I LINEAR ALGEBRA

A. Fields.

A field is a set of elements in which a pair of operations called multiplication and addition is defined analogous to the operations of multiplication and addition in the real number system (which is itself an example of a field). In each field F there exist unique elements called o and 1 which, under the operations of addition and multiplication, behave with respect to all the other elements of F exactly as their correspondents in the real number system. In two respects, the analogy is not complete: 1) multiplication is not assumed to be commutative in every field, and 2) a field may have only a finite number of elements.

More exactly, a field is a set of elements which, under the above mentioned operation of addition, forms an additive abelian group and for which the elements, exclusive of zero, form a multiplicative group and, finally, in which the two group operations are connected by the distributive law. Furthermore, the product of o and any element is defined to be o.

If multiplication in the field is commutative, then the field is called a commutative field.

B. Vector Spaces.

If V is an additive abelian group with elements A, B, . . . , F a field with elements a, b, . . . , and if for each $a \epsilon F$ and $A \epsilon V$

the product aA denotes an element of V, then V is called a (left) vector space over F if the following assumptions hold:

1) $a(A + B) = aA + aB$

2) $(a + b)A = aA + bA$

3) $a(bA) = (ab)A$

4) $1A = A$

The reader may readily verify that if V is a vector space over F, then $oA = O$ and $aO = O$ where o is the zero element of F and O that of V. For example, the first relation follows from the equations:

$$aA = (a + o)A = aA + oA$$

Sometimes products between elements of F and V are written in the form Aa in which case V is called a right vector space over F to distinguish it from the previous case where multiplication by field elements is from the left. If, in the discussion, left and right vector spaces do not occur simultaneously, we shall simply use the term "vector space."

C. Homogeneous Linear Equations.

If in a field F, a_{ij}, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ are $m \cdot n$ elements, it is frequently necessary to know conditions guaranteeing the existence of elements in F such that the following equations are satisfied:

$$\begin{array}{ll} & a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = o. \\ (1) & \quad \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = o. \end{array}$$

The reader will recall that such equations are called linear homogeneous equations, and a set of elements, $x_1, x_2, \ldots, x_n$ of F, for which all the above equations are true, is called a solution of the system. If not all of the elements $x_1, x_2, \ldots, x_n$ are o the solution is called non-trivial; otherwise, it is called trivial.

THEOREM 1. A system of linear homogeneous equations always has a non-trivial solution if the number of unknowns exceeds the number of equations.

The proof of this follows the method familiar to most high school students, namely, successive elimination of unknowns. If no equations in $n > 0$ variables are prescribed, then our unknowns are unrestricted and we may set them all $= 1$.

We shall proceed by complete induction. Let us suppose that each system of k equations in more than k unknowns has a non-trivial solution when $k < m$. In the system of equations (1) we assume that $n > m$, and denote the expression $a_{i1}x_1 + \ldots + a_{in}x_n$ by L_i, $i = 1, 2, \ldots, m$. We seek elements $x_1, \ldots, x_n$ not all o such that $L_1 = L_2 = \ldots = L_m = o$. If $a_{ij} = o$ for each i and j, then any choice of $x_1, \ldots, x_n$ will serve as a solution. If not all a_{ij} are o, then we may assume that $a_{11} \neq o$, for the order in which the equations are written or in which the unknowns are numbered has no influence on the existence or non-existence of a simultaneous solution. We can find a non-trivial solution to our given system of equations, if and only if we can find a non-trivial solution to the following system:

$$
\begin{aligned}
&L_1 = o \\
&L_2 - a_{21}a_{11}^{-1}L_1 = o \\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
&L_m - a_{m1}a_{11}^{-1}L_1 = o
\end{aligned}
$$

For, if $x_1, \ldots, x_n$ is a solution of these latter equations then, since $L_1 = o$, the second term in each of the remaining equations is o and, hence, $L_2 = L_3 = \ldots = L_m = o$. Conversely, if (1) is satisfied, then the new system is clearly satisfied. The reader will notice that the new system was set up in such a way as to "eliminate" x_1 from the last m-1 equations. Furthermore, if a non-trivial solution of the last m-1 equations, when viewed as equations in $x_2, \ldots, x_n$, exists then taking $x_1 = - a_{11}^{-1}(a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n)$ would give us a solution to the whole system. However, the last m-1 equations have a solution by our inductive assumption, from which the theorem follows.

