International Mathematical Forum, 3, 2008, no. 17, 829 - 836

A Note on a Generalization of an Extension of the Cayley-Hamilton Theorem

Issam Kaddoura[1] **and Bassam Mourad**[2]

Department of Mathematics, Lebanese International University
Beirut Campus, Al-Mouseitbeh, P.O. Box 14-6404, Beirut, Lebanon

Abstract

In this note, we give a generalization of some extension of the Cayley-Hamilton theorem in the case of a pair of $n \times n$ commuting matrices to the case of a pair of $n \times n$ non-commuting matrices. The classical Cayley-Hamilton theorem and its extension in the case of pairs of commuting matrices, are special cases of the proposed generalization.

Mathematics Subject Classification: 15A54

Keywords: Cayley-Hamilton theorem, Generalized Cayley-Hamilton Theorem

1 Introduction

Let $\mathbf{C}$ denote the field of complex numbers and $\mathbf{M}_n$ denote the class of all $n \times n$ complex matrices. For any matrix $A \in \mathbf{M}_n$. Let $adj(A)$ denote the classical adjoint matrix of A i.e. the transpose of the matrix of cofactors from A. For any matrix A, $det(A)$ represents the determinant of A and for any eigenvalue λ_i of A, let V_{λ_i} denote its corresponding eigenspace.

Now given any polynomial in λ, $P(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_1\lambda + a_{0,}$ with complex coefficients $a_n, a_{n-1}, ..., a_0$, then we can always define a matrix polynomial for any $A \in \mathbf{M}_n$ by $P(A) = a_nA^n + a_{n-1}A^{n-1} + ... + a_1A + a_0I_n$. Now a simple well-known fact in matrix theory for which the proof can be easily checked (see for example [6]), is the following proposition.

Proposition 1.1 *For any $n \times n$ complex matrix A, there exists a nonzero polynomial $P(\lambda)$ over $\mathbf{C}$ such that $P(A) = 0$.*

[1]issam.kaddoura@liu.edu.lb
[2]corresponding author.Email:bassam.mourad@liu.edu.lb

A less obvious result is the well-known Cayley-Hamilton theorem which can be stated as follows.

Theorem 1.2 ***(Cayley-Hamilton)*** *Let A be an $n \times n$ complex matrix and let $P_A(\lambda)$ be the characteristic polynomial of A, that is $P_A(\lambda) = \det(A - \lambda I_n)$. Then $P_A(A) = 0$.*

The Cayley-Hamilton theorem and its extensions have many applications in control systems, electric circuit and many other areas see for example [2] and the references therein, see also [5]. The Cayley-Hamilton theorem has been extended to rectangular matrices, block matrices and to pairs of commuting matrices(see for example [1, 2, 3, 4]). Our intention, in this note, is to give a generalization of some extension of the Cayley-Hamilton theorem in the case of two $n \times n$ matrices which do not necessarily commute. This paper is organized as follows. In the second section, we consider the extension of the Cayley-Hamilton theorem in the case of two $n \times n$ matrices that commute and we collect all the results needed for our purposes and for completeness we include most of the proofs. In the third section, we give a generalization of the extension of the Cayley-Hamilton theorem given in second section which is the main result and we end up with some observations and conclusions.

2 Extension of the Cayley-Hamilton theorem

In this section, we consider the extension of the Cayley-Hamilton theorem to pairs of commuting matrices. For more details on this topic see for example [1, 2, 3, 4].

First we begin with some auxiliary results. Let x and y be two complex numbers and let A and B be two $n \times n$ matrices with entries in the field $\mathbf{C}$. Now define $P_{A,B}(x, y) = \det(xA - yB)$. Then it can be easily checked that $P_{A,B}(x, y)$ is a homogeneous polynomial of degree n in the two variables x and y. Generally, if $AB = BA$, then it is well-known that A and B have a common set of eigenvectors. More precisely, for each eigenvalue λ_i of A, there exists an eigenvalue μ_i of B and an eigenvector X such that $X \in V_{\lambda_i} \cap V_{\mu_i} \neq \{0\}$. Now, let $f(x, y)$ be a complex polynomial in the two variables x and y such that $f(x, y) = \sum_{k=1}^{\infty} a_k x^{\alpha(k)} y^{\beta(k)}$, and $f(A, B) = \sum_{k=1}^{\infty} a_k A^{\alpha(k)} B^{\beta(k)}$ where $\alpha(k)$ and $\beta(k)$ are non-negative[3] integers and a_k are complex. Now if $X \in V_{\lambda_i} \cap V_{\mu_i}$ such that $X \neq \{0\}$, then clearly $f(A, B)X = \sum_{k=1}^{\infty} a_k A^{\alpha(k)} B^{\beta(k)} X = \sum_{k=1}^{\infty} a_k \lambda_i^{\alpha(k)} \mu_i^{\beta(k)} = f(\lambda_i, \mu_i)X$. Thus we can conclude the following well-known result.

