Eigenvalues of the Sum and Product of Anticommuting Matrices

Vadim Ponomarenko* Louis Selstad†

June 2, 2020

Abstract

Improving on a recent result of Zhong, we characterize the eigenvalues of AB and $A+B$, for square matrices A, B satisfying $AB + BA = 0$.

Square matrices A, B are called anticommuting if $AB = -BA$. Such matrices are important in mathematical physics, e.g. as Pauli spin matrices. They are of continued mathematical interest (see, e.g., [2, 3, 4]). The recent paper [5] presented the following theorem.

Theorem 1 (Zhong)**.** *Let A, B be square matrices with $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}, \sigma(B) = \{\mu_1, \ldots, \mu_n\}$. Suppose that A, B anticommute, and for each eigenvalue λ_i of A, the algebraic multiplicity equals the geometric multiplicity. Then*

1. $\sigma(AB) \subseteq \{\lambda_j \mu_k i : \text{all } j, k\}$*; and*
2. $\sigma(A + B) \subseteq \left\{\pm\sqrt{\lambda_j^2 + \mu_k^2} : \text{all } j, k\right\}$*.*

We improve on the previous theorem by slightly weakening the hypotheses, and finding $\sigma(AB)$ and $\sigma(A + B)$ exactly (in addition to various structural results). We will put A into Jordan canonical form, which will impose a (to be defined) correspondence between eigenvalues of A and of B. Our main result is:

Main Theorem. *Let A, B be square matrices with $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}, \sigma(B) = \{\mu_1, \ldots, \mu_n\}$. Suppose that A, B anticommute, and for each nonzero eigenvalue λ_i of A, the algebraic multiplicity equals the geometric multiplicity. Then*

1. $\sigma(AB) = \{\lambda_j \mu_k i : \text{corresponding } j, k\}$*; and*
2. $\sigma(A + B) = \left\{\pm\sqrt{\lambda_j^2 + \mu_k^2} : \text{corresponding } j, k\right\}$*.*

Proof. Corollaries 7 and 9, to follow. □

*San Diego State University, corresponding author
†San Diego State University

Before proceeding, we set some notation. We define $\mathbb{C}^\bullet$ as $\mathbb{C} \setminus \{0\}$. Given a square $n \times n$ matrix M, we recall the characteristic polynomial $p_M(t) = |tI_n - M|$. A key property of characteristic polynomials is that $p_{AB}(t) = p_{BA}(t)$. If $AB+BA = 0$, then $p_{AB}(t) = p_{BA}(t) = p_{-AB}(t) = |tI_n + AB| = (-1)^n|(-t)I_n - AB| = (-1)^n p_{AB}(-t)$. Consequently, for even (resp. odd) n, the function $p_{AB}(t)$ is even (resp. odd). Since $p_{AB}(t)$ is a polynomial, it must therefore be a sum of monomials in only even (resp. odd) powers of t. In general, $p_{AB}(t)$ cannot be determined from $p_A(t)$ and $p_B(t)$.

We now present a familiar definition, generalized to nonsquare matrices.

Definition 2. *Let $m, n \in \mathbb{N}$, and let C be an $m \times n$ matrix. We say that C is* upper triangular *if it satisfies $C_{i,j} = 0$ for all i, j with $j - i < \max(0, n - m)$.*

For upper triangular matrix C, its nonzero entries are confined to a triangle in the upper-right corner, whose sides are horizontal, vertical, and of slope -1. This triangle has one corner at $C_{1,n}$, and the other corners at either $\{C_{1,1}, C_{n,n}\}$ (if $m \geq n$), or at $\{C_{1,n-m+1}, C_{m,n}\}$ (if $m \leq n$). For $m = n$, this coincides with the usual definition of "upper triangular."

The product of upper triangular matrices satisfies a useful property.

