BOUNDARY MATRICES AND THE MARCUS-DE OLIVEIRA DETERMINANTAL CONJECTURE*

AMEET SHARMA†

Abstract. We present notes on the Marcus-de Oliveira conjecture. The conjecture concerns the region in the complex plane covered by the determinants of the sums of two normal matrices with prescribed eigenvalues. Call this region Δ. This paper focuses on boundary matrices of Δ. We prove 2 theorems regarding these boundary matrices. This paper uses ideas from [1].

Key words. determinantal conjecture, Marcus-de Oliveira, determinants, normal matrices, convex-hull

AMS subject classifications. 15A15, 15A16

1. Introduction.

Marcus [4] and de Oliveira [2] made the following conjecture. Given two normal matrices A and B with prescribed eigenvalues $a_1, a_2 ... a_n$ and $b_1, b_2 ... b_n$ respectively, $det(A + B)$ lies within the region:

$$co\Big\{ \prod (a_i + b_{\sigma(i)}) \Big\}$$

where $\sigma \in S_n$. co denotes the convex hull of the $n!$ points in the complex plane. As described in [1], the problem can be restated as follows. Given two diagonal matrices, $A_0 = diag(a_1, a_2 ... a_n)$ and $B_0 = diag(b_1, b_2 ... b_n)$, let:

$$\Delta = \Big\{ det(A_0 + UB_0U^*) : U \in U(n) \Big\} \tag{1.1}$$

where $U(n)$ is the set of $n \times n$ unitary matrices. Then we can write the conjecture as:

CONJECTURE 1.1 (Marcus-de Oliveira Conjecture).

$$\Delta \subseteq co\Big\{ \prod (a_i + b_{\sigma(i)}) \Big\} \tag{1.2}$$

Let

$$M(U) = det(A_0 + UB_0U^*). \tag{1.3}$$

The paper is organized as follows. In section 2 we define terms and functions that will be used in the rest of the paper. These definitions are necessary to state our results. In section 3, we state 3 lemmas and 2 theorems that form the bulk of the paper. We state them in the order they are proved.

2. Preparatory definitions.

*Submitted to the editors June 6th, 2018.

† (ameet_n_sharma@hotmail.com).

2.1. Terms. Given a unitary matrix U and square, diagonal matrices A_0 and B_0 all of dimension $n \times n$,

- If $M(U)$ is a point on $\partial\Delta$ (the boundary of Δ), we call $M(U)$ a boundary point of Δ and we call U a **boundary matrix** of Δ. See (1.1) and (1.3).
- We define the **B-matrix** of U as UB_0U^*.
- We define the **C-matrix** of U as $A_0 + UB_0U^*$.
- We define the **F-matrix** of U as $C^{-1}A_0 - A_0C^{-1}$ where C is the C-matrix of U. Note that the F-matrix is only defined when C is invertible, or equivalently when $det(C) = M(U) \neq 0$. See (1.3). Also note that since A_0 is diagonal, the F-matrix is a zero-diagonal matrix. The idea for using the F-matrix comes from [1], Theorem 4, p.27.

Throughout the rest of the paper, we'll assume A_0 and B_0 are defined, even if we don't explicitly mention them.

2.2. Functions given a unitary matrix U. Given a unitary matrix U with B-matrix B, C-matrix C and F-matrix F. Given $M(U) \neq 0$. For every skew-hermitian matrix Z, we define the following functions

let

$$U_Z(t) = (e^{Zt})U \tag{2.1}$$

where t is any real number.

Since the exponential of a skew-hermitian matrix is unitary, $U_Z(t)$ is a function of unitary matrices.

let

$$B_Z(t) = U_Z(t)B_0U_Z^*(t) \tag{2.2}$$

let $C_Z(t) = A_0 + B_Z(t)$

We note that $B_Z(0) = B$ and $C_Z(0) = C$.

let

$$R_Z(t) = det(C_Z(t)) \tag{2.3}$$

We can see by (1.1) that $R_Z(t) \subseteq \Delta$.

