

ON A MATRIX DIFFERENTIAL OPERATOR*

BY A. K. MITCHELL

1. *Introduction.* H. W. Turnbull† has defined a matrix differential operator and has given several theorems concerning its properties. The first of these theorems is proved by induction, and for the fourth, use is made of a lemma concerning the principal minors of a determinant for which he gives a long and intricate proof. It is the purpose of this note to give a simple and direct proof of this lemma and of the fourth theorem, and to indicate how each of the first three theorems proved by Turnbull may be obtained directly. In the last paragraph will be given, in addition, the results of applying the operator successively to the coefficients of the characteristic equation of a matrix.

2. *The Matrix Differential Operator.* Let

$$E = \begin{vmatrix} E_1^1 & E_2 & \cdots & E_n^1 \\ E_1^2 & E_2^2 & \cdots & E_n^2 \\ \cdots & \cdots & \cdots & \cdots \\ E_1^n & E_2^n & \cdots & E_n^n \end{vmatrix}$$

be an n -rowed square matrix whose n^2 elements are treated as independent variables. This matrix is characterized by a typical element, and we shall write $E = E_s^r$, where r denotes the row and s the column. In this way if we have another matrix $F = F_s^r$, we shall write,‡ from the law of multiplication of matrices,

$$EF = E_{\alpha}^r F_s^{\alpha}.$$

We observe that

$$EF = E_{\alpha}^r F_s^{\alpha} = F_s^{\alpha} E_{\alpha}^r,$$

whereas

$$FE = F_{\alpha}^r E_s^{\alpha} = E_s^{\alpha} F_{\alpha}^r \neq EF.$$

* Presented to the Society, March 26, 1932.

† H. W. Turnbull, *On differentiating a matrix*, Proceedings of the Edinburgh Mathematical Society, (2), vol. 1, Part 2 (1928), pp. 111-128.

‡ Throughout this paper repeated Greek suffixes will denote summation from 1 to n .

The matrix differential operator Ω is defined as follows :

$$\Omega = \left\| \begin{array}{cccc} \partial/\partial E_1^1 & \partial/\partial E_1^2 & \dots & \partial/\partial E_1^n \\ \partial/\partial E_2^1 & \partial/\partial E_2^2 & \dots & \partial/\partial E_2^n \\ \dots & \dots & \dots & \dots \\ \partial/\partial E_n^1 & \partial/\partial E_n^2 & \dots & \partial/\partial E_n^n \end{array} \right\|.$$

We note that the order of suffixes is transposed from that of E . The effect of the operator Ω is defined by the ordinary law of multiplication of matrices; thus $\Omega F = \partial F_s^\alpha / \partial E_r^\alpha$. For a scalar function f we have $\Omega f = \partial f / \partial E_r^s$.

3. *Proof of Turnbull's Theorems.* Now let I_s denote the sum of all the s -rowed principal minors of the determinant of the matrix E . The fourth theorem given by Turnbull, which we shall prove directly, is then as follows.

THEOREM.

$$\begin{aligned} \delta_s^r - \Omega I_1 &= 0, & E - \delta_s^r I_1 + \Omega I_2 &= 0, \\ \dots & \dots & \dots & \dots \\ E^{n-1} - I_1 E^{n-2} + \dots + (-1)^{n-1} \delta_s^r I_{n-1} + (-1)^n \Omega I_n &= 0. \end{aligned}$$

