

0151.2
E-608

15219

中文书库

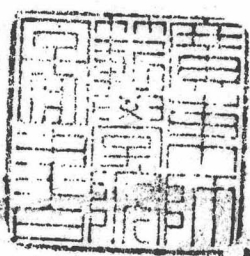
FINITE MATRICES

~~P133~~
F368

BY

W. L. FERRAR, M.A., D.Sc.

FELLOW OF HERTFORD COLLEGE
OXFORD



暨南大學
數學系資料室

OXFORD

AT THE CLARENDON PRESS

1951

Oxford University Press, Amen House, London E.C.4

GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON

BOMBAY CALCUTTA MADRAS CAPE TOWN

Geoffrey Cumberlege, Publisher to the University

PRINTED IN GREAT BRITAIN

PREFACE

THIS book is written for graduate students and for undergraduates whose degree courses include more matrix theory than a text-book of elementary properties will provide. In it I have tried to give an account of the theory of finite matrices, including their invariant factors and elementary divisors, which can be read with reasonable ease by mathematicians who are not specialists in this particular field. I have worked with the ordinary numbers of analysis and have not considered, save for an odd reference or two, the demands of an abstract algebra. My aim throughout has been to make the argument simple and straightforward.

When I began the book I expected that the whole of it would be concerned with the presentation of results long since known in some form or other. On reaching the chapter on functions of matrices I found that, starting from a few 'well-known' facts, the theory unfolded itself naturally and easily, but that only patches of it here and there appeared to have been published before. Accordingly, Chapter V is largely a first essay at a connected account of this part of the theory.

My indebtedness to other books and to research papers is very great. The reader who wishes to acquire a knowledge of the wider field within which my own limited treatment lies should consult, among others:

H. W. Turnbull and A. C. Aitken, *An Introduction to the Theory of Canonical Matrices* (Blackie, 1932);

W. V. D. Hodge and D. Pedoe, *Methods of Algebraic Geometry* (Cambridge, 1947);

G. Julia, *Introduction mathématique aux théories quantiques* (Gauthier-Villars, 1949), Part I on Finite Matrices and Part II on Hilbert Space and Infinite Matrices;

A. A. Albert, *Modern Higher Algebra* (Chicago, 1937);

J. H. M. Wedderburn, *Lectures on Matrices* (Amer. Math. Soc. Colloquium Publications, vol. xvii, 1934);

C. C. MacDuffee, *The Theory of Matrices* (Chelsea Publishing Co., New York, 1946: reprint of first edition).

For anything concerning matrices that was known prior to 1932 MacDuffee's book is invaluable. A similar account of what has been done since 1932 would be a great asset; is it too much to hope that a scholar might one day write it or edit a series of B.Sc. and Ph.D. theses written to that end? The present book makes no pretence to be complete, even in the central topics of finite matrices: it attempts a clear and readable account of the principal theorems and no more.

I end with an acknowledgement of my debt to the staff of the Clarendon Press. I have no immediate plans for another book with which to tax their skill and forbearance and so, on this occasion, I wish particularly to thank all of them for the way in which, over a period of some fifteen years, a series of not too tidy manuscripts has been transformed into well-printed books.

HERTFORD COLLEGE, OXFORD

W. L. F.

April 1951

CONTENTS

I. INTRODUCTION	1
II. EQUIVALENT MATRICES	15
III. EQUIVALENT λ -MATRICES	26
IV. COLLINEATION	48
V. INFINITE SERIES AND FUNCTIONS OF MATRICES	87
VI. CONGRUENCE	135
VII. MATRIX EQUATIONS	153
VIII. MISCELLANEOUS NOTES	168
INDEX	181

CHAPTER I

INTRODUCTION

1. Scope of the chapter

THE aim of this introductory chapter is to provide a résumé of the more elementary properties of matrices. I have thought it useful, for both writer and reader, to note explicitly, even though it be briefly and sketchily, the accepted facts about matrices on which the later chapters will be based. Some proofs, but not all, will be given. I have tried to hold the balance between brevity and clarity and, accordingly, I have omitted many details that would find a place in a full account of what is given here in outline.