$R_Z(0) = A_0 + UB_0U^*$

So by (1.3) we see that $R_Z(0) = M(U)$.

So all the $R_Z(t)$ functions go through $M(U)$ at $t = 0$.

We shall refer to these functions in the rest of the paper with the same notation (for example $R_Z(t)$ for a skew-hermitian matrix Z. $R_{Z_1}(t)$ for a skew-hermitian matrix Z_1). Note that $R_Z(t)$ requires A_0, B_0, U and Z in order to be defined. But we won't explicitly mention A_0 and B_0. All the results in this paper assume there are two diagonal matrices A_0 and B_0 defined in the background.

2.3. Skew-Hermitian matrices Z^{ab} and $Z^{ab,i}$. Given two integers a,b where $1 \leq a, b \leq n$ and $a \neq b$.

We define the $n \times n$ skew-hermitian matrix Z^{ab} as follows. $Z^{ab}_{ab} = -1$ (the element at the ath row and bth column is -1.) $Z^{ab}_{ba} = 1$ (the element at the bth row and ath column is 1.) And all other elements are 0. Note that $Z^{ab} = -Z^{ba}$.

We define the $n \times n$ skew-hermitian matrix $Z^{ab,i}$ as follows. $Z^{ab,i}_{ab} = i$ and $Z^{ab,i}_{ba} = i$. All other elements are zero. Note that $Z^{ab,i} = Z^{ba,i}$.

It is straightforward to verify that Z^{ab} and $Z^{ab,i}$ are skew-hermitian.

3. Main Results.

LEMMA 3.1. *Given a unitary matrix U with $M(U) \neq 0$. Let F be its F-matrix. Then $R'_Z(0) = M(U)tr(ZF)$ for any skew-hermitian matrix Z.*

LEMMA 3.2. *Given an $n \times n$ zero-diagonal matrix W. If for every $n \times n$ skew-hermitian matrix Z, $tr(ZW) = 0$ then W is the zero-matrix.*

LEMMA 3.3. *Given a boundary matrix U with $M(U) \neq 0$ and with F-matrix $F \neq 0$. Given there's a unique tangent line L to Δ at $M(U)$ with direction vector v. Then for every skew-hermitian matrix Z, $tr(ZF) = cv$ where c is some real number.*

THEOREM 3.4. *Given a boundary matrix U with $M(U) \neq 0$ and with F-matrix $F \neq 0$. Given there's a unique tangent line to Δ at $M(U)$. Then F can be written uniquely in the form $F = e^{i\theta}H$ where H is a zero-diagonal hermitian matrix and $0 \leq \theta < \pi$.*

THEOREM 3.5. *Given a boundary matrix U with $M(U) \neq 0$ and with F-matrix $F \neq 0$. Given there's a unique tangent line L to Δ at $M(U)$. By the previous theorem we know that $F = e^{i\theta}H$ for some real $0 \leq \theta < \pi$. Then L makes an angle $arg(M(U)) + \theta + \pi/2$ with the positive real axis.*

4. Proof of Lemma 3.1. The proof given here uses ideas from [1], Theorem 4, p.26-27. But the proof given here is complete on its own.

Proof. We're given a unitary matrix U where $M(U) \neq 0$. So its F-matrix is well-defined and we call it F. Let B be its B-matrix, and C be its C-matrix. Given an arbitrary skew-hermitian matrix Z.

We can use Jacobi's formula [5] on (2.3) to find $R'_Z(t)$

$$R'_Z(t) = tr(det(C_Z(t))C_Z^{-1}(t)C'_Z(t)) \tag{4.1}$$

$R'_Z(0) = tr(det(C_Z(0))C_Z^{-1}(0)C'_Z(0))$

We can substitute C for $C_Z(0)$.