Theorem 3. *Let $m, n \in \mathbb{N}$. Let F be an upper triangular $m \times n$ matrix, and E an upper triangular $n \times m$ matrix. Then FE is an $m \times m$ matrix satisfying $(FE)_{i,j} = 0$ for all i, j with $j - i < |n - m|$. In particular, if $n \neq m$, then FE is strictly upper triangular, and hence satisfies $p_{FE}(t) = t^m$.*

Proof. Note that $(FE)_{i,j} = \sum_{k=1}^n F_{i,k}E_{k,j}$. Suppose first that $n \geq m$, and $j - i < n - m$. Then $F_{i,k} = 0$ for all i, k with $k - i < n - m$. Also, $E_{k,j} = 0$ for all k, j with $j - k < 0$. Hence, if $k > j$, then $E_{k,j} = 0$; if instead $k \leq j$, then $k - i \leq j - i < n - m$, so $F_{i,k} = 0$. Thus $(FE)_{i,j} = 0$. Suppose now that $n < m$, and $j - i < m - n$. Then $F_{i,k} = 0$ for all i, k with $k - i < 0$. Also, $E_{k,j} = 0$ for all k, j with $j - k < m - n$. Hence, if $k < i$, then $F_{i,k} = 0$; if instead $k \geq i$, then $j - k \leq j - i < m - n$, so $E_{k,j} = 0$. Thus $(FE)_{i,j} = 0$. □

We recall the (square) upper shift matrix U_n (of size $n \in \mathbb{N}$), which has ones on the superdiagonal and zeroes elsewhere. Using the Kronecker delta function, we may define $(U_n)_{i,j} = \delta_{i+1,j}$. For future use, we define matrix $U_n^\star$ to be any integer matrix whose entries satisfy $0 \leq (U_n^\star)_{i,j} \leq (U_n)_{i,j}$. We next recall (square) Jordan blocks, which exist for any eigenvalue α and any size $n \in \mathbb{N}$, defined as $J_n(\alpha) = \alpha I_n + U_n$.

We recall that a matrix is in Jordan canonical form if it is a block diagonal matrix, whose diagonal blocks are each Jordan blocks (of any size and any eigenvalues). The celebrated Jordan canonical form theorem states that every square matrix is similar to one in Jordan form. We now apply this similarity to the anticommutative relation. If $AB + BA = 0$, choose P so that PAP^{-1} is in Jordan form. We have $PAP^{-1}PBP^{-1} + PBP^{-1}PAP^{-1} = P0P^{-1} = 0$. Set $A' = PAP^{-1}, B' = PBP^{-1}$; we have $A'B' + B'A' = 0$. A is similar to A' and B is similar to B', and similarity preserves characteristic polynomials (and hence eigenvalues). Henceforth we simply assume that $AB + BA = 0$ and that A is in Jordan form.

We now partition A, B into blocks, based on the block diagonal structure of A's Jordan form. Considering an arbitrary $m \times n$ block C of B we find that the anticommuting relation forces $J_m(\alpha)C + CJ_n(\beta) = 0$, for some diagonal Jordan blocks $J_m(\alpha), J_n(\beta)$ of A. This forces most such blocks C to be zero, and imposes upper triangularity on the rest.

Theorem 4. *Let $m, n \in \mathbb{N}$, and let C be an $m \times n$ matrix. Suppose that $J_m(\alpha)C + CJ_n(\beta) = 0$, for some $\alpha, \beta \in \mathbb{C}$. If $\alpha + \beta \neq 0$, then $C = 0$. If instead $\alpha + \beta = 0$, then C must be upper triangular.*

Proof. Note that $(U_m C)_{m,j} = 0$, and $(U_m C)_{i,j} = C_{i+1,j}$ for $i < m$. Similarly, $(CU_n)_{i,1} = 0$, and $(CU_n)_{i,j} = C_{i,j-1}$ otherwise. Conventionally, define $C_{i,j} = 0$ if $i > m$ or $j < 1$. We have $0 = J_m(\alpha)C + CJ_n(\beta) = (\alpha + \beta)C + U_m C + CU_n$. Now, $U_m C$ moves the rows of C upward, inserting a zero row, while CU_n moves the columns of C to the right, inserting a zero column. Hence, in $U_m C + CU_n$, the southwest frontier of C must move northeast.

We first consider $\alpha + \beta \neq 0$. Set S to be the set of those $(i, j) \in \mathbb{N}^2 \cap [1, m] \times [1, n]$ which correspond to a nonzero entry in C. Suppose, by way of contradiction, that S is nonempty. Take $(i, j) \in K'$ that is minimal with respect to $j - i$, i.e. on the southwest frontier of C. If there is more than one with this minimal $j - i$, take the one with maximal i. Look at the (i, j) entry of $0 = (\alpha + \beta)C + U_m C + CU_n$. We have $0 = (\alpha + \beta)C_{i,j} + C_{i+1,j} + C_{i,j-1}$. Either since $i + 1 > m$ or since $j - (i + 1) < j - i$, we must have $C_{i+1,j} = 0$. Either since $j - 1 < 1$ or since $(j - 1) - i < j - i$, we must have $C_{i,j-1} = 0$. In other words, the southwest frontier of $U_m C + CU_n$ has moved away from (i, j). Hence $0 = (\alpha + \beta)C_{i,j}$. But since $\alpha + \beta \neq 0$, $C_{i,j} = 0$, which is a contradiction. Consequently S is empty.

