F. Determinants.[1]

The theory of determinants that we shall develop in this chapter is not needed in Galois theory. The reader may, therefore, omit this section if he so desires.

We assume our field to be c o m m ut a t i v e and consider the square matrix

1) Of the preceding theory only Theorem 1, for homogeneous equations and the notion of linear dependence are assumed known.

$$(1) \qquad \begin{pmatrix} a_{11} a_{12} \ldots a_{1n} \\ a_{21} a_{22} \ldots a_{2n} \\ \ldots\ldots\ldots\ldots \\ a_{n1} a_{n2} \ldots a_{nn} \end{pmatrix}$$

of n rows and n columns. We shall define a certain function of this matrix whose value is an element of our field. The function will be called the determinant and will be denoted by

$$(2) \qquad \begin{vmatrix} a_{11} a_{12} \ldots a_{1n} \\ a_{21} a_{22} \ldots a_{2n} \\ \ldots\ldots\ldots\ldots \\ a_{n1} a_{n2} \ldots a_{nn} \end{vmatrix}$$

orby $D(A_1, A_2, \ldots A_n)$ if we wish to consider it as a function of the column vectors A,, A,, . . . A,, of (1). If we keep all the columns but A, constant and consider the determinant as a function of A,, then we write $D_k(A_k)$ and sometimes even only D.

Definition. A function of the column vectors is a determinant if it satisfies the following three axioms:

1. Viewed as a function of any column A, it is linear and homogeneous, i.e.,

$$(3) \qquad D_k(A_k + A_k') = D_k(A_k) + D_k(A_k')$$

$$(4) \qquad D_k(cA_k) = c \cdot D_k(A_k)$$

2. Its value is $= 0$[1)] if the adjacent columns A, and A_{k+1} are equal.

3. Its value is = 1 if all A, are the unit vectors U_k where

---

1) Henceforth, 0 will denote the zero element of a field.

$$(5)\quad U_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix};\quad U_2 = \begin{pmatrix} 0 \\ 1 \\ \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \quad \cdot\ \cdot\ \cdot\ \cdot\ \cdot \quad U_n = \begin{pmatrix} 0 \\ 0 \\ \\ \cdot \\ \cdot \\ 1 \end{pmatrix}$$

The question as to whether determinants exist will be left open for the present. But we derive consequences from the axioms:

a) If we put $c = 0$ in (4) we get: a determinant is 0 if one of the columns is 0.

b) $D_k(A_k) = D_k(A_k + cA_{k+1})$ or a determinant remains unchanged if we add a multiple of one column to an adjacent column. Indeed

$$D_k(A_k + cA_{k\pm 1}) = D_k(A_k) + cD_k(A_{k\pm 1}) = D_k(A_k)$$

because of axiom 2.

c) Consider the two columns A, and A_{k+1}. We may replace them by A, and $A_{k+1} + A_k$; subtracting the second from the first we may replace them by $-A_{k+1}$ and A_{k+1} + A,, adding the first to the second we now have $-A_{k+1}$ and A,, finally, we factor out -1. We conclude: a determinant changes sign if we interchange two adjacent columns.

d) A determinant vanishes if any two of its columns are equal. Indeed, we may bring the two columns side by side after an interchange of adjacent columns and then use axiom 2. In the same way as in b) and c) we may now prove the more general rules:

e) Adding a multiple of one column to another does not change the value of the determinant.

f) Interchanging any two columns changes the sign of D.

g) Let $(\nu_1, \nu_2, \ldots \nu_n)$ be a permutation of the subscripts $(1, 2, \ldots n)$. If we rearrange the columns in $D(A_{\nu_1}, A_{\nu_2}, \ldots, A,,_n)$ until they are back in the natural order, we see that

$$D(A_{\nu_1}, A_{\nu_2}, \ldots, A,_n) = \pm D(A_1, A_2, \ldots, A_n).$$

Here $\pm$ is a definite sign that does not depend on the special values of the A,. If we substitute U_k for A, we see that $D(U_{\nu_1}, U_{\nu_2}, \ldots, U_{\nu_n}) = \pm 1$ and that the sign depends only on the permutation of the unit vectors.