$R'_Z(0) = tr(det(C)C^{-1}C'_Z(0))$

$R'_Z(0) = det(C)tr(C^{-1}C'_Z(0))$

We know that $C'_Z(t) = B'_Z(t)$ so

$R'_Z(0) = det(C)tr(C^{-1}B'_Z(0))$

By subsection 2.1 and (1.3) we know that $det(C) = M(U)$

$$R'_Z(0) = M(U)tr(C^{-1}B'_Z(0)) \tag{4.2}$$

Using (2.2),

$$B'_Z(t) = \frac{dU_Z(t)}{dt}B_0U^*_Z(t) + U_Z(t)B_0\frac{dU^*_Z(t)}{dt} \tag{4.3}$$

Using (2.1),

$\frac{dU_Z(t)}{dt} = Ze^{Zt}U$

$U^*_Z(t) = (U^*)e^{-Zt}$

$\frac{dU^*_Z(t)}{dt} = -(U^*)Ze^{-Zt}$

Substitute these and (2.1) into (4.3)

$B'_Z(t) = Ze^{Zt}UB_0(U^*)e^{-Zt} - (e^{Zt})UB_0(U^*)Ze^{-Zt}$

$B'_Z(0) = ZUB_0U^* - UB_0(U^*)Z$

Using the definition of the C-matrix in subsection 2.1

$B'_Z(0) = Z(C - A_0) - (C - A_0)Z$

$B'_Z(0) = ZC - ZA_0 - CZ + A_0Z$

$C^{-1}B'_Z(0) = C^{-1}ZC - C^{-1}ZA_0 - Z + C^{-1}A_0Z$

$tr(C^{-1}B'_Z(0)) = tr(C^{-1}ZC) - tr(C^{-1}ZA_0) - tr(Z) + tr(C^{-1}A_0Z)$

The first and third terms cancel since similar matrices have the same trace.

$tr(C^{-1}B'_Z(0)) = -tr(C^{-1}ZA_0) + tr(C^{-1}A_0Z)$.

Using the idea that $tr(XY) = tr(YX)$

$tr(C^{-1}B'_Z(0)) = -tr(ZA_0C^{-1}) + tr(ZC^{-1}A_0)$

$tr(C^{-1}B'_Z(0)) = tr(ZC^{-1}A_0) - tr(ZA_0C^{-1})$

$tr(C^{-1}B'_Z(0)) = tr(Z(C^{-1}A_0 - A_0C^{-1}))$

$tr(C^{-1}B'_Z(0)) = tr(ZF)$

Substitute this into (4.2) to get

$$R'_Z(0) = M(U)tr(ZF) \tag{4.4}$$

This proves Lemma 3.1. ∎

5. Proof of Lemma 3.2.

Proof. Given an $n \times n$ zero-diagonal matrix W. Given that for all $n \times n$ skew-hermitian matrices Z, $tr(ZW) = 0$.

We can write element $W_{ab} = W_{ab,r} + iW_{ab,i}$, where $W_{ab,r}$ and $W_{ab,i}$ are real. These aren't tensors. $W_{ab,r}$ just denotes the real component of W_{ab} and $W_{ab,i}$ denotes the imaginary component.

$tr(Z^{ab}W) = 0$.

$tr(Z^{ab,i}W) = 0$

(See subsection 2.3 for definitions of Z^{ab} and $Z^{ab,i}$).

by direct computation we see that

$tr(Z^{ab}W) = (W_{ab,r} - W_{ba,r}) + i(W_{ab,i} - W_{ba,i}) = 0$

$tr(Z^{ab,i}W) = (-W_{ab,i} - W_{ba,i}) + i(W_{ab,r} + W_{ba,r}) = 0$

Solving these, we get that $W_{ab} = 0$. This is true for every pair (a,b) where $1 \leq a, b \leq n$ and $a \neq b$. So all the off-diagonal elements of W are zero. Hence W is the zero-matrix.

