

we see that any term which does not correspond to an absolute permutation of the n indices i, j, \dots, k contains an element from the main diagonal and is therefore zero. The value of a term corresponding to an absolute permutation is $+1$ or -1 according as the permutation is even or odd. The value of D is therefore the difference between the number of even absolute permutations, N_e , and the number of odd absolute permutations, N_o , that is

$$(1) \quad N_e - N_o = (-1)^{n-1}(n-1).$$

Moreover we have*

$$(2) \quad N_e + N_o = n! \sum_{r=2}^n \frac{(-1)^r}{r!}.$$

From (1) and (2) N_e and N_o can be calculated.

PRINCETON UNIVERSITY

THE ABSOLUTE VALUE OF THE PRODUCT OF TWO MATRICES†

BY J. H. M. WEDDERBURN

1. *Introduction.* If $a = (a_{pq})$ is a matrix of order n whose elements are ordinary complex numbers, the absolute value of a is defined as $\sqrt{\sum a_{pq} \bar{a}_{pq}}$, where $\bar{a} = (\bar{a}_{pq})$ is the matrix whose coefficients are the conjugates of the corresponding coefficients in a ; we shall denote it here by $|a|$, a special symbol being convenient since the absolute value of a scalar matrix λ is not $|\lambda| \equiv \text{mod } \lambda$ but $n^{1/2}|\lambda|$. This definition has been freely used by writers on differential equations; but, in spite of this, its properties with regard to multiplication have seemingly escaped notice, or are at least not well known.

* Cf. Seelhoff, ARCHIV DER MATHEMATIK UND PHYSIK, (2), vol. 1, p. 100.

† Presented to the Society, May 2, 1925.

The principal properties of $[a]$ are as follows:

- $$\begin{aligned} (1) \quad & [a + b] \leq [a] + [b], \\ (2) \quad & [\lambda] = n^{1/2} |\lambda|, \\ (3) \quad & [\lambda a] = |\lambda| [a], \\ (4) \quad & [ab] \leq [a][b]. \end{aligned} \quad \left. \vphantom{\begin{aligned} (1) \\ (2) \\ (3) \end{aligned}} \right\} (\lambda \text{ scalar})$$

Relation (4), which is the only one which is not obvious, follows immediately from the identity

$$\begin{aligned} [a]^2 [b]^2 &= \sum_{pr} a_{pr} \bar{a}_{pr} \sum_{qs} b_{sq} \bar{b}_{sq} \\ &= \sum_{pq} \left\{ \frac{1}{2} \sum_{rs} (a_{pr} \bar{b}_{sq} - a_{ps} \bar{b}_{rq}) (\bar{a}_{pr} b_{sq} - \bar{a}_{ps} b_{rq}) + \sum_r a_{pr} b_{rq} \sum_s \bar{a}_{ps} \bar{b}_{sq} \right\}, \end{aligned}$$

whence, seeing that the terms of the first summation within the braces are real and positive, we have

$$[a]^2 [b]^2 \geq \sum_{pq} \sum_r a_{pr} b_{rq} \sum_s \bar{a}_{ps} \bar{b}_{sq} = [ab]^2.$$

The equality sign holds in (4) only when a_{rs} and b_{rs} have the form $a_{rs} = \lambda_r \mu_s$, $b_{rs} = \bar{\mu}_r \nu_s$.

Similar matrices do not in general have the same absolute value; $[bab^{-1}]$ and $[a]$ are the same however if $b = uc$, where $ca = ac$ and u is a unitary matrix.*

2. *Negative Powers of a Matrix.* As pointed out by Schur, $[a]^2$ may be defined as the trace of $\bar{a}'a$, which is a positive Hermitian matrix and therefore has positive real roots, say g_1, g_2, \dots, g_n . If

$$s_r = \sum g_1^r, \quad p_r = \sum g_1 g_2 \cdots g_r, \quad s \equiv s_1 = [a]^2, \quad p \equiv p_n = [\det a]^2,$$

* Cf. Schur, *Über die charakteristischen Wurzeln einer linearen Substitution*, MATHEMATISCHE ANNALEN, vol. 66 (1909), pp. 488-510, pp. 490, 492. Since replacing a matrix by a similar one corresponds to a change of coordinates when the matrix is regarded as a linear transformation, it follows from Schur's work that, when the elementary divisors of a are simple, $[a]$ has its minimum value when a is represented as a diagonal matrix, that is, in its normal form; and it seems probable that the normal form also gives the minimum value even if the elementary divisors are not simple.

then it follows readily from the inequality connecting the arithmetic and geometric means that*

$$r^r p_r \leq \sum (g_1 + g_2 + \cdots + g_r)^r \leq \binom{n-1}{r-1} s^r,$$

$$p_r \geq \binom{n}{r} p^{\binom{n-1}{r-1} / \binom{n}{r}} = \binom{n}{r} p^{r/n},$$

whence

$$(5) \quad \binom{n}{r} p^{r/n} \leq p_r \leq r^{-r} \binom{n-1}{r-1} s^r.$$

If $C_r(a)$ is the r th compound of a , † $p_r = [C_r(a)]^2$; hence, if we put α for the determinant of a , (5) becomes

$$(6) \quad \binom{n}{r} |\alpha|^{2r/n} \leq [C_r(a)]^2 \leq r^{-r} \binom{n-1}{r-1} |\alpha|^{2r}.$$

Now $a^{-1} = C_{n-1}(a)/\alpha$, provided $\alpha \neq 0$; hence

$$(7) \quad n^{1/2} |\alpha|^{-1/n} \leq |a^{-1}| \leq (n-1)^{-\frac{(n-4)}{2}} |a|^{n-1} / |\alpha|.$$

This inequality enables us to deal with expressions containing negative powers of a matrix. The following inequality is, however, sometimes more convenient. Since $aa^{-1} = 1$, it follows from (4) that

$$(8) \quad |a^{-1}| \leq \frac{n^{1/2}}{|a|},$$

and, since ‡ $a = e^{\log a}$, $a^{-1} = e^{-\log a}$, both $|a|$ and $|a^{-1}|$ are less than or equal to $e^{|\log a|}$.