Remark: If the linear homogeneous equations had been written in the form $\Sigma x_j a_{ij} = o$, $j = 1, 2, \ldots, n$, the above theorem would still hold and with the same proof although with the order in which terms are written changed in a few instances.

D. Dependence and Independence of Vectors.

In a vector space V over a field F, the vectors $A_1, \ldots, A_n$ are called dependent if there exist elements $x_1, \ldots, x_n$, not all o, of F such that $x_1A_1 + x_2A_2 + \ldots + x_nA_n = O$. If the vectors $A_1, \ldots, A_n$ are not dependent, they are called independent.

The dimension of a vector space V over a field F is the maximum number of independent elements in V. Thus, the dimension of V is n if there are n independent elements in V, but no set of more than n independent elements.

A system $A_1, \ldots, A_m$ of elements in V is called a generating system of V if each element A of V can be expressed linearly in terms of $A_1, \ldots, A_m$, i.e., $A = \sum_{i=1}^{m} a_i A_i$ for a suitable choice of a_i, $i = 1, \ldots, m$, in F.

<u>THEOREM 2.</u> <u>In any generating system the maximum number of independent vectors is equal to the dimension of the vector space.</u>

Let $A_1, \ldots, A_m$ be a generating system of a vector space V of dimension n. Let r be the maximum number of independent elements in the generating system. By a suitable reordering of the generators we may assume $A_1, \ldots, A_r$ independent. By the definition of dimension it follows that $r \leq n$. For each j, $A_1, \ldots, A_r$, A_{r+j} are dependent, and in the relation

$$a_1 A_1 + a_2 A_2 + \ldots + a_r A_r + a_{r+j} A_{r+j} = O$$

expressing this, $a_{r+j} \neq o$, for the contrary would assert the dependence of $A_1, \ldots, A_r$. Thus,

$$A_{r+j} = -a_{r+j}^{-1}[a_1 A_1 + a_2 A_2 + \ldots + a_r A_r].$$

It follows that $A_1, \ldots, A_r$ is also a generating system since in the linear relation for any element of V the terms involving A_{r+j}, $j \neq o$, can all be replaced by linear expressions in $A_1, \ldots, A_r$.

Now, let $B_1, \ldots, B_t$ be any system of vectors in V where $t > r$, then there exist a_{ij} such that $B_j = \sum_{i=1}^{r} a_{ij} A_i$, $j = 1, 2, \ldots, t$, since the A_i's form a generating system. If we can show that $B_1, \ldots, B_t$ are dependent, this will give us $r \geq n$, and the theorem will follow from this together with the previous inequality $r \leq n$. Thus, we must exhibit the existence of a non-trivial solution out of F of the equation

$$x_1 B_1 + x_2 B_2 + \ldots + x_t B_t = O.$$

To this end, it will be sufficient to choose the x_i's so as to satisfy the linear equations $\sum_{j=1}^{t} x_j a_{ij} = o$, $i = 1, 2, \ldots, r$, since these expressions will be the coefficients of A_i when in $\sum_{j=1}^{t} x_j B_j$ the B_j's are replaced by $\sum_{i=1}^{r} a_{ij} A_i$ and terms are collected. A solution to the equations $\sum_{j=1}^{t} x_j a_{ij} = o$, $i = 1, 2, \ldots, r$, always exists by Theorem 1.

Remark: Any n independent vectors $A_1, \ldots, A_n$ in an n dimensional vector space form a generating system. For any vector A, the vectors A, $A_1, \ldots, A_n$ are dependent and the coefficient of A, in the dependence relation, cannot be zero. Solving for A in terms of $A_1, \ldots, A_n$, exhibits $A_1, \ldots, A_n$ as a generating system.