The proof will depend upon the properties of the generalized Kronecker delta* $\delta_{s_1^1 \dots s_k^k}^{r_1^1 \dots r_k^k}$ which, if the superscripts are distinct from each other and the subscripts are the same set of numbers as the superscripts, has the value +1 or -1 according as an even or an odd permutation is required to arrange the superscripts in the same order as the subscripts; and which, in all other cases, has the value zero. In particular $\delta_s^r = 0$ if $r \neq s$, $\delta_s^r = 1$ if $r = s$. Then, I_s being the sum of all the s -rowed principal minors of E , we may write:

$$\begin{aligned} I_1 &= \delta_\beta^\alpha E_\alpha^\beta, \\ I_2 &= \frac{1}{2!} \delta_{\beta_1 \beta_2}^{\alpha_1 \alpha_2} E_{\alpha_1}^{\beta_1} E_{\alpha_2}^{\beta_2}, \\ \dots & \dots \dots \dots \dots \dots \\ I_s &= \frac{1}{s!} \delta_{\beta_1 \dots \beta_s}^{\alpha_1 \dots \alpha_s} E_{\alpha_1}^{\beta_1} \dots E_{\alpha_s}^{\beta_s}, \end{aligned}$$

* See F. D. Murnaghan, *American Mathematical Monthly*, vol. 32 (1925), p. 233; and O. Veblen, *Invariants of Quadratic Differential Forms*, p. 3.

from which we have immediately

$$\begin{aligned} \Omega I_1 &= \partial \delta_s^\alpha E_\alpha^\beta / \partial E_r^s = \delta_s^r, \\ \Omega I_2 &= \delta_s^r I_1 - E_s^r. \end{aligned}$$

Continuing in this way we obtain by an easy calculation*

$$\begin{aligned} \Omega I_n &= \delta_s^r I_{n-1} - I_{n-2} E_s^r + I_{n-3} E_{\alpha_1}^r E_s^{\alpha_1} + \dots + (-1)^{t+1} I_{n-t} E_{\alpha_1}^r \\ &\quad \dots E_s^{\alpha_{t-2}} + \dots + (-1)^{n+1} E_{\alpha_1}^r E_{\alpha_2}^{\alpha_1} \dots E_s^{\alpha_{n-2}}. \end{aligned}$$

Transposing the left members of these equations we have the theorem stated above.

From the foregoing we see that

$$\begin{aligned} \partial I_{p-1} / \partial E_\alpha^s &= \delta_s^\alpha I_{p-2} - I_{p-3} E_s^\alpha + \dots + (-1)^{t-1} I_{p-t} E_{\alpha_1}^\alpha E_{\alpha_2}^{\alpha_1} \dots E_s^{\alpha_{t-2}} \\ &\quad + \dots + (-1)^p E_{\alpha_1}^\alpha E_{\alpha_2}^{\alpha_1} \dots E_s^{\alpha_{p-3}} \end{aligned}$$

and

$$\begin{aligned} \partial I_p / \partial E_r^s &= \delta_s^r I_{p-1} - I_{p-2} E_s^r + I_{p-3} E_\alpha^r E_s^\alpha + \dots + (-1)^{t-1} I_{p-t} E_{\alpha_1}^r \\ &\quad \dots E_s^{\alpha_{t-2}} + \dots + (-1)^{p+1} E_{\alpha_1}^r \dots E_s^{\alpha_{p-3}}, \end{aligned}$$

from which it is evident that

$$\partial I_p / \partial E_r^s = \delta_s^r I_{p-1} - (\partial I_{p-1} / \partial E_\alpha^s) E_\alpha^r.$$

This is the lemma used by Turnbull.

By induction Turnbull proves the following formula:

$$\Omega E^r = nE^{r-1} + s_1 E^{r-2} + \dots + s_{j-1} E^{r-j} + \dots + s_{r-1},$$

where s_r is the sum of the r th powers of the latent roots of the matrix E (that is, $s_r = \lambda_1^r + \lambda_2^r + \dots + \lambda_n^r$).

Now by a theorem due to Sylvester,† the latent roots of any rational function of a matrix E are the corresponding functions of the latent roots of E . In particular the latent roots of E^r are $\lambda_1^r, \lambda_2^r, \dots, \lambda_n^r$. From

$$\Omega E^r = \partial E_{\alpha_1}^\beta E_{\alpha_2}^{\alpha_1} \dots E_s^{\alpha_{r-1}} / \partial E_r^s,$$

by differentiation, we obtain

* See *The derivation of tensors from tensor functions*, American Journal of Mathematics, vol. 53 (1931), p. 198.