We now consider the case of $\alpha + \beta = 0$. First we take the case $m \geq n$. Set K to be the set of those $(i, j) \in \mathbb{N}^2 \cap [1, m] \times [1, n]$ such that $j - i < 0$. Let K' be those elements of K which correspond to a nonzero entry in C. Suppose, by way of contradiction, that K' is nonempty. Take $(i, j) \in K'$ that is minimal with respect to $j - i$. If there is more than one with this minimal $j - i$, take the one with minimal i. We have $0 = (\alpha + \beta)C + U_m C + CU_n = U_m C + CU_n$. Note that if $i = 1$ then $j < 1$, which is impossible. Hence $i \geq 2$. Look at the $(i - 1, j)$ entry of this matrix. We have $0 = C_{i,j} + C_{i-1,j-1}$. Either since $j < 1$, or since $(j - 1) - (i - 1) = j - i$ and $i - 1 < i$, we must have $C_{i-1,j-1} = 0$. But then $C_{i,j} = 0$, a contradiction. Hence K' is empty.

Lastly, we consider the case $m > n$. Set K to be the set of those $(i, j) \in \mathbb{N}^2 \cap [1, m] \times [1, n]$ such that $j - i < n - m$. Let K' be those elements of K which correspond to a nonzero entry in C. Suppose, by way of contradiction, that K' is nonempty. Take $(i, j) \in K'$ that is minimal with respect to $j - i$. If there is more than one with this minimal $j - i$, take the one with maximal i. We have $0 = (\alpha + \beta)C + U_m C + CU_n = U_m C + CU_n$. Note that if $j = n$ then $-i < -m$, or $i > m$, which is impossible. Hence $j \leq n - 1$. Look at the $(i, j + 1)$ entry of this matrix. We have $0 = C_{i+1,j+1} + C_{i,j}$. Either since $i + 1 > m$, or since

$(j+1)-(i+1)=j-i$ and $i+1>i$, we must have $C_{i+1,j+1}=0$. But then $C_{i,j}=0$, a contradiction. Hence K' is empty. □

It turns out that the upper triangular matrices C from Theorem 4 have additional banded structure, which we will not explore.

We now impose a particular order to the diagonal Jordan blocks of A. By reordering rows and columns if necessary, we collect together all of the Jordan blocks of eigenvalue 0 (if any), into one big block $A(0)=U_n^\star$, for some $n\in\mathbb{N}_0$ (here n is the algebraic multiplicity of eigenvalue 0 in A). The corresponding diagonal big block of B, which we call $B(0)$, is upper triangular by Theorem 4.

For each nonzero eigenvalue α, we also collect together all of the Jordan blocks of $\pm\alpha$, into one big block $A(\alpha)$, which has n copies of α on the diagonal, followed by m copies of $-\alpha$. By reversing the roles of $\alpha,-\alpha$ if necessary, we assume that $n\geq m$. We have $A(\alpha)=\begin{pmatrix}\alpha I_n+U_n^\star & 0\\ 0 & -\alpha I_m+U_m^\star\end{pmatrix}$. We now repartition A, B into big blocks, based on the block diagonal structure induced by these big blocks. We call by $B(\alpha)=\begin{pmatrix}0 & E\\ F & 0\end{pmatrix}$ the diagonal big block of B corresponding to $A(\alpha)$. We call this the big block Jordan form for B. By Theorem 4 again, B will now also be block diagonal. Hence $p_B(t)$ is just the product of all the $p_{B(\alpha)}(t)$.

These big blocks induce a correspondence between the eigenvalues of $A(\alpha)$, namely $\pm\alpha$, and the eigenvalues of $B(\alpha)$.

We can now compute the characteristic polynomial of $B(\alpha)$ (for nonzero α).