Now we replace each vector A, by the following linear combination A'_k of $A_1, A_2, \ldots, A_n$:

(6) $A; = b_{1k}A_1 + b_{2k}A_2 + \ldots + b_{nk}A_n.$

In computing $D(A'_1, A'_2, \ldots, A'_n)$ we first apply axiom 1 on A'_1 breaking up the determinant into a sum; then in each term we do the same with A'_2 and so on. We get

$$(7)\quad D(A'_1, A'_2, \ldots, A'_n) = \sum_{\nu_1, \nu_2, \ldots, \nu_n} D(b_{\nu_1 1}A_{\nu_1}, b_{\nu_2 2}A_{\nu_2}, \ldots, b_{\nu_n n}A_{\nu_n})$$
$$= \sum_{\nu_1, \nu_2, \ldots, \nu_n} b_{\nu_1 1} \cdot b_{\nu_2 2} \cdot \ldots \cdot b_{\nu_n n} D(A_{\nu_1}, A,,_2 \ldots, A,_n)$$

where each ν_i runs independently from 1 to n. Should two of the indices ν_i be equal, then $D(A_{\nu_1}, A,_2, \ldots, A_{\nu_n}) = 0$; we need therefore keep only those terras in which (vi, $\nu_2, \ldots, \nu_n$) is a permutation of $(1, 2, \ldots, n)$. This gives

(8) $$D(A'_1, A'_2, \ldots, A'_n)$$
$$= D(A_1, A_2, \ldots, A,). \sum_{(\nu_1, \ldots, \nu_n)} \pm b_{\nu_1 1} \cdot b_{\nu_2 2} \cdot \ldots \cdot b_{\nu_n n}$$

where $(\nu_1, \nu_2, \ldots, \nu_n)$ runs through all the permutations of $(1, 2, \ldots, n)$ and where $\pm$ stands for the sign associated with that permutation. It is important to remark that we would have arrived at the same formula (8) if our function D satisfied only the first two of our axioms.

Many conclusions may be derived from (8).

We first assume axiom 3 and specialize the A_k to the unit vectors U_k of (5). This makes $A_k' = B_k$ where B_k is the column vector of the matrix of the b_{ik}. (8) yields now:

$$(9) \quad D(B_1, B_2, \ldots, B_n) = \sum_{(\nu_1, \nu_2, \ldots, \nu_n)} \pm\, b_{\nu_1 1} \cdot b_{\nu_2 2} \cdot \ldots \cdot b_{\nu_n n}$$

giving us an explicit formula for determinants and showing that they are uniquely determined by our axioms provided they exist at all.

With expression (9) we retum to formula (8) and get

$$(10) \quad D(A_1', A_2', \ldots, A_n') = D(A_1, A_2, \ldots, A_n) D(B_1, B_2, \ldots, B_n).$$

This is the so-called multiplication theorem for determinants. At the left of (10) we have the determinant of an n-rowed matrix whose elements c_{ik} are given by

$$(11) \qquad c_{ik} = \sum_{\nu=1}^{n} a_{i\nu} b_{\nu k}.$$

c_{ik} is obtained by multiplying the elements of the i - th row of $D(A_1, A_2, \ldots, A_n)$ by those of the k-th column of $D(B_1, B_2, \ldots, B_n)$ and adding.

Let us now replace D in (8) by a function $F(A_1, \ldots, A_n)$ that satisfies only the first two axioms. Comparing with (9) we find

$$F(A_1', A_2', \ldots, A_n') = F(A_1, \ldots, A_n) D(B_1, B_2, \ldots, B_n).$$

Specializing A_k to the unit vectors U_k leads to

$$(12) \quad F(B_1, B_2, \ldots, B_n) = c \cdot D(B_1, B_2, \ldots, B_n)$$

with $\quad c = F(U_1, U_2, \ldots, U_n)$.

