ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Linear Algebra and its Applications 422 (2007) 318–325

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

OMC for scalar multiples of unitaries

S.W. Drury [*,1]

Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 ouest, rue Sherbrooke Street West, Montréal, Que., Canada H3A 2K6

Received 29 June 2006; accepted 22 October 2006
Available online 1 December 2006
Submitted by S. Fallat

Abstract

We establish the following case of the Determinantal Conjecture of Marcus [M. Marcus, Derivations, Plücker relations and the numerical range, Indiana Univ. Math. J. 22 (1973) 1137–1149] and de Oliveira [G.N. de Oliveira, Research problem: normal matrices, Linear and Multilinear Algebra 12 (1982) 153–154]. Let A and B be unitary $n \times n$ matrices with prescribed eigenvalues $a_1, \dots, a_n$ and $b_1, \dots, b_n$, respectively. Then for any scalars t and s

$$\det(tA - sB) \in \operatorname{co}\left\{\prod_{j=1}^{n}(ta_j - sb_{\sigma(j)}); \sigma \in S_n\right\},$$

where S_n denotes the group of all permutations of $\{1, \dots, n\}$ and co the convex hull taken in the complex plane.

AMS classification: 15A15; 15A42; 15A57

Keywords: Determinant; Eigenvalue; Marcus–de Oliveira conjecture

* Tel.: +1 514 398 3830; fax: +1 514 398 3899.
E-mail address: drury@math.mcgill.ca

1 Research supported in part by the Natural Sciences and Engineering Research Council of Canada. Grant RGPIN 8548-96.

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.10.022

The celebrated Determinantal Conjecture of Marcus [5] and de Oliveira [6] can be stated as follows.

Conjecture 1. *Let A and B be normal $n \times n$ matrices with prescribed complex eigenvalues $a_1, \dots, a_n$ and $b_1, \dots, b_n$, respectively. Let Δ_0 be the subset of $\mathbb{C}$ given by $\Delta_0 = \mathrm{co}\big\{ \prod_{j=1}^n (a_j - b_{\sigma(j)}); \sigma \in S_n \big\}$. Then*

$$\det(A - B) \in \Delta_0,$$

where S_n denotes the group of all permutations of $\{1, \dots, n\}$ and co *the convex hull taken in the complex plane.*

The purpose of this paper is to establish the conjecture in case A and B are scalar multiples of unitary matrices.

Theorem 2. *Conjecture 1 holds in case A and B are scalar multiples of unitary matrices.*

Conjecture 1 is known in a great many special cases. The case in which A and B are both hermitian was settled by Fiedler [3]. The case in which A is positive definite and B is skew-hermitian was settled by da Providência and Bebiano [7]. The proof of Theorem 2 borrows many ideas from that paper. The case in which A and B are both unitary was settled by Bebiano and da Providência [1], the key observation being that Conjecture 1 is unchanged under a simultaneous application of a fractional linear (Möbius) transformation of the eigenvalues a_j and b_j, allowing a reduction to the case in which A and B are hermitian. The same idea does not work in case that the a_j and b_j lie in different concentric circles.

Theorem 2 has the following corollary.

Corollary 3. *Let C_A and C_B be circles in the complex plane that do not intersect. Let $a_1, \dots, a_n \in C_A$ and $b_1, \dots, b_n \in C_B$, then Conjecture 1 holds.*

Proof. Any pair of non-intersecting circles in $\mathbb{C}$ can be mapped by some Möbius transformation φ to a pair of concentric circles centred at the origin. Using the invariance under Möbius transformations [1, Theorem 2], it is equivalent to replace A and B with $\varphi(A)$ and $\varphi(B)$, respectively. But $\varphi(A)$ and $\varphi(B)$ are scalar multiples of unitary matrices and Theorem 2 establishes the result in that case. $\square$

We start by stating some ideas extracted from [7,2].

