perpendicular to it. Hence $|\mathbf{x}-\mathbf{y}|^2 = |\mathbf{a}_{\parallel}|^2 + |\mathbf{a}_{\perp}|^2 \geq |\mathbf{a}_{\perp}|^2$. The part $\mathbf{a}_{\parallel}$ can vary by adding vectors parallel to L, while $\mathbf{a}_{\perp}$ is unique as long as $\mathbf{x}$ is given. It is clear now that the shortest length so obtained is for $\mathbf{x}-\mathbf{y} = \mathbf{a}_{\perp}$, i.e., when $\mathbf{y}$ is the end of the perpendicular dropped on L from the point $\mathbf{x}$. (Draw a picture.) Let O be some fixed point of the plane L and let vectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$ 'span' L so that an arbitrary point of L has the appearance $O + c_1\mathbf{a}_1 + \ldots + c_k\mathbf{a}_k$. Then we can consider a k-dimensional parallelipiped $\Pi(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ and a $(k+1)$-dimensional parallelipiped $\Pi(\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{x}-O)$, both with basepoint O. Clearly, the desired distance is the height of $\Pi(\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{x}-O)$.

Corollary 2.2. *The distance between a point* $\mathbf{x}$ *and a plane* L *through a point* O *in the direction of vectors* $\mathbf{a}_1, \ldots, \mathbf{a}_k$ *is given by the formula*

$$\frac{\operatorname{vol}\Pi(\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{x}-O)}{\operatorname{vol}\Pi(\mathbf{a}_1, \ldots, \mathbf{a}_k)} = \frac{\sqrt{\det G(\mathbf{a}_1, \ldots, \mathbf{a}_k, \mathbf{x}-O)}}{\sqrt{\det G(\mathbf{a}_1, \ldots, \mathbf{a}_k)}}$$

Example 2.7. Given a plane through the points $A = (1,0,0)$, $B = (0,1,0)$, $C = (0,0,1)$. Find the distance between it and the point $D = (3,2,-1)$. Solution: denote the desired distance by h. Consider the vectors $\mathbf{a} = \overrightarrow{CA} = (1,0,-1)$, $\mathbf{b} = \overrightarrow{CB} = (0,1,-1)$ and $\mathbf{d} = \overrightarrow{CD} = (3,2,-2)$. We have

$$h = \frac{\operatorname{vol}\Pi(\mathbf{a},\mathbf{b},\mathbf{d})}{\operatorname{area}\Pi(\mathbf{a},\mathbf{b})}\,.$$

The numerator is easier to find by using formula (4). We have

$$\begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 3 & 2 & -2 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 2 & -2 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 3 & 2 \end{vmatrix} = 3$$

hence $\operatorname{vol}\Pi(\mathbf{a},\mathbf{b},\mathbf{d}) = 3$. For the denominator we calculate the Gram determinant:

$$\det G(\mathbf{a},\mathbf{b}) = \begin{vmatrix} (\mathbf{a},\mathbf{a}) & (\mathbf{a},\mathbf{b}) \\ (\mathbf{b},\mathbf{a}) & (\mathbf{b},\mathbf{b}) \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 1\,;$$

hence $\operatorname{area}\Pi(\mathbf{a},\mathbf{b}) = \sqrt{\det G(\mathbf{a},\mathbf{b})} = 1$. Finally, $h = 3$.

2.4.2 Area of a piece of the plane or a surface

Knowing how to find the area of a parallelogram allows to find the areas of more general objects. The same is true for volumes in higher dimensions.

Consider first the problem of calculating the area of a domain in $\mathbb{R}^2$ using so-called "curvilinear coordinates". By a system of *curvilinear coordinates* in a part of the plane $\mathbb{R}^2$ is understood a way of specifying points $\mathbf{x} = (x, y)$ by suitable parameters u, v instead of their 'standard' coordinates x, y.

Example 2.8. Polar coordinates r, θ where $x = r\cos\theta$, $y = r\sin\theta$.

Let us choose such a system u, v. Curves $v = \text{const}$ and $u = \text{const}$ are called *coordinate lines*. For example, for standard coordinates x, y they are horizontal and vertical lines, respectively. For polar coordinates r, θ they are rays emanating from the origin $O = (0, 0)$ and circles with center O.

Consider a region of the plane bounded by the coordinate lines $v = v_0$, $v = v_0 + \Delta v$, $u = u_0$ and $u = u_0 + \Delta u$. For small increments $\Delta u, \Delta v$ it is a 'curvilinear quadrangle' with the vertices A where $u = u_0, v = v_0$, B where $u = u_0 + \Delta u, v = v_0$, C where $u = u_0 + \Delta u, v = v_0 + \Delta v$, and D where $u = u_0, v = v_0 + \Delta v$. Denote its area by ΔS. For small $\Delta u, \Delta v$, this curvilinear quadrangle $ABCD$ is very close to the parallelogram built on the vectors

$$\mathbf{a} = \mathbf{e}_u\,\Delta u, \quad \mathbf{b} = \mathbf{e}_v\,\Delta v$$

where $\mathbf{e}_u = \frac{\partial\mathbf{x}}{\partial u}(u_0, v_0)$, $\mathbf{e}_v = \frac{\partial\mathbf{x}}{\partial v}(u_0, v_0)$, with basepoint A. (Draw a picture!) Hence the area ΔS can be approximated by

$$\operatorname{area}\Pi(\mathbf{e}_u\,\Delta u, \mathbf{e}_v\,\Delta v) = \operatorname{area}\Pi(\mathbf{e}_u, \mathbf{e}_v)\Delta u\Delta v = \sqrt{\det G(\mathbf{e}_u, \mathbf{e}_v)}\,\Delta u\Delta v$$

Denote for brevity $\det G(\mathbf{e}_u, \mathbf{e}_v)$ by g. It is a function of a point in the plane.