Theorem 5. *Let $m,n\in\mathbb{N}$ with $n\geq m$. Suppose E is an $n\times m$ matrix, and F is an $m\times n$ matrix. Set $B(\alpha)=\begin{pmatrix}0 & E\\ F & 0\end{pmatrix}$. Then $p_{B(\alpha)}(t)=t^{n-m}p_{FE}(t^2)$.*

Proof. We calculate $p_{B(\alpha)}(t)=|tI-B(\alpha)|=\begin{vmatrix}tI_n & -E\\ -F & tI_m\end{vmatrix}$. Since tI_n is invertible, this equals $|tI_n||tI_m-(-F)(tI_n)^{-1}(-E)|=t^n|tI_m-t^{-1}FE|=t^n|t^{-1}I_m||t^2I_m-FE|=t^{n-m}p_{FE}(t^2)$. □

If $n>m$, then by Theorem 3, $p_{FE}(t)=t^m$. Hence, $p_{B(\alpha)}=t^{n-m}t^{2m}=t^{n+m}$, and $B(\alpha)$ has only 0 as an eigenvalue. If we assume instead that big block $B(\alpha)$ has a nonzero eigenvalue, then $n=m$. This makes $p_{A(\alpha)}(t)=(t-\alpha)^m(t+\alpha)^m=(t^2-\alpha^2)^m$ and $p_{B(\alpha)}(t)=p_{FE}(t^2)$ both even. If we assume that A is invertible, and that every big block $B(\alpha)$ has a nonzero eigenvalue, then $p_A(t)$ and $p_B(t)$ must both be even.

We now focus our attention on the case when big block $A(\alpha)$ is diagonalizable, for nonzero α. In this case, the geometric multiplicity and algebraic multiplicity of α coincide.

Theorem 6. *Let $\alpha\in\mathbb{C}^\bullet$, and let $m,n\in\mathbb{N}$. Suppose we have $A(\alpha)=\begin{pmatrix}\alpha I_n & 0\\ 0 & -\alpha I_m\end{pmatrix}$, $B(\alpha)=\begin{pmatrix}0 & E\\ F & 0\end{pmatrix}$. Then $p_{A(\alpha)B(\alpha)}(t)=(i\alpha)^{n+m}p_{B(\alpha)}(\frac{t}{i\alpha})$, and $p_{A(\alpha)+B(\alpha)}(t)=(t-\alpha)^{n-m}p_{FE}(t^2-\alpha^2)$.*

Proof. We calculate $p_{A(\alpha)B(\alpha)}(t) = |tI - A(\alpha)B(\alpha)| = \begin{vmatrix} tI_n & -\alpha E \\ \alpha F & tI_m \end{vmatrix}$. Since tI_n is invertible, this equals $|tI_n||tI_m - (\alpha F)(tI_n)^{-1}(-\alpha E)| = t^n|tI_m - (-\alpha^2 t^{-1})FE| = t^n|(-\alpha^2)t^{-1}I_m||(-\frac{t^2}{\alpha^2}I_m - FE)| = t^{n-m}(-\alpha^2)^m p_{FE}((\frac{t}{i\alpha})^2) =$
$= (\frac{t}{i\alpha})^{n-m}(i\alpha)^{n-m}(i\alpha)^{2m}p_{FE}((\frac{t}{i\alpha})^2) = (i\alpha)^{n+m}p_{B(\alpha)}(\frac{t}{i\alpha})$.

We now calculate $p_{A(\alpha)+B(\alpha)}(t) = |tI - A(\alpha) - B(\alpha)| = \begin{vmatrix} (t-\alpha)I_n & -E \\ F & (t+\alpha)I_m \end{vmatrix}$. Since $(t-\alpha)I_n$ is invertible, this equals $|(t-\alpha)I_n||(t+\alpha)I_m - F((t-\alpha)I_n)^{-1}E| = (t-\alpha)^n|(t+\alpha)I_m - (t-\alpha)^{-1}FE| = (t-\alpha)^n|(t-\alpha)^{-1}I_m||(t+\alpha)(t-\alpha)I_m - FE| = (t-\alpha)^{n-m}p_{FE}(t^2 - \alpha^2)$. ☐

From our characteristic polynomial calculations, we can derive the eigenvalues of $A(\alpha)B(\alpha)$ and $A(\alpha) + B(\alpha)$ from the eigenvalues of $A(\alpha)$ and $B(\alpha)$.