2. Notation

(a) A set of mn numbers arranged in m columns and n rows is called a matrix.† Thus

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

is a matrix. The square bracket is a conventional symbol which is read as 'the matrix'. The individual numbers are referred to as the **ELEMENTS OF THE MATRIX**.

We shall normally use the square bracket whenever we wish to indicate that an array of numbers is to be considered as a matrix: thus

$$[x_1, x_2, \dots, x_n]$$

indicates that the n letters x_1, x_2, \dots, x_n are to be considered as a **ONE-ROW MATRIX**. To indicate that n letters x_1, x_2, \dots, x_n are to be considered as a **ONE-COLUMN MATRIX** we use the special notation

$$\{x_1, x_2, \dots, x_n\}.$$

When a matrix has n rows and n columns we refer to it as a **SQUARE MATRIX OF ORDER n** .

† Some writers insist that the laws of addition and multiplication must be laid down before the use of the word matrix can be justified.

(b) Capital italic letters A, B, \dots, X will be used to denote matrices, in general square matrices of order n . To indicate the actual elements of a matrix we shall write down the element in the i th row and k th column; thus

$$A = [a_{ik}], \quad B = [\xi_{ki}]$$

means that A has the element a_{ik} in the i th row and k th column, while B has the element ξ_{ki} in the i th row and k th column.

The DETERMINANT whose elements are precisely those of a square matrix A is denoted by $|A|$. When $|A| = 0$ the matrix is said to be SINGULAR and when $|A| \neq 0$ the matrix is said to be NON-SINGULAR.

We sometimes use a special notation for matrices having a single row or a single column. Clarendon letters $\mathbf{a}, \mathbf{b}, \dots, \mathbf{x}$ denote single-column matrices and (anticipating the later definition of a transpose) $\mathbf{a}', \mathbf{b}', \dots, \mathbf{x}'$ denote single-row matrices; thus

$$\mathbf{x} = \{x_1, \dots, x_n\}, \quad \mathbf{x}' = [x_1, \dots, x_n]$$

means that \mathbf{x} is the single-column matrix and \mathbf{x}' the single-row matrix having the elements shown.

(c) It being understood that, unless the contrary is stated, all literal suffixes run from 1 to n , we shall use the SUMMATION CONVENTION for repeated literal suffixes. With this convention

$$a_{rs}x_s \quad \text{denotes} \quad \sum_{s=1}^n a_{rs}x_s,$$

$$\text{and} \quad a_{rs}x_{rs} \quad \text{denotes} \quad \sum_{r=1}^n \sum_{s=1}^n a_{rs}x_{rs}.$$

On the other hand, a repeated numerical suffix, such as the suffix 1 in $a_{r1}x_1$, will not imply a summation.

When the occasion arises we shall enclose a literal suffix in brackets to denote that there is to be no summation with respect to that particular suffix; thus

$$a_{rs}x_{r(s)} \quad \text{denotes} \quad \sum_{r=1}^n a_{rs}x_{rs},$$

there being no summation with respect to s .

(d) The square matrix of order n having unity in each place on its leading diagonal and zero elsewhere is called the UNIT

MATRIX of order n ; it is denoted by I . Sometimes we use I_r , I_s, \dots to denote unit matrix of order r, s, \dots

A matrix having zero in every place is called a NULL MATRIX and is denoted by 0.

A square matrix of order n whose only non-zero elements occur in its leading diagonal is called a DIAGONAL MATRIX.

3. Addition and multiplication

(a) *Addition.* $[a_{ik}] + [b_{ik}] = [a_{ik} + b_{ik}].$ (1)

The definition applies to any two matrices A and B , not necessarily square, provided that each has the same number of rows and each has the same number of columns. Moreover, from the definition,

$$A + B = B + A.$$

(b) *Multiplication.*

$$[a_{ik}] \times [b_{ik}] = [a_{i\lambda} b_{\lambda k}].$$
 (2)

The definition applies to any two matrices A and B , not necessarily square, provided that the number of columns in A is equal to the number of rows in B . The product AB has as many rows as A and as many columns as B .