Let Δ be a closed bounded subset of $\mathbb{C}$. Let z be an extreme point of $\mathrm{co}(\Delta)$ which therefore necessarily lies in Δ. We will say that z is *almost flat* if there is a smooth curve segment passing through z, lying entirely inside Δ and having zero curvature at z.

Lemma 4. *The set Δ is contained in the closed convex hull of those extreme points of* $\mathrm{co}(\Delta)$ *that are not almost flat.*

Proof. Let N be the set of extreme points of $\mathrm{co}(\Delta)$ which are not almost flat. Let K be the closed convex hull of N. We claim that $\mathrm{co}(\Delta) \subseteq K$. Suppose not, then there is a point z_1 in $\mathrm{co}(\Delta)$ not in K. But now it is possible to find a closed disk B with $z_1 \notin B$ but $K \subseteq B$ see Valentine [8]. Let w be the centre of this disk. Let z be a furthest point of $\mathrm{co}(\Delta)$ from w. Then certainly $z \notin B$.

There can be no line segment lying in $\mathrm{co}(\Delta)$ and passing through z (for otherwise there would be points of $\mathrm{co}(\Delta)$ further from w than z). So z is an extreme point of $\mathrm{co}(\Delta)$ and hence $z \in \Delta$ and it does not lie in $N \subseteq K \subseteq B$. Therefore, z is an almost flat point of Δ. But then there is a curve segment, passing through z lying in Δ and with zero curvature at z. This contradicts the definition of z since then there must be points of $\mathrm{co}(\Delta)$ further from w than z. □

Lemma 5. *Let A and B be normal matrices and P skew-hermitian. Let $T = A - B$ and consider a variation $T(t) = \exp(tP)A\ \exp(-tP) - B$. Then the expansion*

$$\frac{\det(T(t))}{\det(T)} = 1 + u_1 t + u_2 t^2 + \cdots$$

is valid about $t = 0$ where $u_1 = \mathrm{tr}(T^{-1}[P, A]) = \mathrm{tr}(T^{-1}[P, B])$ and

$$\begin{aligned} u_2 &= \frac{1}{2}\left(\mathrm{tr}(T^{-1}[P, A])\mathrm{tr}(T^{-1}[P, B]) - \mathrm{tr}(T^{-1}[P, A]T^{-1}[P, B])\right) \\ &= \mathrm{tr}(T^{-1}[P, A] \wedge T^{-1}[P, B]). \end{aligned} \tag{1}$$

We note that for S an operator on an inner product space E, the operator $S \wedge S$ is defined on the inner product space $E \wedge E$ by extending $(S \wedge S)(v_1 \wedge v_2) = Sv_1 \wedge Sv_2$ by linearity. This definition is further extended to $S_1 \wedge S_2$ for possibly different operators S_1 and S_2 on E either by the polarization identity

$$S_1 \wedge S_2 = \frac{1}{4}(S_1 + S_2) \wedge ((S_1 + S_2) - (S_1 - S_2) \wedge (S_1 - S_2))$$

or by $(S_1 \wedge S_2)(v_1 \wedge v_2) = \frac{1}{2}(S_1 v_1 \wedge S_2 v_2 + S_2 v_1 \wedge S_1 v_2)$. We remark in particular that if S is a rank one operator, then $S \wedge S$ is zero. We refer the reader to Lang [4] for issues relating to exterior products.

Lemma 6. *Let R be a non-zero $n_1 \times n_2$ matrix. Then there exists a $n_1 \times n_2$ matrix G such that $GR^\star + RG^\star$ and $R^\star G + G^\star R$ are both rank one matrices with non-zero trace.*

Proof. By an application of the singular value decomposition, it suffices to prove the result in the special case that R is rectangular diagonal with non-negative diagonal entries. Let the diagonal values be $r_j \geqslant 0$ for $j = 1, 2, \dots, m$ where $m = \min(n_1, n_2)$. Now define $g_{jk} = 1/(r_j + r_k)$ for $1 \leqslant j, k \leqslant m$ and $r_j > 0$ and $r_k > 0$ and $g_{jk} = 0$ in all other cases. Then G is non-zero and the result is evident. □

We now specialize to the case

$$\Delta = \{\det(U^\star \mathrm{diag}(a_1, \dots, a_n)U - V^\star \mathrm{diag}(b_1, \dots, b_n)V);\ U, V \text{ unitary}\}$$

of interest in this article by proving the following.