[3] This assumption could be dropped on the account of considering A and B to be invertible.

Proposition 2.1 *Let A and B be any two $n \times n$ complex matrices that commute and let $f(x,y)$ be any polynomial in the variables x and y. In addition, let $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ and $\{\mu_1, \mu_2, ..., \mu_n\}$ be respectively the eigenvalues of A and B such that $V_{\lambda_i} \cap V_{\mu_i} \neq \{0\}$, for all $i = 1, 2, ..., n$. Then $f(\lambda_i, \mu_i)$ is an eigenvalue of $f(A, B)$.*

The extension of the Cayley-Hamilton theorem giving in [1, 2, 3, 4] can be stated as follows. For completeness, we include a somewhat different proof.

Theorem 2.2 *[1, 2, 3, 4] Let A and B be two $n \times n$ matrices over $\mathbf{C}$ and let $P_{A,B}(x, y) = \det(xA - yB)$. If A and B commute, then $P_{A,B}(B, A) = 0$.*

Proof. Without loss of generality, we can assume that

$$P_{A,B}(x, y) = a_n x^n + a_{n-1} x^{n-1} y + ... + a_1 x y^{n-1} + a_0 y^n.$$

Let N be the adjoint matrix of the matrix $xA - yB$ that is $N = adj(xA - yB)$. Clearly, N can be written in the form $N = N_{n-1}x^{n-1} + N_{n-2}x^{n-2}y + ... + N_1 x y^{n-2} + N_0 y^{n-1}$ for some $n \times n$ complex matrices $N_{n-1}, N_{n-2}, ..., N_1, N_0$. Now using the fact that for any complex matrix A, $A.adj(A) = \det(A) I_n$ then $(xA - yB)N = P_{A,B}(x, y) I_n$ or $(xA - yB)(N_{n-1}x^{n-1} + N_{n-2}x^{n-2}y + ... + N_1 x y^{n-2} + N_0 y^{n-1}) = P_{A,B}(x, y) I_n$. Expanding the left-hand side and comparing the coefficients of x^{ij} in both sides, we obtain the following $n+1$ equations:

$$\begin{cases} AN_{n-1} = a_n I_n \\ AN_{n-2} - BN_{n-1} = a_{n-1} I_n \\ AN_{n-3} - BN_{n-2} = a_{n-2} I_n \\ \cdot \\ \cdot \\ \cdot \\ AN_0 - BN_1 = a_1 I_n \\ -BN_0 = a_0 I_n \end{cases}$$

Multiplying these equations to the left by $B^n, B^{n-1}A, B^{n-2}A^2, ..., A^n$ respectively, and using the fact that A and B commute, we obtain the following equations:

$$\begin{cases} B^n A N_{n-1} = a_n B^n I_n \\ B^{n-1} A^2 N_{n-2} - B^n A N_{n-1} = a_{n-1} B^{n-1} A I_n \\ B^{n-2} A^3 N_{n-3} - B^{n-1} A^2 N_{n-2} = a_{n-2} B^{n-2} A^2 I_n \\ \cdot \\ \cdot \\ \cdot \\ B A^n N_0 - B^2 A^{n-1} N_1 = a_1 B A^{n-1} I_n \\ -B A^n N_0 = a_0 A^n I_n \end{cases}$$

and the proof is completed by adding up these equations. ■

Remark 2.3 *In the above theorem, if $x = 1$ and $B = I_n$, then we obtain the classical Cayley-Hamilton theorem.*

Using Proposition 2.1, we end this section by the following corollary.

Corollary 2.4 *Let A and B be two $n \times n$ complex matrices that commute, and let $P_{A,B}(x, y) = \det(xA - yB)$. If $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ and $\{\mu_1, \mu_2, ..., \mu_n\}$ are respectively the eigenvalues of A and B such that $V_{\lambda_i} \cap V_{\mu_i} \neq \{0\}$, for all $i = 1, 2, ..., n$. Then $P_{A,B}(\mu_i, \lambda_i) = 0$.*

3 Generalized Cayley-Hamilton Theorem

In this section, we prove a generalization of the extension giving in Theorem 2.2. However first, we need the following three lemmas.