A subset of a vector space is called a subspace if it is a subgroup of the vector space and if, in addition, the multiplication of any element in the subset by any element of the field is also in the subset. If $A_1, \ldots, A_s$ are elements of a vector space V, then the set of all elements of the form $a_1 A_1 + \ldots + a_s A_s$ clearly forms a subspace of V. It is also evident, from the definition of dimension, that the dimension of any subspace never exceeds the dimension of the whole vector space.

An s-tuple of elements $(a_1, \ldots, a_s)$ in a field F will be called a row vector. The totality of such s-tuples form a vector space if we define

α) $(a_1, a_2, \ldots, a_s) = (b_1, b_2, \ldots, b_s)$ if and only if $a_1 = b_1$, $i = 1, \ldots, s$,

β) $(a_1, a_2, \ldots, a_s) + (b_1, b_2, \ldots, b_s) = (a_1 + b_1, a_2 + b_2, \ldots, a_s + b_s)$,

γ) $b(a_1, a_2, \ldots, a_s) = (ba_1, ba_2, \ldots, ba_s)$, for b an element of F.

When the s-tuples are written vertically, $\begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_s \end{pmatrix}$

they will be called column vectors.

THEOREM 3. The row (column) vector space F^n of all n-tuples from a field F is a vector space of dimension n over F.

The n elements

$$\begin{aligned} \epsilon_1 &= (1, 0, 0, \ldots, 0) \\ \epsilon_2 &= (0, 1, 0, \ldots, 0) \\ &\cdot \\ &\cdot \\ &\cdot \\ \epsilon_n &= (0, 0, \ldots, 0, 1) \end{aligned}$$

are independent and generate F^n. Both remarks follow from the relation $(a_1, a_2, \ldots, a_n) = \Sigma a_i \epsilon_i$.

We call a rectangular array

$$\begin{pmatrix} a_{11} a_{12} \ldots a_{1n} \\ a_{21} a_{22} \ldots a_{2n} \\ \cdot \quad \cdot \quad \quad \cdot \\ \cdot \quad \cdot \quad \quad \cdot \\ \cdot \quad \cdot \quad \quad \cdot \\ a_{m1} a_{m2} \ldots a_{mn} \end{pmatrix}$$

of elements of a field F a matrix. By the right row rank of a matrix, we mean the maximum number of independent row vectors among the rows $(a_{i1}, \ldots, a_{in})$ of the matrix when multiplication by field elements is from the right. Similarly, we define left row rank, right column rank and left column rank.

THEOREM 4. In any matrix the right column rank equals the left row rank and the left column rank equals the right row rank. If the field is commutative, these four numbers are equal to each other and are called the rank of the matrix.

Call the column vectors of the matrix $C_1, \ldots, C_n$ and the row vectors $R_1, \ldots, R_m$. The column vector O is $\begin{pmatrix} o \\ o \\ \cdot \\ \cdot \\ o \end{pmatrix}$ and any dependence $C_1 x_1 + C_2 x_2 + \ldots + C_n x_n = O$ is equivalent to a solution of the equations

$$
(1) \qquad \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = o \\ \cdot \qquad\qquad\qquad\qquad \cdot \\ \cdot \qquad\qquad\qquad\qquad \cdot \\ \cdot \qquad\qquad\qquad\qquad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = o. \end{array}
$$