Proposition 7. *Let $n \geqslant 2$ and $0 < t < 1$. Let A be an $n \times n$ unitary matrix and B an $n \times n$ matrix with $t^{-1}B$ unitary. Suppose that A and B have no common non-trivial invariant linear subspace and that $\det(A - B)$ is an extreme point of $\mathrm{co}(\Delta)$. Let X denote the unitary matrix $t^{-1}BA^{-1}$. Then X has exactly two distinct eigenvalues ω_1 and ω_2 given by*

$$\omega_1, \omega_2 = \frac{\mathrm{i}(2\sin(\theta) + \mu(1 + t^2)) \pm \sqrt{4t^2 - (2\sin(\theta) + \mu(1 - t^2))^2}}{2t(\mathrm{i}\mu + \omega)}, \tag{2}$$

where μ is a real constant and $\omega = e^{i\theta}$ is a complex number of absolute value 1 such that $|2\sin(\theta) + \mu(1-t^2)| < 2t$. Further, we may write $X = \omega_1 E_1 + \omega_2 E_2$ where E_1 and E_2 are complementary orthogonal projections arising from the spectral decomposition of X.

Proof. If $\xi \in \ker(A-B)$, then $A\xi = B\xi$ and $\|\xi\| = \|A\xi\| = \|B\xi\| = t\|\xi\|$ a contradiction. So $T = A - B$ is always invertible. Since $\det(A-B)$ is an extreme point of $\mathrm{co}(\Delta)$, it possesses a supporting hyperplane. We choose ω to be a complex number of unit modulus such that the direction $\omega\det(A-B)$ is normal to this hyperplane. It is now clear that for every choice of skew-hermitian P the function

$$t \mapsto \Re\bar{\omega}\frac{\det(T(t))}{\det(T)}$$

has a critical point at $t = 0$. Consequently, $\Re\bar{\omega}u_1 = 0$. Thus, for all P skew-hermitian, we have

$$\Re\bar{\omega}\,\mathrm{tr}(P[T^{-1}, A]) = 0$$

or equivalently $H = \bar{\omega}[T^{-1}, A]$ is hermitian. Letting $X = t^{-1}BA^{-1}$, $Y = X^\star = t^{-1}A^{-1}B = A^{-1}XA$ we may rewrite $H = H^\star$ as

$$\bar{\omega}\big((I - tX)^{-1} - (I - tY)^{-1}\big) = \omega\big((I - tX^{-1})^{-1} - (I - tY^{-1})^{-1}\big),$$

or

$$\begin{aligned}\bar{\omega}(I - tX)^{-1} - \omega(I - tX^{-1})^{-1} &= \bar{\omega}(I - tY)^{-1} - \omega\big(I - tY^{-1}\big)^{-1}\\ &= A^{-1}\big(\bar{\omega}(I - tX)^{-1} - \omega(I - tX^{-1})^{-1}\big)A\end{aligned}$$

so that,

$$\big[A, \bar{\omega}(I - tX)^{-1} - \omega(I - tX^{-1})^{-1}\big] = 0.$$

We see that the matrix $\bar{\omega}(I - tX)^{-1} - \omega(I - tX^{-1})^{-1}$ commutes with A and with X are therefore also with B. Any eigenspace of this operator is necessarily invariant under both A and B. It follows that the skew-hermitian operator $\bar{\omega}(I - tX)^{-1} - \omega(I - tX^{-1})^{-1} = i\mu I$ for some suitable real μ. Thus X satisfies the quadratic equation

$$\bar{\omega}(X - tI) - \omega(X - tX^2) = i\mu\big(-tI + (1+t^2)X - tX^2\big) \tag{3}$$

or equivalently

$$t(i\mu + \omega)X^2 - i(2\sin(\theta) + \mu(1+t^2))X + t(i\mu - \bar{\omega})I = 0. \tag{4}$$