From the definition,

$$A(B + C) = AB + AC$$

and $(B + C)A = BA + CA.$

Further, $[a_{ik}] \times [b_{ik}] \times [c_{ik}]$,
whether the triple product be formed by

$$(AB)C \quad \text{or} \quad A(BC),$$

is equal to $[a_{i\lambda} b_{\lambda\mu} c_{\mu k}]$ (3)

and is commonly denoted by ABC . Products of four or more matrices are formed on the same pattern: thus

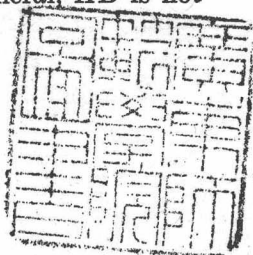
$$ABC \dots Z = [a_{i\lambda} b_{\lambda\mu} c_{\mu\nu} \dots z_{\rho k}],$$

the i, k being the only suffixes that do not imply summation.

By their definitions, AB and BA are matrices whose elements are formed by different processes and, in general, AB is not equal to BA . On the other hand

$$AI = IA = A$$

for every square matrix A of order n .



Finally, the equation $AB = 0$ may be true when neither A nor B is a null matrix; but if either A or B is known to be non-singular when $AB = 0$, then the other is necessarily a null matrix.

4. Related matrices

(a) *The reciprocal of a matrix.*

When $A = [a_{ik}]$, the determinant obtained from $|A|$ by deleting the r th row and s th column and multiplying by the sign-factor $(-1)^{r+s}$ is denoted by A_{rs} ; it is called the cofactor of a_{rs} in $|A|$. The matrix† $[A_{ki}]$

is called the ADJUGATE or ADJOINT of A .

It is a well-known result in the theory of determinants that

$$a_{ij}A_{kj} = 0 \quad (i \neq k) \quad (4)$$

and
$$a_{ij}A_{(ij)} = |A| \quad (i = 1, 2, \dots, n). \quad (5)$$

Hence, when A is a non-singular matrix,

$$[a_{ik}] \times [A_{ki}/|A|] = [a_{ij}A_{kj}/|A|] = I$$

and similarly, on working with columns instead of rows at lines (4) and (5),

$$[A_{ki}/|A|] \times [a_{ik}] = I.$$

Accordingly, we call the matrix

$$[A_{ki}/|A|]$$

the RECIPROCAL of A and denote it by A^{-1} .

When A is a singular matrix, $|A| = 0$ and the division by $|A|$ is no longer valid: the reciprocal is not then definable.

Moreover A^{-1} is the only matrix with the property that its product by A is equal to I . For, if $RA = I$, then

$$(R - A^{-1})A = RA - I = 0;$$

on multiplying by A^{-1} , we get

$$(R - A^{-1})AA^{-1} = 0,$$

or, since $AA^{-1} = I$, $R - A^{-1} = 0$.

Thus a matrix R for which $RA = I$ must be equal to A^{-1} . Similarly, $AR = I$ implies $R = A^{-1}$.

† Notice that A_{ki} comes in the i th row and k th column.

*(b) The powers of a matrix.*

The notations A^2, A^3, \dots , stand for AA, AA^2, \dots ; A^{-2}, A^{-3}, \dots stand for $A^{-1}A^{-1}, A^{-1}A^{-2}, \dots$; and with this notation

$$A^r \times A^s = A^s \times A^r = A^{r+s}$$

for all integer values of r and s (positive, negative, or zero) provided that A^0 is interpreted to be I .

(c) The transpose of a matrix.

The matrix whose r th column ($r = 1, 2, \dots$) is the r th row of A is called the TRANSPOSE of A and is commonly denoted by A' . Thus

$$A = [a_{ik}] \text{ gives } A' = [a_{ki}].$$

When $A = A'$, that is when $a_{ik} = a_{ki}$, the matrix A is said to be SYMMETRICAL.

(d) Functions of a matrix.