6. Proof of Lemma 3.3.

Proof. Given a boundary matrix U with $M(U) \neq 0$ and with F-matrix $F \neq 0$. Given there's a unique tangent line L to Δ at $M(U)$. Let v be the direction vector of the line L. Note that v is just a non-zero complex number.

Let Z be a skew-hermitian matrix. By Lemma 3.1 we know that $R'_Z(0) = M(U)tr(ZF)$.

Since $R_Z(t) \subseteq \Delta$ and $R_Z(0) = M(U)$, we know that $R'_Z(0) = kv$ for some real number k. (if L is the unique tangent to the region Δ at $M(U)$, then it must the tangent to every curve that lies in Δ and goes through $M(U)$ and has a well-defined derivative at $M(U)$).

So, $M(U)tr(ZF) = kv$

$tr(ZF) = (\frac{k}{M(U)})v$ ∎

7. Proof of Theorem 3.4.

Proof. Given a boundary matrix U with $M(U) \neq 0$ and with F-matrix $F \neq 0$. Given there's a unique tangent line to Δ at $M(U)$.

We pick an arbitrary pair $\{a, b\}$ such that $1 \leq a, b \leq n$ and $a \neq b$

We have two skew-hermitian matrices Z^{ab} and $Z^{ab,i}$ defined as per subsection 2.3.

By direct computation we see that

$tr(Z^{ab}F) = F_{ab} - F_{ba}$

$tr(Z^{ab,i}F) = (F_{ab} + F_{ba})i$

Given $F_{ab} = F_{ab,r} + iF_{ab,i}$. We can substitute this in to get

$$tr(Z_{ab}F) = (F_{ab,r} - F_{ba,r}) + i(F_{ab,i} - F_{ba,i}) \tag{7.1}$$

$$tr(Z_{ab,i}F) = (-F_{ab,i} - F_{ba,i}) + i(F_{ab,r} + F_{ba,r}) \tag{7.2}$$

We know by Lemma 3.3 that these are collinear vectors in the complex plane.

So we know that

$(F_{ab,i} - F_{ba,i})(-F_{ab,i} - F_{ba,i}) = (F_{ab,r} + F_{ba,r})(F_{ab,r} - F_{ba,r})$

We can simplify this to get:

$F_{ab,r}^2 + F_{ab,i}^2 = F_{ba,r}^2 + F_{ba,i}^2$

$|F_{ab}| = |F_{ba}|$

We can write:

$F_{ab} = |F_{ab}|\angle\theta_{ab}$

$F_{ba} = |F_{ab}|\angle\theta_{ba}$

For the remainder of the proof we will divide the possibilities for F into multiple cases. Note that we are given that $F \neq 0$. First we split all cases into two. The first is when only one pair of elements of the F-matrix, F_{ab} and F_{ba} is nonzero. The second case is when multiple pairs of elements of the F-matrix are nonzero. We shall further subdivide the second case using the fact that all tr(ZF) values are collinear. We can divide these cases into 3 possibilities: 1. All nonzero $tr(ZF)$ values are imaginary. 2. All nonzero $tr(ZF)$ values are real. 3. All nonzero $tr(ZF)$ values are not real or imaginary. (note that since F is nonzero, we don't have to deal with the possibility that tr(ZF) is 0 for all skew-hermitian matrices Z).

So we have 4 cases to deal with. Note that we already know by subsection 2.1 that F is zero-diagonal.

Case 1: $|F_{ab}|$ is non-zero for only one pair $\{a, b\}$ where $a \neq b$

In this case,

$H = e^{-(\theta_{ab}+\theta_{ba})/2}F$ is a hermitian matrix, and we're finished.

Case 2: $|F_{ab}|$ is non-zero for multiple pairs $\{a, b\}$ where $a \neq b$. For any skew-hermitian Z, when tr(ZF) is non-zero, it is imaginary.

If $|F_{ab}| \neq 0$, then by (7.1) and (7.2), $\theta_{ab} = -\theta_{ba}$. This holds for all distinct pairs {a,b}, so our F-matrix is already hermitian, and we're done.