Since X is a unitary, we can therefore write

$$X = \omega_1 E_1 + \omega_2 E_2,$$

where E_1 and E_2 are complementary orthogonal projections and ω_1 and ω_2 are complex numbers of absolute value 1 and the roots of the Eq. (3) are given by (2). The quantity under the square root in (2) is non-negative if and only if the roots have equal absolute value. Thus we have

$$|2\sin(\theta) + \mu(1-t^2)| < 2t.$$

Equal roots are not allowed since then X is a scalar multiple of the identity, A and B commute and possess a common non-trivial invariant subspace since $n \geqslant 2$. $\square$

We now tackle the second-order term.

Proposition 8. *With the hypotheses and notations of Proposition 7, we obtain*

$$\Re\bar{\omega}\big(u_2 - C_1\operatorname{tr}((E_1Z)\wedge(E_1Z)) - C_2\operatorname{tr}((E_2Z)\wedge(E_2Z))\big) = 0,$$

where $Q = APA^{-1}$ and $Z = P - Q$ are skew-hermitian matrices for suitable scalars C_1 and C_2.

Proof. The first part of (1) involves

$$\begin{aligned}
&\operatorname{tr}\big(T^{-1}[P,A]\big)\operatorname{tr}\big(T^{-1}[P,B]\big)\\
&\quad= t\operatorname{tr}\big(A^{-1}(I-tX)^{-1}(PA-AP)\big)\operatorname{tr}\big(A^{-1}(I-tX)^{-1}(PXA-XAP)\big)\\
&\quad= t\operatorname{tr}((I-tX)^{-1}Z)\operatorname{tr}(X(I-tX)^{-1}Z).
\end{aligned}$$

We write $(I-tX)^{-1} = c_1E_1 + c_2E_2$ where $c_j = (1-t\omega_j)^{-1}$ are scalars and we obtain similarly $tX(I-tX)^{-1} = (c_1-1)E_1 + (c_2-1)E_2$. Splitting into main and cross terms we get

$$\operatorname{tr}\big(T^{-1}[P,A]\big)\operatorname{tr}\big(T^{-1}[P,B]\big) = \alpha_1 + \beta_1,$$

where

$$\alpha_1 = c_1(c_1-1)(\operatorname{tr}(E_1Z))^2 + c_2(c_2-1)(\operatorname{tr}(E_2Z))^2$$

and

$$\beta_1 = (c_1(c_2-1) + c_2(c_1-1))\operatorname{tr}(E_1Z)\operatorname{tr}(E_2Z). \tag{5}$$

For this term, the key point is that

$$\begin{aligned}
c_1(c_2-1)+c_2(c_1-1) &= \frac{t(\omega_1+\omega_2)}{1-t(\omega_1+\omega_2)+t^2\omega_1\omega_2}\\
&= \frac{\mathrm{i}(2\sin(\theta)+\mu(1+t^2))}{(\mathrm{i}\mu+\omega)-\mathrm{i}(2\sin(\theta)+\mu(1+t^2))+t^2(\mathrm{i}\mu-\bar{\omega})}\\
&= \mathrm{i}\omega\frac{2\sin(\theta)+\mu(1+t^2)}{1-t^2}
\end{aligned}$$

using expressions for $\omega_1+\omega_2$ and $\omega_1\omega_2$ obtained from (4) and using $2\mathrm{i}\sin(\theta) = \omega - \bar{\omega}$. It follows that $(c_1(c_2-1)+c_2(c_1-1))$ is a pure imaginary multiple of ω and since the two remaining factors in (5) are pure imaginary, it follows that the term β_1 is purely tangential (i.e. a real multiple of u_1).