* For $r = n$ this inequality, which is an extension of Hadamard's expression for the maximum of the absolute value of a determinant, is given (loc. cit., p. 496) by Schur, who also derives a number of other interesting inequalities.

† It must be remembered here that the order of $C_r(a)$ is $\binom{n}{r}$, so that, if for example $a = 1$, then $[C_r(a)]^2 = \binom{n}{r}$ and not n . References to the literature of compound matrices will be found in Pascal, *REPERTORIUM DER HÖHEREN MATHEMATIK*, vol. 1, pp. 138-146.

‡ An elementary discussion of the logarithm of a matrix is given by the author, *The automorphic transformation of a bilinear form*, *ANNALS OF MATHEMATICS*, (2), vol. 23 (1921), pp. 122-125.

3. *Infinite Series and Products of Matrices.* The use of (4) makes the discussion of infinite series and products of matrices somewhat more compact than is the case with the usual methods.* We shall give here only a few elementary illustrations.

If a_r ($r = 0, 1, 2, \dots$) is an infinite sequence of matrices and x a variable matrix, and we set $[a_r] = \alpha_r$, $[x] = \xi$, then $\sum a_r x^r$ converges absolutely if $\sum [a_r x^r]$ converges, and hence also when $\sum \alpha_r \xi^r$ is convergent. In particular, $e^x = \sum_0^\infty x^r / r!$ converges for every matrix x and $[e^x] \leq e^{[x]}$. Similarly $\log(1+x) = \sum_1^\infty (-1)^{r-1} x^r / r$ converges for $[x] < 1$. Here, however, the condition is much too strong; for, if $x = 1/3$, say, then $[x] = n^{1/2}/3$, which is greater than 1 if $n > 9$, while the series converges absolutely in the ordinary sense. A better form of the condition in this and similar cases is $[xc]/[c] < 1$, where c is some non-singular matrix. Any criterion of this sort has distinct limitations; for it is not difficult to show that all that is necessary for the convergence of $\sum a_r x^r$ is that the series converge when x is replaced by the ordinary absolute value of the root of x of greatest absolute value.

4. *Examples.* The following examples† illustrate the application to infinite products. Let c be an arbitrary matrix, and set

$$(9) \quad \begin{aligned} P_m &= (1 + a_1)(1 + a_2) \cdots (1 + a_m), \\ Q_m &= (1 + [a_1])(1 + [a_2]) \cdots (1 + [a_m]); \end{aligned}$$

then

$$(10) \quad [P_m - 1] \leq Q_m - 1 < e^{\sum [a_r]} - 1,$$

* For applications to differential equations, for example, reference may be made to Schlesinger, *Vorlesungen über lineare Differentialgleichungen*, Leipzig, 1908, pp. 86-90, 41-42.

† The inequalities given here are those required in proving the existence of a solution of a system of linear differential equations by the method of successive approximations; cf. Schlesinger, loc. cit., pp. 3-16.

$$(11) \quad [P_m c] \leq e^{\sum [a_r]} [c],$$

$$(12) \quad [P_m - P_k] \leq Q_m - Q_k \leq \sum_k^m [a_r] e^{\sum [a_r]}.$$

It is sufficient here to give the proofs of (10) and (11). Expanding (9), we have

$$P_m = 1 + \sum_p a_p + \sum_{p < q} a_p a_q + \sum_{p < q < r} a_p a_q a_r + \dots,$$

therefore

$$[P_m - 1] \leq \sum_p [a_p] + \sum_{p < q} [a_p][a_q] + \dots = Q_m - 1,$$

$$[P_m c] \leq (1 + \sum_p [a_p] + \sum_{p < q} [a_p][a_q] + \dots) [c] = Q_m [c].$$

As another example we may take the proof of the existence of a solution of the differential equation

$$(13) \quad \frac{dx}{dt} = ax,$$

where $a = a_0 + a_1 t + a_2 t^2 + \dots$ is a matrix which is analytic in the scalar variable t at $t = 0$. Let C be the circle of convergence of the series for a , that is, if $[a_r] = \alpha_r$; then $\alpha(t) = \sum \alpha_r t^r$ converges in C ; and set $x = \sum x_r t^r$ in (13). Comparing like powers of t we get

$$s x_s = a_0 x_{s-1} + a_1 x_{s-2} + \dots + a_{s-1} x_0,$$

so that $s [x_s] \leq \alpha_0 [x_{s-1}] + \alpha_1 [x_{s-2}] + \dots + \alpha^{s-1} [x_0]$. Let

$$\xi_0 = [x_0], \quad \xi(t) = \xi_0 e^{\int_0^t \alpha(t) dt} = \sum \xi_r t^r, \quad \text{then} \quad d\xi/dt = \alpha(t) \xi,$$

so that $s \xi_s = \alpha_0 \xi_{s-1} + \alpha_1 \xi_{s-2} + \dots + \alpha_{s-1} \xi_0$. It follows that $[x_s] \leq \xi_s$; for if this is true for $s = 1, 2, \dots, i-1$, then

$$i [x_i] \leq \sum_j \alpha_j [x_{i-j-1}] \leq \sum_j \alpha_j \xi_{i-j-1} = i \xi_i.$$

The series for x therefore also converges in C and $x = x_0$ when $t = 0$.