Next we specialize (10) in the following way: If i is a certain subscript from 1 to n-l we put $A_k = U_k$ for $k \neq i$, i+ 1 $A_i = U_i + U_{i+1}$, $A_{i+1} = 0$. Then $D(A_1, A_2, \ldots, A_n) = 0$ since one column is 0. Thus, $D(A_1', A_2', \ldots, A_n') = 0$; but this determinant differs from that of the elements b_{jk} only in the respect that the $i+1$-st row has been made equal to the i-tb. We therefore see:

A determinant vanishes if two adjacent rows are equal.

Each term in (9) is a product where precisely one factor cornes from a given row, say, the i-th. This shows that the determinant is linear and homogeneous if çonsidered as function of this row. If, finally, we select for eaeh row the corresponding unit vector, the determinant is = 1 since the matrix is the same as that in which the columns are unit vectors. This shows that a determinant satisfies our three axioms if we consider it as function of the row vectors. In view of the uniqueness it follows:

A determinant remains unchanged if we transpose the row vectors into column vectors, that is, if we rotate the matrix about its main diagonal.

A determinant vanishes if any two rows are equal. It changes sign if we interchange any two rows. It remains unchanged if we add a multiple of one row to another.

We shall now prove the existence of determinants. For a 1-rowed matrix a $_{11}$ the element a_{11} itself is the determinant. Let us assume the existence of (n - 1) - rowed determinants. If we consider the n-rowed matrix (1) we may associate with it certain (n - 1) - rowed determinants in the following way: Let a_{ik} be a particular element in (1). We cancel the i-th row and k-th column in (1) and take the determinant of the remaining (n - 1) - rowed matrix. This determinant multiplied by $(-1)^{i+k}$ will be called the cofactor of a_{ik} and be denoted by A_{ik}. The distribution of the sign $(-1)^{i+k}$ follows the chessboard pattern, namely,

$$\begin{pmatrix} + & - & + & - & . & . & . & . \\ - & + & - & + & . & . & . & . \\ + & - & + & - & . & . & . & . \\ - & + & - & + & . & . & . & . \\ . & . & . & . & . & . & . & . \end{pmatrix}$$

Let i be any number from 1 to n. We consider the following function D of the matrix (1):

(13) $$D = a_{i1}A_{i1} + a_{i2}A_{i2} + \ . \ . + a_{in}A_{in} .$$

It is the sum of the products of the i-th row and their cofactors.

Consider this D in its dependence on a given column, say, A,. For $\nu \neq k$, $A_{i\nu}$ depends linearly on A, and $a_{i\nu}$ does not depend on it; for $\nu = k$, A_{ik} does not depend on A, but a_{ik} is one element of this column. Thus, axiom 1 is satisfied. Assume next that two adjacent columns A, and A_{k+1} are equal. For $\nu \neq k$, k + 1 we have then two equal columns in $A_{i\nu}$ so that A,, = 0. The determinants used in the computation of $A_{i\ k}$ and $A_{i\ k+1}$ are the same but the signs are opposite hence, $A_{i\ k} = -A_{i\ k+1}$ whereas $a_{i\ k}$ = a, $_{k+1}$. Thus D = 0 and axiom 2 holds. For the special case A, $= U_\nu(\nu = 1, 2, . . , n)$ we have $a_{i\nu} = 0$ for $\nu \neq i$ while a,, = 1, $A_{ii} = 1$. Hence, D = 1 and this is axiom 3. This proves both the existence of an n-rowed determinant as well as the truth of formula (13), the so-called development of a determinant according to its i-th row. (13) may be generalized as follows: In our determinant replace the i-th row by the j-th row and develop according to this new row. For $i \neq j$ that determinant is 0 and for $i = j$ it is D:

$$(14) \quad a_{j1}A_{i1} + a_{j2}A_{i2} + \cdots + a_{jn}A_{in} = \begin{cases} D & \text{for } j = i \\ 0 & \text{for } j \neq i \end{cases}$$