Any change in the order in which the rows of the matrix are written gives rise to the same system of equations and, hence, does not change the column rank of the matrix, but also does not change the row rank since the changed matrix would have the same set of row vectors. Call c the right column rank and r the left row rank of the matrix. By the above remarks we may assume that the first r rows are independent row vectors. The row vector space generated by all the rows of the matrix has, by Theorem 1, the dimension r and is even generated by the first r rows. Thus, each row after the r^{th} is linearly expressible in terms of the first r rows. Consequently, any solution of the first r equations in (1) will be a solution of the entire system since any of the last n-r equations is obtainable as a linear combination of the first r. Conversely, any solution of (1) will also be a solution of the first r equations. This means that the matrix

$$\begin{pmatrix} a_{11} a_{12} \dots a_{1n} \\ \cdot \quad \cdot \quad\quad \cdot \\ \cdot \quad \cdot \quad\quad \cdot \\ \cdot \quad \cdot \quad\quad \cdot \\ a_{r1}\ a_{r2} \dots a_{rn} \end{pmatrix}$$

consisting of the first r rows of the original matrix has the same right column rank as the original. It has also the same left row rank since the r rows were chosen independent. But the column rank of the amputated matrix cannot exceed r by Theorem 3. Hence, $c \leq r$. Similarly, calling c' the left column rank and r' the right row rank, $c' \leq r'$.

If we form the transpose of the original matrix, that is, replace rows by columns and columns by rows, then the left row rank of the transposed matrix equals the left column rank of the original. If then to the transposed matrix we apply the above considerations we arrive at $r \leq c$ and $r' \leq c'$.

E. Non-homogeneous Linear Equations.

The system of non-homogeneous linear equations

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \dots\dots\dots\dots + a_{2n}x_n = b_2 \\ \cdot \qquad\qquad\qquad\qquad\quad \cdot \qquad \cdot \\ \cdot \qquad\qquad\qquad\qquad\quad \cdot \qquad \cdot \\ \cdot \qquad\qquad\qquad\qquad\quad \cdot \qquad \cdot \\ a_{m1}x_1 + \dots\dots\dots\dots + a_{mn}x_m = b_m \end{array} \tag{2}$$

has a solution if and only if the column vector $\begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix}$ lies

in the space generated by the vectors $\begin{pmatrix} a_{11} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{pmatrix}$

This means that there is a solution if and only if the right column rank of the matrix $\begin{pmatrix} a_{11} \dots a_{1n} \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ \cdot \quad \cdot \\ a_{m1} \dots a_{mn} \end{pmatrix}$ is the same as the right column rank of the augmented matrix $\begin{pmatrix} a_{11} \dots a_{1n} b_1 \\ \cdot \quad \cdot \; \cdot \\ \cdot \quad \cdot \; \cdot \\ \cdot \quad \cdot \; \cdot \\ a_{m1} \dots a_{mn} b_m \end{pmatrix}$ since the vector space generated by the original must be the same as the vector space generated by the augmented matrix and in either case the dimension is the same as the rank of the matrix by Theorem 2.

By Theorem 4, this means that the row ranks are equal. Conversely, if the row rank of the augmented matrix is the same as the row rank of the original matrix, the column ranks will be the same and the equations will have a solution.

If the equations (2) have a solution, then any relation among the rows of the original matrix subsists among the rows of the augmented matrix. For equations (2) this merely means that like combinations of equals are equal. Conversely, if each relation which subsists between the rows of the original matrix also subsists between the rows of the augmented matrix, then the row rank of the augmented matrix is the same as the row rank of the original matrix. In terms of the equations this means that there will exist a solution if and only if the equations are consistent, i.e., if and only if any dependence between the left hand sides of the equations also holds between the right sides.

<u>THEOREM 5. If in equations (2) $m = n$, there exists a unique solution if and only if the corresponding homogeneous equations</u>

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ \quad\vdots \qquad\qquad\qquad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = 0 \end{array}$$

<u>have only the trivial solution.</u>

If they have only the trivial solution, then the column vectors are independent. It follows that the original n equations in n unknowns will have a unique solution if they have any solution, since the difference, term by term, of two distinct solutions would be a non-trivial solution of the homogeneous equations. A solution would exist since the n independent column vectors form a generating system for the n-dimensional space of column vectors.

Conversely, let us suppose our equations have one and only one solution. In this case, the homogeneous equations added term by term to a solution of the original equations would yield a new solution to the original equations. Hence, the homogeneous equations have only the trivial solution.

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 *commutative* 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.