Lemma 3.1 *Let A and B be any two $n \times n$ complex matrices. Then for all $r, s \in \mathbf{C}$, the two matrices $r \det(A)I_n + sA.adj(B)$ and $s \det(B)I_n + rB.adj(A)$ commute and so does the matrix $r \det(B)I_n + s.adj(A).B$ with $s \det(A)I_n + r.adj(B).A$.*

Proof. It suffices to notice that $A.adj(B)$ and $B.adj(A)$ commute and so do $adj(A)B$ and $adj(B)A$. ■

Lemma 3.2 *Let A and B be any two $n \times n$ complex matrices. If A or B is invertible, then there exist $r, s \in \mathbf{C}$ such that the matrix $r.adj(A) + s.adj(B)$ is invertible.*

Proof. It suffices to take $r = 0$ and $s \neq 0$ if B is invertible and $r \neq 0$ and $s = 0$ if A is invertible. ■

Lemma 3.3 *Let A and B be any two $n \times n$ complex matrices. Let $P_{A,B}(x, y) = \det(xA - yB) = \sum_{k=0}^{n} a_k x^k y^{n-k}$. Then $a_n = \det(A)$ and $a_0 = (-1)^n \det(B)$*

Proof. It suffices to notice that for $x = 0$ and $y = 1$, $f(x, y) = \det(-B) = (-1)^n \det(B)$, and for $x = 1$ and $y = 0$ then $f(x, y) = \det(A)$. ■

Now we are ready to prove the following main result.

Theorem 3.4 *Let A and B be any two $n \times n$ matrices with entries in $\mathbf{C}$ and let $P_{A,B}(x, y) = \det(xA - yB)$. Then there exists $r, s \in \mathbf{C}$, not both zeros at the same time such that $P_{A,B}\left(s \det(B)I_n + rB.adj(A), r \det(A)I_n + sA.adj(B)\right) = 0$, and $P_{A,B}(r \det(B)I_n + s.adj(A).B, s \det(A)I_n + r.adj(B).A) = 0$.*

Proof. We split the proof into two cases:
Case one: If one of the two matrices A or B is invertible, then by Lemma 3.2, there exist $r, s \in \mathbf{C}$ such that $\det[r.adj(A)+s.adj(B)] \neq 0$. Since $\det[r.adj(A)+s.adj(B)]$ is a nonzero complex constant, then the two polynomials $P_{A,B}(x,y) = \det(xA-yB)$ and $\det(xA-yB)\det[r.adj(A)+s.adj(B)]$ have the same zeros. clearly $\det[(xA-yB)(r.adj(A)+s.adj(B))] = \det[x(r\det(A)I_n+sA.adj(B))-y(s\det(B)I_n+rB.adj(A))] = P_{M,N}(x,y)$ where $M = r\det(A)I_n+sA.adj(B)$ and $N = s\det(B)I_n+rB.adj(A)$. By Lemma 3.1, the two matrices M and N commute, and therefore the proof is complete by using Theorem 2.2.
Case two: If both matrices A and B are singular, then $P_{A,B}(x,y) = \det(xA-yB) = \sum_{k=0}^{n} a_k x^k y^{n-k}$ with $a_0 = a_n = \det(A) = \det(B) = 0$ (see Lemma 3.3) and now a simple check shows that

$$P_{A,B}\left(s\det(B)I_n + rB.adj(A), r\det(A)I_n + sA.adj(B)\right) =$$

$$P_{A,B}\left(rB.adj(A), sA.adj(B)\right) = 0,$$

for all $r, s \in \mathbf{C}$. The proof of the second part is done by a similar method. ■

