ON THE CHARACTERISTIC POLYNOMIAL OF THE
PRODUCT OF SEVERAL MATRICES

WILLIAM E. ROTH

We shall prove two theorems.

**Theorem I.** If $A$ is an $n \times n$ matrix with elements in the field $F$, if $R$ and $S_i$, $i = 1, 2, \ldots, r$, are $1 \times n$ matrices with elements in $F$, and $D_i = R^T S_i$, where $R^T$ is the transpose of $R$, and if the characteristic polynomial of $A_i = A + D_i$ is

$$|xI - A_i| = m_{i0} + m_{i1}x + m_{i2}x^2 + \cdots + m_{i,r-1}x^{r-1},$$

where $m_{i,j-1}$, $i, j = 1, 2, \ldots, r$, are polynomials in $x^r$ with coefficients in $F$, then the characteristic polynomial of the product $P = A_1 A_2 \cdots A_r$ is given by $(-1)^{(r-1)n}|\Delta(x)|$, where

$$\Delta(x^r) = \begin{bmatrix}
m_{10} & m_{1,r-1}x^{r-1} & m_{1,r-2}x^{r-2} & \cdots & m_{11}x \\
m_{21}x & m_{20} & m_{2,r-1}x^{r-1} & \cdots & m_{22}x^2 \\
m_{32}x^2 & m_{31}x & m_{30} & \cdots & m_{33}x^3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
&m_{r,r-1}x^{r-1} & m_{r,r-2}x^{r-2} & m_{r,r-3}x^{r-3} & \cdots & m_{r0}
\end{bmatrix}.$$

This proposition has been proved by the writer [1] for the case $r = 2$. Recently Parker [2] generalized that result and Goddard [3] gave an alternate proof of it and extended his method to the product of three matrices. This latter result does not come under the theorem above. Schneider [5] proved the theorem for the case $A_i A_j = 0$, $i < j$, $i, j = 1, 2, \ldots, r$.

Capital letters and expressions in bold faced parentheses will indicate matrices with elements in the field $F$ or in $F(\omega)$, the extension of $F$ by the adjunction of $\omega$ a primitive $r$th root of unity to it, and in $F(x)$ the polynomial domain of $F(\omega)$. The direct product of $B$ and $C$ is $(b_i c) = B(C)$. The product indicated by $\Pi$ will run from 1 to $r$.

If $R$ is not zero a nonsingular matrix $Q$ with elements in $F$ exists such that $QR^T = (1, 0, \cdots, 0)^T$; as a result

$$QD_iQ^T = (1, 0, \cdots, 0)^TS_iQ^T = E_i,$$

where $Q^t$ is the inverse of $Q$ and $E_i$ has nonzero elements in only the first row. Now let

Presented to the Society, April 23, 1955; received by the editors April 4, 1955 and, in revised form, August 26, 1955.

578
\[ QA_k Q' = M_k = Q(A + D_k)Q' = M + E_k, \]

where \( Q \) is the matrix defined above and \( QAQ' = M = (m_{ij}) \). Consequently \( M_k = (m_{ij}^{(k)}) \), where \( m_{ij}^{(k)} = m_{ij} + e_{ij}^{(k)} \) and \( m_{ij}^{(k)} = m_{ij} \) for \( i > 1 \). That is, the matrices \( M_k \) differ only in the elements of their first rows. As a result the elements of the first columns of the adjoints \([xI - M_k]^A\) and \([xI - M]^A\) of \( xI - M_k \) and \( xI - M \) respectively are identical for \( k = 1, 2, \ldots, r \), since all these matrices agree in the elements of their last \( n - 1 \) rows and for the same reason

\[
N_k(x) = [xI - M_k][xI - M]^A,
\]

\[
\begin{bmatrix}
m_k(x) & * & * & \cdots & * \\
0 & m(x) & 0 & \cdots & 0 \\
0 & 0 & m(x) & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & m(x)
\end{bmatrix},
\]

\[ k = 1, 2, \ldots, r, \] where asterisks indicate nonzero elements in \( F(x) \) and \( |xI - M| = m(x) \).

Let \( W = (\omega_{ij}) = (\omega_{(i-1)(j-1)}), \quad i, j = 1, 2, \ldots, r; \) then

\[
|W(I_k)| = |W|^k,
\]

where \( I_k \) is the identity matrix of order \( k \). The determinantal equation holds because \( W(I_k) \) can be transformed by the interchange of rows and corresponding columns to the direct sum \( W + W + \cdots + W \) of \( k \) summands.