When a_0, a_1, \dots, a_p are numbers and

$$f(x) \equiv a_0 + a_1 x + \dots + a_p x^p$$

is a polynomial in a single variable x , the matrix-sum†

$$a_0 I + a_1 A + \dots + a_p A^p$$

is a single matrix that is conveniently denoted by $f(A)$. If this matrix is non-singular, it has a reciprocal and this is conveniently denoted by $1/f(A)$.

The product of this by $g(A)$, where $g(x)$ is another polynomial in x , yields a matrix that is conveniently denoted by $g(A)/f(A)$. The resulting matrix is said to be a rational function of A .

5. The law of reversal for transposes and reciprocals

$$\text{Let } A = [a_{ik}], \quad B = [b_{ik}].$$

$$\text{Then } AB = [a_{ij} b_{jk}], \quad (AB)' = [a_{kj} b_{ji}].$$

$$\begin{aligned} \text{But } B'A' &= [b_{ki}] \times [a_{ki}] \\ &= [b_{ji} a_{kj}] \\ &= [a_{kj} b_{ji}] = (AB)'. \end{aligned}$$

† The matrix $a_1 A$, where a_1 is an ordinary number, is defined to be the matrix whose elements are a_1 times the elements of A ; e.g.

$$2 \begin{bmatrix} 1 & 4 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 8 \\ 6 & 10 \end{bmatrix}.$$

Hence, *the transpose of a product is the product of the transposes taken in the reverse order.*

By extension, $(ABC)' = C'B'A'$

and $(AB...K)' = K'...B'A'.$

Again

$$(B^{-1}A^{-1}) \times (AB) = B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I$$

and, similarly, $(AB) \times (B^{-1}A^{-1}) = I.$

Accordingly, *the reciprocal of a product is the product of the reciprocals taken in the reverse order.*

By extension, $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$

and $(AB...K)^{-1} = K^{-1}...B^{-1}A^{-1}.$

6. Simple matrix equations

(a) When A , B are given square matrices of order n and B is non-singular, the matrix equations

$$A = BX, \quad A = YB$$

in the 'unknowns' X and Y have the unique solutions

$$X = B^{-1}A, \quad Y = AB^{-1}.$$

That these are solutions follows at once from the fact that

$$BB^{-1} = B^{-1}B = I.$$

Moreover, each solution is unique: if, for example, BR is also equal to A , then $B(R-X) = 0$ and, since B is non-singular,

$$R-X = 0.$$

(b) The unique solution of the matrix equation

$$Ax = b,$$

where A is a non-singular square matrix of order n and x , b are single-column matrices of n rows, is $x = A^{-1}b$. It provides the solution of the n linear equations

$$a_{ik}x_k = b_i \quad (i = 1, 2, \dots, n).$$

Similarly, the matrix equation

$$y'A = b'$$

has the solution $\mathbf{y}' = \mathbf{b}'\mathbf{A}^{-1}$ and provides the solution of the n equations

$$y_k a_{ki} = b_i \quad (i = 1, 2, \dots, n).$$

7. Submatrices

A matrix
$$P \equiv \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

can be denoted, on introducing symbols P_1, P_2, P_3, P_4 , where

$$P_1 \equiv \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad P_2 \equiv \begin{bmatrix} p_{13} \\ p_{23} \end{bmatrix},$$

$$P_3 \equiv [p_{31} \quad p_{32}], \quad P_4 \equiv [p_{33}],$$

by
$$\begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix}.$$

The matrices P_1, P_2, P_3, P_4 are called submatrices of P .

When a second matrix Q is divided into submatrices on the same pattern as P , a little calculation shows that

$$P+Q = \begin{bmatrix} P_1+Q_1 & P_2+Q_2 \\ P_3+Q_3 & P_4+Q_4 \end{bmatrix}$$

and
$$PQ = \begin{bmatrix} P_1 Q_1 + P_2 Q_3 & P_1 Q_2 + P_2 Q_4 \\ P_3 Q_1 + P_4 Q_3 & P_3 Q_2 + P_4 Q_4 \end{bmatrix}.$$

In this sum and product the 'elements' are themselves matrices; for example, P_1+Q_1 is a matrix of two rows and two columns, while $P_3 Q_1 + P_4 Q_3$ is a matrix of one row and two columns.