Case 3: $|F_{ab}|$ is non-zero for multiple pairs $\{a, b\}$ where $a \neq b$. For any skew-hermitian Z, when tr(ZF) is non-zero, it is real.

If $|F_{ab}| \neq 0$, then by (7.1) and (7.2), $\theta_{ab} = \pi - \theta_{ba}$. This holds for all distinct pairs {a,b}

$H = e^{-(\frac{\pi}{2})}F$ is hermitian and we're done.

Case 4: $|F_{ab}|$ is non-zero for multiple pairs $\{a,b\}$ where $a \neq b$. For any skew-hermitian matrix Z, when tr(ZF) is non-zero, it isn't real or imaginary.

Suppose $|F_{ab}| \neq 0$ and $|F_{cd}| \neq 0$

if $tr(Z_{ab}F) \neq 0$, then

slope of $tr(Z_{ab}F) = \frac{\sin(\theta_{ab})-\sin(\theta_{ba})}{\cos(\theta_{ab})-\cos(\theta_{ba})} = -\cot(\frac{\theta_{ab}+\theta_{ba}}{2})$

if $tr(Z_{ab,i}F) \neq 0$:

slope of $tr(Z_{ab,i}F) = \frac{\cos(\theta_{ab})+\cos(\theta_{ba})}{-\sin(\theta_{ab})-\sin(\theta_{ba})} = -\cot(\frac{\theta_{ab}+\theta_{ba}}{2})$

We know that since $|F_{ab}| \neq 0$, at least one of $tr(Z_{ab}F)$ or $tr(Z_{ab,i}F)$ is non-zero.

similarly,

if $tr(Z_{cd}F) \neq 0$, then

slope of $tr(Z_{cd}F) = -\cot(\frac{\theta_{cd}+\theta_{dc}}{2})$

if $tr(Z_{cd,i}F) \neq 0$:

slope of $tr(Z_{cd,i}F) = -\cot(\frac{\theta_{cd}+\theta_{dc}}{2})$

We know that since $|F_{cd}| \neq 0$, at least one of $tr(Z_{cd}F)$ or $tr(Z_{cd,i}F)$ is non-zero.

So we have:

$\cot(\frac{\theta_{cd}+\theta_{dc}}{2}) = \cot(\frac{\theta_{ab}+\theta_{ba}}{2})$ (Lemma 3.3)

therefore:

$\frac{\theta_{cd}+\theta_{dc}}{2} = \frac{\theta_{ab}+\theta_{ba}}{2} + n\pi$ for some integer n.

We can freely adjust θ_{cd} by $-2n\pi$. It makes no difference since $|F_{cd}|\angle\theta_{cd} = |F_{cd}|\angle(\theta_{cd} - 2n\pi)$

So after the adjustment we have:

$\frac{\theta_{cd}+\theta_{dc}}{2} = \frac{\theta_{ab}+\theta_{ba}}{2}$.

We make the same adjustment for any pair $\{c,d\} \neq \{a,b\}$ where $|F_{cd}| \neq 0$

We set $\beta = \frac{\theta_{ab}+\theta_{ba}}{2}$

let $H = e^{-i\beta}F$

For some pair $\{x,y\}$ where $x \neq y$ and $|H_{xy}| \neq 0$,

$H_{xy} = |H_{xy}|\angle\alpha_{xy}$

$\alpha_{xy} = -(\frac{\theta_{ab}+\theta_{ba}}{2}) + \theta_{xy}$

$\alpha_{yx} = -(\frac{\theta_{ab}+\theta_{ba}}{2}) + \theta_{yx}$

But because of our adjustments,

$\frac{\theta_{ab}+\theta_{ba}}{2} = \frac{\theta_{xy}+\theta_{yx}}{2}$

Plugging this into the above two formulas we have

$\alpha_{xy} = \frac{\theta_{xy}-\theta_{yx}}{2}$

$\alpha_{yx} = -(\frac{\theta_{xy}-\theta_{yx}}{2})$

Therefore H is zero-diagonal, with transpositional elements of equal magnitude and opposite arguments. Therefore H is hermitian.