We shall operate in \( F(x) \) on the matrix

\[
M(x) = (\delta_{ij}x - \delta_{i+1,j}M_i) \quad (\delta_{r+1,1} = 1)
\]

\[
= \begin{bmatrix}
xI & -M_1 & 0 & \cdots & 0 \\
0 & xI & -M_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -M_r \\
-M_r & 0 & 0 & \cdots & xI
\end{bmatrix}
\]

If we multiply this matrix on the right by one whose first row is \( I, M_1x^{-1}, M_1M_2x^{-2}, \ldots, M_1M_2 \cdots M_{r-1}x^{-r+1} \), whose second row is \( 0, I, M_2x^{-1}, \ldots, M_2M_3 \cdots M_{r-1}x^{-r+2} \), and whose last row is
0, 0, 0, \ldots, I, we find that the determinant of the product is 
\( |x^rI - M_rM_{r-1}| \) and is therefore equal to \( |x^r - P| \). The 
proof of Theorem I will consist in showing that 
\[
(5) \quad |M(x)| = (-1)^{(r-1)n} |\Delta(x^r)|.
\]

We now proceed to establish this equation.
\[
M(x)W(I) = (\delta_{ij}xI - \delta_{i1}xI)(\omega^{(k-1)(1-\rho)}I),
\]
\[
= (\omega^{(i-1)(1-\rho)}xI - \omega^{i(1-\rho)}M_i),
\]
\[
= (\omega^{1-i}\{\omega^{i-1}xI - M_i\}).
\]
The number \( \omega^{1-i} \) is a common multiplier of the \( n \times n \) matrices in the 
jth column of the \( nr \times nr \) matrix in right member above. Consequently 
the determinant of this matrix has the factor \( \pi\omega^{(1-k)n} = \omega^{-r(r-1)n/2} \) 
\( = \omega^{r(r-1)/2} = (-1)^{(r-1)n} \). The determinantal equation obtained from 
the matric equation above is as a result:
\[
(6) \quad |M(x)| \cdot |W|^n = (-1)^{(r-1)n} |\omega_{ij}[\omega^{i-1}xI - M_i]|.
\]

According to (3) the product
\[
(7) \quad (\omega_{ik}[\omega^{k-1}xI - M_i])(\delta_{kj}[\omega^{j-1}xI - M_j]) = (\omega_{ij}N_i(\omega^{j-1}x)).
\]
The \( nr \times nr \) matrix of the right member of this equation can be transformed 
by the interchange of corresponding rows and columns to a 
similar one having the form
\[
\begin{pmatrix}
(\omega_{ij}m_i(\omega^{i-1}x)), & * & \cdots & * \\
0, & (\omega_{ij}m(\omega^{i-1}x)), & \cdots, & 0 \\
\cdots & \cdots & \cdots & \cdots \\
0, & 0 & \cdots & (\omega_{ij}m(\omega^{i-1}x))
\end{pmatrix},
\]
where asterisks represent \( r \times r \) matrices with elements in \( F(x) \) and 
the zeros are \( r \times r \) zero matrices. The determinant of this matrix is
\[
(8) \quad |W|^n[\prod m(\omega^{i-1}x)]^{n-1} |(\omega_{ij}m_i(\omega^{i-1}x))|
\]
for each of the matrices \( (\omega_{ij}m(\omega^{i-1}x)) \) has \( m(\omega^{i-1}x) \) as a divisor of all 
elements in the jth column. The determinant of the direct sum 
\( (\delta_{ij}[\omega^{j-1}xI - M_j]) \) in equation (7) is \( [\prod m(\omega^{j-1}x)]^{n-1} \); consequently 
the determinantal equation which follows from (7) and (8) is
\[
\begin{align*}
|\omega_{ij}[\omega^{j-1}xI - M_i]| & \cdot [\prod m(\omega^{i-1}x)]^{n-1} \\
& = |W|^n[\prod m(\omega^{j-1}x)]^{n-1} |(\omega_{ij}m_i(\omega^{j-1}x))|,
\end{align*}
\]
or
\[
(9) \quad |\omega_{ij}[\omega^{i-1}xI - M_i]| = |W|^{n-1} |(\omega_{ij}m_i(\omega^{i-1}x))|,
\]
where the determinant of the left member is that of an \( nr \times nr \) matrix and those in the right members are of order \( r \). From (6) and (9) we have 
\[
\frac{|M(x)|}{|W|} = (-1)^{(r-1)n} |\omega_{ij} M_i(\omega^{j-1}x)|
\]
where the determinant of the left member is that of an \( nr \times nr \) matrix and those in the right members are of order \( r \). From (6) and (9) we have 
\[
|M(x)| = W(-1)^{(r-1)n} |\omega_{ij} M_i(\omega^{j-1}x)|
\]
From (6) and (9) we have 
\[
|M(x)| = W(-1)^{(r-1)n} |\omega_{ij} M_i(\omega^{j-1}x)|
\]
It remains to be shown that the right member here is 
\[
(-1)^{(r-1)n} \Delta(x^r)
\]
this is easily accomplished by multiplying \( \Delta(x^r) \) in (1) on the right by \( W \). Here-with equation (5) is established and the proof of Theorem I is completed.