Note that in the proof of the above theorem, we have shown that if A and B are any two $n \times n$ complex singular matrices that do not necessarily commute and if $P_{A,B}(x,y) = \det(xA-yB) = \sum_{k=0}^{n} a_k x^k y^{n-k}$, then for all $r, s \in \mathbf{C}$, $P_{A,B}(sB.adj(A), rA.adj(B)) = 0$. Now if one of the two matrices A or B is invertible, then there exist $r, s \in \mathbf{C}$, such that $\det[r.adj(A)+s.adj(B)] \neq 0$, and the two matrices $M = r\det(A)I_n+sA.adj(B)$ and $N = s\det(B)I_n+rB.adj(A)$ commute. Moreover if

$$P_{M,N}(x,y) = \det(xA-yB)\det[r.adj(A)+s.adj(B)],$$

then we have proved that

$$P_{M,N}(N,M) = P_{A,B}(N,M)\det[r.adj(A)+s.adj(B)] = 0.$$

Next we will prove that this last equation for commuting matrices A and B implies that $P_{A,B}(B,A) = \sum_{k=0}^{n} a_k B^k A^{n-k} = 0$. For, if $P_{A,B}(N,M)\det[r.adj(A)+s.adj(B)] = 0$, then clearly $P_{A,B}(N,M) = \sum_{k=0}^{n} a_k N^k M^{n-k} = \sum_{k=0}^{n} a_k (s\det(B)I_n + rB.adj(A))^k (r\det(A)I_n + sA.adj(B))^{n-k} = 0$. Now using the fact that for any square matrix A, $adj(A)$ is a polynomial in A, then if $AB = BA$, we obtain $adj(A)B = B.adj(A)$, and $A.adj(B) = adj(B).A$. So that A commutes with $adj(B)$ and similarly B commutes with $adj(A)$. Hence A and B both commute with $r.adj(A) + s.adj(B)$ and therefore they both commute with

$(r.adj(A) + s.adj(B))^k$ for any positive integer k. Thus a simple inspection shows that $P_{A,B}(N, M) = (r.adj(A) + s.adj(B))^n P_{A,B}(B, A) = 0$. As a consequence, we have the following corollary.

Corollary 3.5 *with the same notation as in the above theorem, we have the following:*

1. *If $r = 0$, $s = 1$, $B = I_n$ and $x = 1$, then the above theorem becomes the classical Cayley-Hamilton theorem.*
2. *If A and B commute, then $P_{A,B}(B, A) = 0$, and we obtain Theorem 2.2.*

Finally, we end this section with the following theorem.

Theorem 3.6 *Let A and B be any two $n \times n$ matrices with entries in $\mathbf{C}$, and let $P_{A,B}(x, y) = \det(xA - yB)$. For all $r, s \in \mathbf{C}$, define $M = s\det(B)I_n + rB.adj(A)$, and $N = r\det(A)I_n + sA.adj(B)$. Now, let the eigenvalues of the matrices M and N be respectively $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ and $\{\mu_1, \mu_2, ..., \mu_n\}$ (counted with their multiplicity) ordered in the way such that $V_{\lambda_i} \cap V_{\mu_i} \neq \{0\}$, for all $i = 1, ..., n$.. Then $MN - s\det(B)M - r\det(A)N = 0$, for all $i = 1, ..., n$. Moreover,*

$$P_{M,N}(\mu_i, \lambda_i) = 0, \tag{1}$$

and,

$$\lambda_i \mu_i - sdet(B)\lambda_i - r\det(A)\mu_i = 0 \tag{2}$$

for all $i = 1, ..., n$.

Proof. Clearly

$$MN = r^2 \det(A)B.adj(A) + 2rs\det(AB)I_n + s^2 \det(B)A.adj(B),$$

$$s\det(B)M = rs\det(A)\det(B)I_n + s^2 \det(B)A.adj(B),$$

and

$$r\det(A)N = rs\det(A)\det(B)I_n + r^2 \det(A)B.adj(A).$$

So that $MN - s\det(B)M - r\det(A)N = 0$. For the second part, it suffices to see that since M and N commute, then by Corollary 2.4, $P_{M,N}(\mu_i, \lambda_i) = 0$, for all $i = 1, ..., n$.The proof of the last part is completed by Proposition 2.1. ∎

Concluding Remarks

One could think of the polynomial $P_{A,B}(x,y) = \det(xA - yB)$ as the joint characteristic polynomial of A and B. Now two cases present themselves:

1. If A and B commute and if, in addition, the eigenvalues of A and B are respectively $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ and $\{\mu_1, \mu_2, ..., \mu_n\}$ (counted with their multiplicity) ordered in the way such that $V_{\lambda_i} \cap V_{\mu_i} \neq \{0\}$, for all $i = 1, ..., n$, then $P_{A,B}(B, A) = 0$ and $P_{A,B}(\mu_i, \lambda_i) = 0$. Then clearly $P_{A,B}(x,y) = \alpha \prod_{i=1}^{n} (\lambda_i x - \mu_i y)$ for some constant α in $\mathbf{C}$.