The symmetry of the previous example is not an essential feature of the process. In multiplication, for example, what is essential is that, in each $P_r Q_s$ that occurs, P_r shall have as many columns as Q_s has rows. As an illustration of a non-symmetrical arrangement (it has no other interest and is devised purely as an illustration), let

$$\begin{aligned} P_{11} &= [p_{11} \quad p_{12}], & P_{12} &= [p_{13}], \\ Q_{11} &= \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \end{bmatrix}, & Q_{12} &= \begin{bmatrix} q_{14} \\ q_{24} \end{bmatrix}, \\ Q_{21} &= [q_{31} \quad q_{32} \quad q_{33}], & Q_{22} &= [q_{34}]. \end{aligned}$$

Then

$$\begin{aligned}
 [p_{11} \quad p_{12} \quad p_{13}] &\times \left[\begin{array}{ccc|c} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ \hline q_{31} & q_{32} & q_{33} & q_{34} \end{array} \right] \\
 &= [P_{11} \quad P_{12}] \times \left[\begin{array}{cc} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{array} \right] \\
 &= [P_{11} Q_{11} + P_{12} Q_{21} \quad P_{11} Q_{12} + P_{12} Q_{22}] \\
 &= [p_{1k} q_{k1} \quad p_{1k} q_{k2} \quad p_{1k} q_{k3} \quad p_{1k} q_{k4}],
 \end{aligned}$$

the summation with regard to k being for $k = 1, 2, 3$.

8. The rank of a matrix

(a) Minors.

Let A be a matrix, not necessarily square. From it delete all rows save a certain r rows and all columns save a certain r columns. When $r > 1$ the elements that remain form a square matrix of order r and the determinant of this matrix is called a minor of A of order r . A single element of A may be considered to be a minor of order 1.

(b) Definition of rank.

A matrix has rank r (≥ 1) when r is the largest integer for which we can state that 'not ALL minors of order r are zero'.

To understand the definition we note that a minor of order $k+1$ can be expanded by its first row as a sum of multiples of minors of order k , so that if all minors of order k are zero, then all minors of order $k+1$ are zero. The converse is not true; for example, in

$$\begin{bmatrix} 1 & 2 & 4 & 7 \\ 0 & 5 & 1 & 6 \\ 1 & 0 & 2 & 3 \\ 2 & 1 & 3 & 6 \end{bmatrix}$$

the only minor of order 4 is the determinant of the matrix and its value is zero, but the minor of order 3

$$\begin{vmatrix} 1 & 2 & 4 \\ 0 & 5 & 1 \\ 1 & 0 & 2 \end{vmatrix}$$

is not equal to zero.

It is sometimes convenient to speak of a null matrix, in which every element is zero, as being of rank zero.

(c) *Linear dependence.*

Consider an array of three rows

$$\begin{array}{cccc} a_1, & b_1, & \dots, & z_1, \\ a_2, & b_2, & \dots, & z_2, \\ a_3, & b_3, & \dots, & z_3. \end{array}$$

If the three rows are related in such a way that there are numbers λ_1 and λ_2 for which

$$\rho_3 = \lambda_1 \rho_1 + \lambda_2 \rho_2 \quad (\rho = a, b, \dots, z) \quad (6)$$

we say that the third row is the SUM OF MULTIPLES (λ_1 and λ_2) of the first and second rows. When the three rows, or two of them, are related in such a way that there are numbers λ_1 , λ_2 , λ_3 , of which two at least are not zero, and for which

$$\lambda_1 \rho_1 + \lambda_2 \rho_2 + \lambda_3 \rho_3 = 0 \quad (\rho = a, b, \dots, z), \quad (7)$$

we say that the three rows are LINEARLY DEPENDENT. We say that the rows are LINEARLY INDEPENDENT if (7) is satisfied only when $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

The definitions extend to any number of rows or of columns.

(d) *Rank and linear dependence.*

The rank of a matrix is equal to the number of linearly independent rows in the matrix, as the following theorem shows.