If we interchange the rows and the columns we get the following formula:

$$(15) \quad a_{1h}A_{1k} + a_{2h}A_{2k} + \cdots + a_{nh}A_{nk} = \begin{cases} D & \text{for } h = k \\ 0 & \text{for } h \neq k \end{cases}$$

Now let A represent an n-rowed and B an m-rowed square matrix. By $|A|$, $|B|$ we mean their determinants. Let C be a matrix of n rows and m columns and form the square matrix of $n + m$ rows

$$(16) \qquad \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

where 0 stands for a zero matrix with m rows and n columns. If we consider the determinant of the matrix (16) as a function of the columns of A only, it satisfies obviously the first two of our axioms. Because of (12) its value is $c \cdot |A|$ where c is the determinant of (16) after substituting unit vectors for the columns of A. This c still depends on B and considered as function of the rows of B satisfies the first two axioms. Therefore the determinant of (16) is $d \cdot |A| \cdot |B|$ where d is the special case of the determinant of (16) with unit vectors for the columns of A as well as of B. Subtracting multiples of the columns of A from C we can replace C by 0. This shows $d = 1$ and hence the formula

(17) $$\begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = |A| \cdot |B|.$$

In a similar fashion we could have shown

(18) $$\begin{vmatrix} A & 0 \\ C & B \end{vmatrix} = |A| \cdot |B|.$$

The formulas (17), (18) are special cases of a general theorem by Lagrange that can be derived from them. We refer the reader to any textbook on determinants since in most applications (17) and (18) are sufficient.

We now investigate what it means for a matrix if its determinant is zero. We can easily establish the following facts:

a) If A,, A,, . , . , A_n are linearly dependent, then $D(A_1$, A,, . . . , A,) = 0. Indeed one of the vectors, say A,, is then a linear combination of the other columns; subtracting this linear combination from the column A, reduces it to 0 and so D = 0.

b) If any vector B can be expressed as linear combination of A,, A,, . . . , A, then $D(A_1, A_2, \ldots,$ A,,) $\neq$ 0. Returning to (6) and (10) we may select the values for b_{ik} in such a fashion that every $A_i' = U_i$. For this choice the left side in (10) is 1 and hence $D(A_1, A_2, \ldots,$ A,) on the right side $\neq$ 0.

c) Let A,, A,, . . . , A,, be linearly independent and B any other vector. If we go back to the components in the equation $A_1x_1 + A_2x_2 + \ldots + A_nx_n +$ By = 0 we obtain n linear homogeneous equations in the n + 1 unknowns $x_1, x_2, \ldots, x_n$, y. Consequently, there is a non-trivial solution. y must be $\neq$ 0 or else the $A_1, A_2, \ldots, A_n$ would be linearly dependent. But then we can compute B out of this equation as a linear combination of A,, A,, . . . , A_n.

Combining these results we obtain:

A determinant vanishes if and only if the column vectors (or the row vectors) are linearly dependent.

Another way of expressing this result is:

The set of n linear homogeneous equations

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = 0 \qquad (i = 1, 2, \dots, n)$$

in n unknowns has a non-trivial solution if and only if the determinant of the coefficients is zero.

Another result that can be deduced is:

If $A_1, A_2, \dots, A_n$ are given, then their linear combinations can represent any other vector B if and only if $D(A_1, A_2, \dots, A_n) \neq 0$.

Or:

The set of linear equations

$$(19) \quad a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \qquad (i = 1, 2, \dots, n)$$

has a solution for arbitrary values of the b_i if and only if the determinant of a_{ik} is $\neq 0$. In that case the solution is unique.

We finally express the solution of (19) by means of determinants if the determinant D of the a_{ik} is $\neq 0$.

We multiply for a given k the i-th equation with A_{ik} and add the equations. (15) gives

$$(20) \quad D \cdot x_k = A_{1k}b_1 + A_{2k}b_2 + \dots + A_{nk}b_n \qquad (k = 1, 2, \dots, n)$$

and this gives x_k. The right side in (12) may also be written as the determinant obtained from D by replacing the k-th column by $b_1, b_2, \dots, b_n$. The rule thus obtained is known as Cramer's rule.