2. If A and B do not commute and if, in addition, the eigenvalues of $M = r\det(A)I_n + sA.adj(B)$ and $N = s\det(B)I_n + rB.adj(A)$, for some complex numbers r and s, are respectively $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ and $\{\mu_1, \mu_2, ..., \mu_n\}$ (counted with their multiplicity) ordered in the way such that $V_{\lambda_i} \cap V_{\mu_i} \neq \{0\}$, for all $i = 1, ..., n$, then $P_{A,B}(N, M) = 0$ and $P_{A,B}(\mu_i, \lambda_i) = 0$. So that we do not have much information on the eigenvalues of A and B. Here all we can hope for are the relations between each pair of eigenvalues of M and N given by (1) and (2).

Finally, recall that for any monic polynomial in one variable $P(x)$ of degree n, then there exists an $n \times n$ companion matrix C such that $P(x)$ is the characteristic polynomial of C. Now for any homogeneous polynomial of degree n in two variables, say, $w(x,y) = \sum_{k=0}^{n} a_k x^k y^{n-k}$ where all the a_k are complex, define the two $n \times n$ matrices A and B by $A = \begin{bmatrix} 0 & 0 & ... & 0 & -\frac{a_0}{a_n} \\ 1 & 0 & ... & 0 & -\frac{a_1}{a_n} \\ 0 & 1 & ... & 0 & -\frac{a_2}{a_n} \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & .... & 1 & -\frac{a_{n-1}}{a_n} \end{bmatrix}$ for $a_n \neq 0$ and $B = \begin{bmatrix} 0 & 0 & ... & 0 & -\frac{a_n}{a_0} \\ 1 & 0 & ... & 0 & -\frac{a_{n-1}}{a_0} \\ 0 & 1 & ... & 0 & -\frac{a_{n-2}}{a_0} \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & .... & 1 & -\frac{a_1}{a_0} \end{bmatrix}$ for $a_0 \neq 0$. An easy inspection shows the following lemma.

Lemma 3.7 *With the same notation as above, we have:*

1. *If $a_n \neq 0$, then $w(x,y) = a_n \det(xI_n - yA) = a_n P_{I_n,A}(x,y)$*

2. *If* $a_0 \neq 0$, *then* $w(x,y) = a_0 \det(yI_n - xB) = a_0 P_{B,I_n}(x,y)$.

As a result we have the following theorem.

Theorem 3.8 *Let* $w(x,y) = \sum_{k=0}^{n} a_k x^k y^{n-k}$ *where all the* a_k *are complex, be any homogeneous polynomial in the two variable* x *and* y, *and let* A *and* B *be defined as above. Then for all complex numbers* r *and* s, *we have:*

1. *If* $a_n \neq 0$, *then* $w(rA + s\det(A)I_n, rI_n + s.adj(A)) = 0$.

2. *If* $a_0 \neq 0$, *then* $w(sI_n + r.adj(B), r\det(B)I_n + s.B) = 0$.

Moreover, the zeroes of $w(x,y)$ *can be determined by either finding the eigenvalues of the matrix* $rA + s\det(A)I_n$ *or those of the matrix* $rdet(B)I_n + sB$.

Proof. The first part is an immediate consequence of the above lemma and Theorem 3.4, and the second part is a result of Theorem 3.6. ■

References

[1] Chang F.R. and Chan C.N., The generalized Cayley-Hamilton theorem for standard pencils; *System and Control Lett.*, vol. **18**(1992), 179-182.

[2] Kaczorek T., New extensions of the Cayley-Hamilton theorem with applications; *Proceeding of the 19th European Conference on Modelling and Simulation,* 2005.

[3] Kaczorek T., An existence of the Cayley-Hamilton theorem for nonsquare block matrices and computation of the left and right inverses; *Bull. Pol. Acad. Techn. Sci.*, vol. **43**, No 1 (1995), 49-56.

[4] Lewis L., Further remarks on the Cayley-Hamilton theorem and Fadeev's method for the matrix pencil $[sE - A]$, *IEEE Trans. Automat. Contr.,* **31** (1986), 869-870.

[5] Urrutia L.F. and Morales N., The Cayley-Hamilton theorem for supermatrices, *J. Phys. A: Math. Gen.,* **27** (1994), 1981-1997.

[6] Zhang F., *Matrix theory; Basic results and techniques,* Springer, New York, 1999.

Received: October 12, 2007