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

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

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

or by $D(A_1, A_2, \ldots A_n)$ if we wish to consider it as a function of the column vectors $A_1, A_2, \ldots A_n$ of (1). If we keep all the columns but A_k constant and consider the determinant as a function of A_k, 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_k it is linear and homogeneous, i.e.,

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

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

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

3. Its value is $= 1$ if all A_k are the unit vectors U_k where

---

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

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

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

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

because of axiom 2.

c) Consider the two columns A_k and A_{k+1}. We may replace them by A_k and $A_{k+1} + A_k$; subtracting the second from the first we may replace them by $-A_{k+1}$ and $A_{k+1} + A_k$; adding the first to the second we now have $-A_{k+1}$ and A_k; 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 $D(A_{\nu_1}, A_{\nu_2}, \ldots, A_{\nu_n})$ until they are back in the natural order, we see that

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

Here $\pm$ is a definite sign that does not depend on the special values of the A_k. If we substitute U_k for A_k 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_k by the following linear combination A_k' of $A_1, A_2, \ldots, A_n$:

(6) $\quad A_k' = b_{1k}A_1 + b_{2k}A_2 + \ldots + b_{nk}A_n.$

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

$$(7)\quad D(A_1', A_2', \ldots, A_n') = \sum_{\nu_1, \nu_2, \ldots, \nu_n} D(b_{\nu_1 1}A_{\nu_1}, b_{\nu_2 2}A_{\nu_2}, \ldots, b_{\nu_n n}A_{\nu_n})$$
$$= \sum_{\nu_1, \nu_2, \ldots, \nu_n} b_{\nu_1 1} \cdot b_{\nu_2 2} \cdot \ldots \cdot b_{\nu_n n} D(A_{\nu_1}, A_{\nu_2}, \ldots, A_{\nu_n})$$

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

$$(8)\quad D(A_1', A_2', \ldots, A_n')$$
$$= D(A_1, A_2, \ldots, A_n) \cdot \sum_{(\nu_1, \ldots, \nu_n)} \pm b_{\nu_1 1} \cdot b_{\nu_2 2} \cdot \ldots \cdot b_{\nu_n n}$$

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

Next we specialize (10) in the following way: If i is a certain subscript from 1 to n-1 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-th. We therefore see:

A determinant vanishes if two adjacent rows are equal.

Each term in (9) is a product where precisely one factor comes from a given row, say, the i-th. This shows that the determinant is linear and homogeneous if considered as function of this row. If, finally, we select for each 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,

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

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

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

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_k. For $\nu \neq k$, $A_{i\nu}$ depends linearly on A_k and $a_{i\nu}$ does not depend on it; for $\nu = k$, A_{ik} does not depend on A_k but a_{ik} is one element of this column. Thus, axiom 1 is satisfied. Assume next that two adjacent columns A_k and A_{k+1} are equal. For $\nu \neq k, k+1$ we have then two equal columns in $A_{i\nu}$ so that $A_{i\nu} = 0$. 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 $a_{i,k} = a_{i,k+1}$. Thus $D = 0$ and axiom 2 holds. For the special case $A_\nu = U_\nu (\nu = 1, 2, \dots, n)$ we have $a_{i\nu} = 0$ for $\nu \neq i$ while $a_{ii} = 1$, $A_{ii} = 1$. 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:

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

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

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

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

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

In a similar fashion we could have shown

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

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 $A_1, A_2, \ldots, A_n$ are linearly dependent, then $D(A_1, A_2, \ldots, A_n) = 0$. Indeed one of the vectors, say A_k, is then a linear combination of the other columns; subtracting this linear combination from the column A_k reduces it to 0 and so $D = 0$.

b) If any vector B can be expressed as linear combination of $A_1, A_2, \ldots, A_n$ then $D(A_1, A_2, \ldots, A_n) \neq 0$. 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 $D(A_1, A_2, \ldots, A_n)$ on the right side $\neq 0$.

c) Let $A_1, A_2, \ldots, A_n$ 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 $A_1, A_2, \ldots, A_n$.

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.