†Sylvester, *Mathematical Papers*, vol. 4, p. 133.

$$\begin{aligned} \Omega E^r &= \delta_\beta^\beta \delta_{\alpha_1}^r E_{\alpha_2}^{\alpha_1} \cdots E_s^{\alpha_{r-1}} + E_{\alpha_1}^\beta \delta_\beta^{\alpha_1, r} \delta_{\alpha_2} E_{\alpha_3}^{\alpha_2} \cdots E_s^{\alpha_{r-1}} \\ &+ \cdots + E_{\alpha_1}^\beta E_{\alpha_2}^{\alpha_1} \cdots E_{\alpha_j}^{\alpha_{j-1}} \delta_\beta^{\alpha_j} \delta_{\alpha_{j+1}}^r E_{\alpha_{j+2}}^{\alpha_{j+1}} \cdots E_s^{\alpha_{r-1}} \\ &+ \cdots + E_{\alpha_1}^\beta E_{\alpha_2}^{\alpha_1} \cdots E_\beta^{\alpha_{r-2}} \delta_s^r. \end{aligned}$$

But the sum of the latent roots of a matrix is equal to the sum of the elements of the leading diagonal of the matrix, so that we have by Sylvester's theorem, $s_1 = E_\alpha^\alpha$, $s_2 = E_\alpha^\alpha$, $E_{\alpha_1}^{\alpha_1}, \dots$, $s^r = E_{\alpha_1}^{\alpha_1} E_{\alpha_2}^{\alpha_2} \cdots E_{\alpha_{r-1}}^{\alpha_{r-1}}$. Hence, from the above expression for ΩE^r ,

$$\Omega E^r = nE^{r-1} + s_1E^{r-2} + \cdots + s_{j-1}E^{r-j} + \cdots + s_{r-1}.$$

In like manner each of the Theorems I and III of Turnbull may be proved directly.

4. *Conclusion.* Let us now consider the effect of repeating the application of the operator Ω on the functions I_1, I_2, \dots, I_n defined above. The I 's so defined are the coefficients of the characteristic equation* of the matrix E . From the definition of I_p and of Ω it is readily seen that

$$\begin{aligned} \Omega I_p &= [1/(p-1)!] \delta_s^{\alpha_2 \cdots \alpha_p} E_{\alpha_2}^{\beta_2} \cdots E_{\alpha_p}^{\beta_p}, \\ \Omega^2 I_p &= [1/(p-1)!] (\partial/\partial E_r^\alpha) \delta_s^{\alpha_2 \cdots \alpha_p} E_{\alpha_2}^{\beta_2} \cdots E_{\alpha_p}^{\beta_p} \\ &= - [(p-1)/(p-1)!] \delta_s^{\alpha_3 \cdots \alpha_p} E_{\alpha_3}^{\beta_3} \cdots E_{\alpha_p}^{\beta_p} \\ &= - [(n-p+1)/(n-p)!] \Omega I_{p-1}, \end{aligned}$$

and

$$\Omega^2 I_{p-1} = - (n-p+2) \Omega I_{p-2}.$$

Therefore

$$\begin{aligned} \Omega^3 I_p &= [(n-p+2)/(n-p)!] \Omega I_{p-2}, \\ &\dots \dots \dots \\ \Omega^p I_p &= (-1)^{p+1} [(n-1)/(n-p)!] \Omega I_1 \\ &= (-1)^{p+1} [(n-1)/(n-p)!] \delta_s^r. \end{aligned}$$

When $n = p$ this becomes

$$\Omega^n I_n = (-1)^{n+1} (n-1)! \delta_s^r.$$

Thus if the operator Ω is applied n times to the n -rowed determinant $|E_s^r|$ we obtain $(-1)^{n+1} (n-1)!$ times the unit matrix.

TRINITY COLLEGE

* See *A note on the characteristic equation of a matrix*, American Mathematical Monthly, vol. 38 (1931), p. 386.