Let A be a matrix of rank r and let a non-zero minor of A of order r have elements from the α th, β th, ..., κ th rows of A (r rows in all). Let A have a further row, say the θ th. Then† there are numbers $\lambda_\alpha, \lambda_\beta, \dots, \lambda_\kappa$ for which

$$\rho_\theta = \lambda_\alpha \rho_\alpha + \lambda_\beta \rho_\beta + \dots + \lambda_\kappa \rho_\kappa,$$

so that the θ th row is the sum of multiples of the α th, β th, ..., κ th rows.

Thus we can, when the number of rows of A exceeds its rank r , select r rows of A and express every other row as a sum of multiples of the r selected rows. Moreover, it is not possible to

† This and other properties noted in this section are proved in Ferrar, *Algebra* (Oxford, 1941), at chapter viii. Further references to this book will be indicated by F. and the appropriate page number.

select q rows of A , where $q < r$, and then express every other row as a sum of multiples of the q selected rows.

There are similar results for columns.

9. Linear equations

Consider the m linear equations

$$\sum_{k=1}^n a_{ik} x_k = b_i \quad (i = 1, \dots, m) \quad (8)$$

in the n unknowns x_1, \dots, x_n . Let the matrices A, B be given by

$$A = \begin{bmatrix} a_{11} & \cdot & \cdot & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & \cdot & \cdot & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & \cdot & \cdot & a_{1n} & b_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & \cdot & \cdot & a_{mn} & b_m \end{bmatrix};$$

let A be of rank r and B of rank r' . Then, from the nature of definition of rank, $r \leq r'$; moreover

When $r = r'$ the equations (8) are consistent (that is, there is at least one set of values of the unknowns that satisfies all the equations) and when $r < r'$ the equations (8) are not consistent.

When all the b_i in (8) are zero, we are concerned with what are known as HOMOGENEOUS LINEAR EQUATIONS, namely

$$\sum_{k=1}^n a_{ik} x_k = 0 \quad (i = 1, \dots, m). \quad (9)$$

The ranks r and r' of the above discussion are now necessarily equal and the equations are always consistent. On the other hand, the equations are always satisfied by

$$x_1 = x_2 = \dots = x_n = 0$$

and this may be the only solution. The following theorem† summarizes the more important results about such a set of equations.

Let A , in (9), be of rank r . Then $r \leq n$, since A has n columns.

(i) *When $r = n$, the equations (9) have no solution other than*

$$x_1 = x_2 = \dots = x_n = 0.$$

† For proofs and further details, see F. 98-105.

(ii) When $r = n-1$, the equations have effectively only one non-zero solution† and if this is

$$\xi_1, \xi_2, \dots, \xi_n,$$

all other non-zero solutions are of the form

$$\lambda\xi_1, \lambda\xi_2, \dots, \lambda\xi_n.$$

(iii) When $r < n-1$, the equations (9) have $n-r$ linearly independent non-zero solutions and every non-zero solution can be expressed as a sum of multiples of these $n-r$ solutions.

10. The rank of a product of two matrices

We note two important theorems of frequent application.

(a) The rank of a product AB cannot exceed the rank of either factor.

(b) When B is a non-singular square matrix of the same order as the square matrix A , the matrices

$$A, AB, BA$$

all have the same rank.

The proof of these theorems follows fairly directly from the following result,‡ one that is often used apart from its immediate connexion with the ranks of matrices.

(c) Let A have n_1 rows and n columns and let B have n rows and n_2 columns; then AB has n_1 rows and n_2 columns. Every minor of AB of order greater than n , if there are such minors, is equal to zero; and every minor of AB of order $t \leq n$ is either the product of a t -rowed minor of A by a t -rowed minor of B or is the sum of a number of such products.

This contains as a special case the more elementary result

When A and B are square matrices of the same order, the determinant $|AB| = |A| \times |B|$. The proof of this follows at once from (2) of § 3. Its extension to a product of three or more matrices,

$$|ABC \dots K| = |A| \cdot |B| \cdot |C| \cdot \dots \cdot |K|,$$

is also a direct consequence of the definition of a product of matrices.

† 'Non-zero' because at least one ξ is different from zero.

‡ F. 109.