So in all 4 cases we can write $F = e^{i\beta}H$ for some hermitian matrix H and some real β. But we've not arrived at a unique representation for F yet.

Suppose

$F = e^{i\beta_1}H_1 = e^{i\beta_2}H_2$

$e^{i(\beta_1-\beta_2)}H_1 = H_2$

$e^{i(\beta_1-\beta_2)}H_1 = H_2 = H_2^* = e^{i(\beta_2-\beta_1)}H_1^* = e^{i(\beta_2-\beta_1)}H_1$

So

$(e^{i(\beta_1-\beta_2)} - e^{i(\beta_2-\beta_1)})H_1 = 0$

Since $F \neq 0$, we know $H_1 \neq 0$ so

$e^{i(\beta_1-\beta_2)} - e^{i(\beta_2-\beta_1)} = 0$

$e^{i(\beta_1-\beta_2)} = e^{i(\beta_2-\beta_1)}$

Then

$\beta_1 - \beta_2 = \beta_2 - \beta_1 + 2k\pi$, for any integer k

$\beta_1 = \beta_2 + k\pi$

So if we restrict all β to $0 \leq \beta < \pi$, we have a unique representation since k is forced to 0.

This completes our proof of Theorem 3.4. ∎

8. Proof of Theorem 3.5. Given an ordinary boundary matrix U with $M(U) \neq 0$ and F-matrix $F \neq 0$. Given $\partial\Delta$ has the unique tangent line L at $M(U)$.

Proof. By Theorem 3.4 we know that

$$F = e^{i\theta}H \tag{8.1}$$

for some real $0 \leq \theta < \pi$ and some zero-diagonal hermitian matrix H.

We can substitute (8.1) into (7.1) and (7.2) and simplify to get:

$$tr(Z_{ab}F) = 2H_{ab,i}e^{i(\theta+\pi/2)} \tag{8.2}$$

$$tr(Z_{ab,i}F) = 2H_{ab,r}e^{i(\theta+\pi/2)} \tag{8.3}$$

By Lemma 3.2 we know that at least one of the above equations is nonzero for some pair $\{a, b\}$. So then using Lemma 3.1 we know that $R'_Z(0) = M(U)tr(ZF) \neq 0$ for some skew-hermitian matrix Z.

So by (8.2) and (8.3) we see that for some skew-hermitian matrix Z, $tr(ZF)$ forms an angle of $(\theta+\pi/2)$ or $(\theta+3\pi/2)$ with the positive real axis (depending on whether the coefficient is negative or not). Therefore $R'_Z(0)$ forms an angle $arg(M(U)) + \theta + \pi/2$ or $arg(M(U)) + \theta + 3\pi/2$ with the positive real axis.

Therefore the line L forms an angle $arg(M(U)) + \theta + \pi/2$ with the positive real axis (since this is a line as opposed to a vector, a rotation of π makes no difference).

This completes our proof of Theorem 3.5. ∎

REFERENCES

[1] N. Bebiano and J. Querió, *The determinant of the sum of two normal matrices with prescribed eigenvalues*, Linear Algebra and its Applications, 71 (1985), pp. 23–28.
[2] G. N. de Oliveira, *Research problem: Normal matrices*, Linear and Multilinear Algebra, 12 (1982), pp. 153–154.
[3] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.
[4] M. Marcus, *Derivations, plücker relations and the numerical range*, Indiana University Math Journal, 22 (1973), pp. 1137–1149.
[5] Wikipedia contributors, *Jacobi's formula — Wikipedia, the free encyclopedia*, 2018, https://en.wikipedia.org/w/index.php?title=Jacobi%27s_formula&oldid=838689851. [Online; accessed 16-September-2018].