Example 2.9. (Continuation of Example 2.8.) We have, for polar coordinates,

$$\mathbf{e}_r = \frac{\partial\mathbf{x}}{\partial r} = (\cos\theta, \sin\theta), \quad \mathbf{e}_\theta = \frac{\partial\mathbf{x}}{\partial\theta} = (-r\sin\theta, r\cos\theta)\,.$$

Hence

$$G(\mathbf{e}_r, \mathbf{e}_\theta) = \begin{pmatrix} \mathbf{e}_r\cdot\mathbf{e}_r & \mathbf{e}_r\cdot\mathbf{e}_\theta \\ \mathbf{e}_\theta\cdot\mathbf{e}_r & \mathbf{e}_\theta\cdot\mathbf{e}_\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

and $g = \det G(\mathbf{e}_r, \mathbf{e}_\theta) = r^2$.

How can we use this? Suppose we want to calculate the area of a certain domain $D \subset \mathbb{R}^2$. Choose some system of curvilinear coordinates in which D can be conveniently described. Consider a partition of D by the coordinate lines $v = v_k$, $u = u_l$, $k, l = 1, \ldots, N$, where $u_{l+1} - u_l = \Delta u$, $v_{k+1} - v_k = \Delta v$. Then the area of D is approximated by the sum of the areas of the curvilinear quadrangles ΔS as above, where u varies between u_l and u_{l+1}, and v, between v_k and v_{k+1}. We can also approximate each of ΔS by the area of the parallelogram $\Pi(\mathbf{e}_u \Delta u, \mathbf{e}_v \Delta v)$. The basepoints are on the grid, and the vectors $\mathbf{e}_u, \mathbf{e}_v$ are calculated at the corresponding points of the grid. Hence

$$\operatorname{area} D = \lim \sum_{k,l} \Delta S = \lim \sum_{k,l} \operatorname{area} \Pi(\mathbf{e}_u \Delta u, \mathbf{e}_v \Delta v) = \lim \sum_{k,l} \sqrt{g} \Delta u \Delta v$$

where $g = \det G(\mathbf{e}_u, \mathbf{e}_v)$. It is just the integral sum for a double integral, and we arrive at the following statement.

Proposition 2.3. *For the area of a domain $D \subset \mathbb{R}^2$ we have*

$$\operatorname{area} D = \iint_D dS \quad \textit{where } dS = \sqrt{g} du dv\,.$$

The expression dS under the integral sign is called the *element of area.*

Example 2.10. The element of area in the standard coordinates x, y and polar coordinates r, θ will be

$$dS = dxdy = rdrd\theta$$

using the result of Example 2.9.

Example 2.11. Find the area of a disk D_R of radius R using the above formulas. Solution: let the center of the disk be at the origin; in polar coordinates we have $0 \leq r \leq R$, $0 \leq \theta \leq 2\pi$, and

$$\operatorname{area} D_R = \iint_{D_R} dS = \iint_{D_R} rdrd\theta = \int_0^{2\pi} d\theta \int_0^R rdr = 2\pi \frac{1}{2} R^2 = \pi R^2.$$

Example 2.12. Similarly, for a sector of the same disk, of angle $\Delta\theta$, we have (denoting the sector by S)

$$\operatorname{area} S = \iint_S dS = \iint_S rdrd\theta = \int_0^{\Delta\theta} d\theta \int_0^R rdr = \frac{1}{2} R^2 \Delta\theta.$$

Exactly in the same way this works for pieces of surfaces in $\mathbb{R}^3$. If a surface M is parametrized by parameters u, v, then we again have vectors

$$\mathbf{e}_u = \frac{\partial \mathbf{x}}{\partial u}, \quad \mathbf{e}_v = \frac{\partial \mathbf{x}}{\partial v},$$

which are now in the tangent planes at points of M (so they vary from point to point). We have parallelograms $\Pi(\mathbf{e}_u \Delta u, \mathbf{e}_v \Delta v)$ in the tangent planes, approximating infinitesimal pieces of the surface M. Hence for the element of area we have the same formula

$$dS = \sqrt{g}\, dudv \quad \text{where } g = \det G(\mathbf{e}_u, \mathbf{e}_v),$$

and the area of (a piece of) M is given by a double integral:

$$\operatorname{area} M = \iint_M dS = \iint_M \sqrt{g}\, dudv\,.$$

Example 2.13. Points $\mathbf{x} = (x, y, z)$ of the sphere of radius R with center at the origin can be parametrized by angles θ, φ, where $0 \leq \theta \leq \pi$, $0 \leq \varphi \leq 2\pi$, as follows: $x = R \sin\theta \cos\varphi$, $y = R \sin\theta \sin\varphi$, $z = R\cos\theta$. Hence we obtain

$$\begin{aligned}
\mathbf{e}_\theta &= \frac{\partial \mathbf{x}}{\partial \theta} = (R\cos\theta\cos\varphi, R\cos\theta\sin\varphi, -R\sin\theta) \\
\mathbf{e}_\varphi &= \frac{\partial \mathbf{x}}{\partial \varphi} = (-R\sin\theta\sin\varphi, R\sin\theta\cos\varphi, 0)\,.
\end{aligned}$$

An immediate calculation gives the following Gram matrix:

$$G(\mathbf{e}_\theta, \mathbf{e}_\varphi) = \begin{pmatrix} \mathbf{e}_\theta \cdot \mathbf{e}_\theta & \mathbf{e}_\theta \cdot \mathbf{e}_\varphi \\ \mathbf{e}_\varphi \cdot \mathbf{e}_\theta & \mathbf{e}_\varphi \cdot \mathbf{e}_\varphi \end{pmatrix} = \begin{pmatrix} R^2 & 0 \\ 0 & R^2\sin^2\theta \end{pmatrix}.$$

Hence $g = R^4 \sin^2\theta$, and for the element of area for the sphere we get

$$dS = R^2 \sin\theta\, d\theta d\varphi.$$

2.5 Problems

Problem 2.1. Find the areas and volumes (signed where indicated):
(a) $\operatorname{Area}\Pi(\mathbf{a}, \mathbf{b})$ if $\mathbf{a} = (-\cos\alpha, -\sin\alpha)$, $\mathbf{b} = (-\sin\alpha, \cos\alpha)$ in $\mathbb{R}^2$; make a sketch;
(b) $\operatorname{Vol}\Pi(\mathbf{a}, \mathbf{b}, \mathbf{c})$ if $\mathbf{a} = (3, 2, -1)$, $\mathbf{b} = (2, 2, 5)$, $\mathbf{c} = (0, 0, 1)$ in $\mathbb{R}^3$;
(c) $\operatorname{area}(\mathbf{a}, \mathbf{b})$ if $\mathbf{a} = (1, -1, 2, 3)$, $\mathbf{b} = (0, 3, 1, 2)$ in $\mathbb{R}^4$.
(*Ans.:* **(a)** -1; **(b)** 2; **(c)** 15.)

Problem 2.2. Verify by a direct calculation that the Gram determinant $\det G(\mathbf{a}, \mathbf{b}, \mathbf{c})$ vanishes if one of the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is a linear combination of the others. (Geometrically that means that they are in the same plane.) What can be said about the volume of the parallelipiped $\Pi(\mathbf{a}, \mathbf{b}, \mathbf{c})$?

Problem 2.3. Show that the oriented area of a triangle ABC in $\mathbb{R}^2$ is given by the formula

$$\operatorname{area}(ABC) = \frac{1}{2} \begin{vmatrix} A_1 & A_2 & 1 \\ B_1 & B_2 & 1 \\ C_1 & C_2 & 1 \end{vmatrix}$$

if $A = (A_1, A_2)$, $B = (B_1, B_2)$, $C = (C_1, C_2)$.

Problem 2.4. (See Problem 1.5.)
(a) Show that the so-called *triple* or *mixed product* $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ in $\mathbb{R}^3$ defined as $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ is the oriented volume of $\Pi(\mathbf{a}, \mathbf{b}, \mathbf{c})$.
(b) Apply the result of part (a) to prove that the length of the vector product $\mathbf{a} \times \mathbf{b}$ in $\mathbb{R}^3$ is the area of the parallelogram $\Pi(\mathbf{a}, \mathbf{b})$.
Hint: for part (b), consider the parallelipiped $\Pi(\mathbf{a}, \mathbf{b}, \mathbf{n})$ where $\mathbf{n}$ is a unit vector perpendicular to $\mathbf{a}$ and $\mathbf{b}$, and apply Theorem 2.3.

Problem 2.5. Show that $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}| \cdot \sin \alpha$ where α is the angle between $\mathbf{a}$ and $\mathbf{b}$.
Hint: use the result of the previous problem, part (b), and express the area by the Gram determinant.

Problem 2.6. Show that the distance of a point $C = (C_1, C_2)$ from the straight line passing through points $A = (A_1, A_2)$ and $B = (B_1, B_2)$ in the plane is given by the absolute value of the expression

$$\frac{1}{|B - A|} \begin{vmatrix} A_1 & A_2 & 1 \\ B_1 & B_2 & 1 \\ C_1 & C_2 & 1 \end{vmatrix}.$$

Problem 2.7. Find the area of the sphere S^2_R: $x^2 + y^2 + z^2 = R^2$ as the integral $\iint_{S^2_R} dS$ using the result of Example 2.13. (*Answer:* $4\pi R^2$.)