**Corollary.** Under the hypotheses of Theorem I and if \( B \) is an \( n \times n \) matrix with elements in \( F \) and if \( B_i = B + S_i^T R \) and
\[
| xI - B_i | = n_{i0} + n_{i1}x + n_{i2}x^2 + \cdots + n_{i,r-1}x^{r-1};
\]
then the characteristic polynomial of \( B_1 B_2 \cdots B_r \) is given by
\[
(10) \quad | xI - A | = m_0 + m_1x + m_2x^2 + \cdots + m_{r-1}x^{r-1},
\]
where \( m_{i-1}, i = 1, 2, \cdots, r \), are polynomials in \( x^r \) with coefficients in \( F \), and the characteristic polynomial of the product \( P = A_1 A_2 \cdots A_r \) is 
\[
\Delta(x^r) = \begin{pmatrix}
  \omega_{r,0} & \omega_{r-1,1}x^{r-1} & \cdots & \omega_{r,1}x
  \\
  \omega_{r-1,1}x & \omega_{r-1,0} & \cdots & \omega_{r-1,2}x^2
  \\
  \cdots & \cdots & \cdots & \cdots
  \\
  \omega_{1,1}x^{r-1} & \omega_{1,0} & \omega_{1,2}x^2 & \omega_{1,3}x^3 & \cdots & \cdots & \cdots & \cdots
\end{pmatrix}
\]
This case can be made to come under Theorem I for \( B_i^T = B_i^{T} + R^T S_i \) where \( B_i^T \) now satisfies the conditions imposed upon \( A_i \). Moreover 
\[
| xI - B_i | = | xI - B_i^T |.
\]
Since \( (B_1 B_2 \cdots B_r)^T = B_r^T B_{r-1}^T \cdots B_1^T \), it follows that in \( \Delta(x^r) \) of (1) we must replace the elements \( m_i - jx^{j-1} \) by \( m_{i-1} - jx^{j-1} \) in forming the matrix \( \Delta'(x^r) \) above. This proves the corollary.

**Theorem II.** If \( D_i, i = 1, 2, \cdots, r \), are \( n \times n \) matrices with elements in \( F \), each of which is nilpotent and commutative with the others and with \( A \), which also has elements in \( F \), then the characteristic polynomials of \( A_i = A + D_i, i = 1, 2, \cdots, r \), are given by
\[
| xI - A | = m_0 + m_1x + m_2x^2 + \cdots + m_{r-1}x^{r-1},
\]
where \( m_{i-1}, i = 1, 2, \cdots, r \), are polynomials in \( x^r \) with coefficients in \( F \), and the characteristic polynomial of the product \( P = A_1 A_2 \cdots A_r \) is
\[
\Delta(x^r) = \begin{pmatrix}
  m_0 & m_{r-1}x^{r-1} & m_{r-2}x^{r-2} & \cdots & m_1x
  \\
  m_1x & m_0 & m_{r-1}x^{r-1} & \cdots & m_2x^2
  \\
  m_2x^2 & m_1x & m_0 & \cdots & m_3x^3
  \\
  \cdots & \cdots & \cdots & \cdots & \cdots
  \\
  m_{r-1}x^{r-1} & m_{r-2}x^{r-2} & m_{r-3}x^{r-3} & \cdots & m_0
\end{pmatrix}
\]
According to a theorem by Frobenius [4], the determinant of the
matrix $B + C$ is equal to that of $B$ if $B$ and $C$ are commutative matrices and $C$ is nilpotent. This establishes equation (10) as giving the characteristic polynomial of $A_i$, $i = 1, 2, \ldots, r$. By Theorem I the determinant

$$|xI - A^r| = (-1)^{(r-1)n} |\Delta(x)|.$$ 

We shall proceed by induction. Let $P_i = A_1 A_2 \cdot \cdot \cdot A_i$, then

$$|xI - P_i A^{r-1}| = |xI - (A + D_i) A^{r-1}| = |xI - A^r - D_i A^{r-1}|.$$ 

Now the matrix $D_i A^{r-1}$ is nilpotent and commutative with $xI - A^r$ consequently the determinants above are equal to $|xI - A^r|$. We assume that

$$|xI - P_i A^{r-i}| = |xI - A^r| ;$$

then

$$|xI - P_{i+1} A^{r-i-1}| = |xI - P_i A^{r-i} - P_i D_{i+1} A^{r-i-1}| .$$ 

Here $P_i D_{i+1} A^{r-i-1}$ is commutative with $xI - P_i A^{r-i}$ and is nilpotent because $D_{i+1}$ is nilpotent and commutative with both $P_i$ and $A$; hence by Frobenius' theorem

$$|xI - P_{i+1} A^{r-i-1}| = |xI - P_i A^{r-i}| = |xI - A^r| .$$ 

Consequently

$$|xI - P| = |xI - A^r| = (-1)^{(r-1)n} |\Delta(x)| ,$$

and the theorem is proved.

References


MISSISSIPPI CITY, MISS.