Corollary 7. *Let $\alpha \in \mathbb{C}^\bullet$, and let $m, n \in \mathbb{N}$. Suppose we have $A(\alpha) = \begin{pmatrix} \alpha I_n & 0 \\ 0 & -\alpha I_m \end{pmatrix}$, $B(\alpha) = \begin{pmatrix} 0 & E \\ F & 0 \end{pmatrix}$. Let $\{\lambda_k\}$ denote the eigenvalues of $B(\alpha)$, with multiplicity. Then the eigenvalues of $A(\alpha)B(\alpha)$ are exactly $\{\pm i\alpha\lambda_k\}$, and the eigenvalues of $A(\alpha) + B(\alpha)$ are exactly $\{\pm\sqrt{\alpha^2 + \lambda_k^2}\}$.*

Proof. Let $p_{FE}(t) = \prod_{k=1}^m (t - \mu_k)$, where the μ_k are the eigenvalues of FE, not assumed distinct. Then $p_{B(\alpha)}(t) = t^{n-m}\prod_{k=1}^m (t^2 - \mu_k)$; hence the eigenvalues of $B(\alpha)$ consist of (a) $n - m$ copies of 0; and (b) $\pm\sqrt{\mu_k}$, for each eigenvalue of FE. We can interpret the complex square root as a principal value, but it doesn't matter since we get both square roots.

Now $p_{A(\alpha)B(\alpha)}(t) = (\frac{t}{i\alpha})^{n-m}(i\alpha)^{n+m}\prod_{k=1}^m ((\frac{t}{i\alpha})^2 - \mu_k) = t^{n-m}\prod_{k=1}^m (t^2 - (-\alpha^2\mu_k))$; hence its eigenvalues are (a) $n - m$ copies of 0; and (b) $\pm\sqrt{-\alpha^2\mu_k}$, for each eigenvalue μ_k of FE.

Finally $p_{A(\alpha)+B(\alpha)}(t) = (t-\alpha)^{n-m}p_{FE}(t^2 - \alpha^2) = (t-\alpha)^{n-m}\prod_{k=1}^m (t^2 - \alpha^2 - \mu_k)$; hence its eigenvalues are (a) $n - m$ copies of α; and (b) $\pm\sqrt{\alpha^2 + \mu_k}$. ☐

Lastly, we extend these results to $A(0)$ and $B(0)$.

Theorem 8. *Let 0 be an eigenvalue of A of algebraic multiplicity $n \geq 1$. Let $A(0)$, $B(0)$ be the corresponding big blocks. Then $p_{A(0)B(0)}(t) = p_{A(0)}(t) = t^n$, and $p_{A(0)+B(0)}(t) = p_{B(0)}(t)$.*

Proof. $A(0)$ is strictly upper triangular, and $B(0)$ is upper triangular. Hence $A(0)B(0)$ is strictly upper triangular, while $A(0) + B(0)$ is upper triangular with the same diagonal entries as $B(0)$. ☐

Corollary 9. *Let 0 be an eigenvalue of A of algebraic multiplicity $n \geq 1$. Let $A(0)$, $B(0)$ be the corresponding big blocks. Let $\{\lambda_k\}$ denote the eigenvalues of $B(\alpha)$, with multiplicity. Then the eigenvalues of $A(\alpha)B(\alpha)$ are all 0, and the eigenvalues of $A(\alpha) + B(\alpha)$ are exactly $\{\lambda_k\}$.*

We close by observing that it appears that the diagonalizability hypothesis on $A(\alpha)$ can be weakened or removed entirely, but we are unable to prove this.

Acknowledgments

The authors would like to thank the anonymous referee, who pointed out Theorem 4.4.6 and Lemma 4.4.11 in [1], which the reader may wish to consult as related to our Theorem 4.

References

[1] Roger A. Horn and Charles R. Johnson. *Topics in matrix analysis.* Cambridge University Press, Cambridge, 1991.

[2] Pavel Hrubeš. On families of anticommuting matrices. *Linear Algebra Appl.*, 493:494–507, 2016.

[3] Daniel B. Shapiro and Reiner Martin. Problems and Solutions: Solutions: Anticommuting Matrices: 10456. *Amer. Math. Monthly*, 105(6):565–566, 1998.

[4] Olga Taussky. Sums of squares. *Amer. Math. Monthly*, 77:805–830, 1970.

[5] Wenti Zhong. On the eigenvalues of anticommuting matrices. *College Math. J.*, 48(5):368–369, 2017.