The second part of (1) involves

$$\begin{aligned}
&\operatorname{tr}\big(T^{-1}[P,A]T^{-1}[P,B]\big)\\
&\quad= t\operatorname{tr}\big(A^{-1}(I-tX)^{-1}(PA-AP)A^{-1}(I-tX)^{-1}(PXA-XAP)\big)\\
&\quad= t\operatorname{tr}(X(I-tX)^{-1}P(I-tX)^{-1}P) - 2t\operatorname{tr}\big(X(I-tX)^{-1}Q(I-tX)^{-1}P\big)\\
&\qquad+ t\operatorname{tr}\big(X(I-tX)^{-1}Q(I-tX)^{-1}Q\big)\\
&\quad= \alpha_2 + \beta_2,
\end{aligned}$$

where

$$\begin{aligned}\alpha_2 &= c_1(c_1-1)\mathrm{tr}(E_1PE_1P - 2E_1QE_1P + E_1QE_1Q)\\ &\quad + c_2(c_2-1)\mathrm{tr}(E_2PE_2P - 2E_2QE_2P + E_2QE_2Q)\\ &= c_1(c_1-1)\mathrm{tr}((E_1Z)^2) + c_2(c_2-1)\mathrm{tr}((E_2Z)^2)\end{aligned}$$

and

$$\begin{aligned}\beta_2 &= (c_1(c_2-1) + c_2(c_1-1))(\mathrm{tr}(E_1PE_2P) + \mathrm{tr}(E_1QE_2Q))\\ &\quad - 2c_1(c_2-1)\mathrm{tr}(E_2QE_1P) - 2c_2(c_1-1)\mathrm{tr}(E_1QE_2P)\end{aligned}$$

and we claim that β_2 is again purely tangential. First we remark that

$$\mathrm{tr}(E_1PE_2P) = \mathrm{tr}(E_1P(I-E_1)P) = \mathrm{tr}(E_1P^2) - \mathrm{tr}\big((E_1PE_1)^2\big)$$

and similarly $\mathrm{tr}(E_1QE_2Q)$ are real, that $\mathrm{tr}(E_2QE_1P)$ and $\mathrm{tr}(E_1QE_2P)$ are complex conjugates and that

$$c_1(c_2-1) - c_2(c_1-1) = \pm\frac{\omega\sqrt{4t^2 - (2\sin(\theta) + \mu(1-t^2))^2}}{t^2-1}$$

is aligned normally. Therefore, the second-order term can be written

$$c_1(c_1-1)\mathrm{tr}((E_1Z)\wedge(E_1Z)) + c_2(c_2-1)\mathrm{tr}((E_2Z)\wedge(E_2Z)) + \frac{1}{2}(\beta_1-\beta_2)$$

with $\beta_1-\beta_2$ purely tangential. Hence the result. □

Proposition 9. *With the hypotheses and notations of Proposition 7 and the additional hypothesis that -1 is not an eigenvalue of A, it follows that* $\det(A-B)$ *is an almost flat extreme point of Δ.*

Proof. Let S be the skew-hermitian logarithm of A with eigenvalues $\mathrm{i}\phi_j$ for $j = 1, 2, \ldots, n$ where $-\pi < \phi_j < \pi$ for all j. We use the standard notation $\mathrm{ad}_S(Q) = SQ - QS$ to denote the adjoint representation. We view ad_S as a linear operator on the space of $n\times n$ matrices. Then ad_S is also skew-hermitian and has eigenvalues $\mathrm{i}(\phi_j - \phi_k)$ for the n^2 values given by $1 \leqslant j, k \leqslant n$. All of these eigenvalues lie within distance 2π from the origin.

We write S in terms of an orthonormal basis $(e_j)_{j=1}^n$ of V which respects the orthogonal sum $V = V_1 \oplus V_2$ corresponding to the projections E_1 and E_2:

$$S = \begin{pmatrix} S_{11} & S_{12}\\ S_{21} & S_{22}\end{pmatrix},$$

where S_{11} and S_{22} are skew-hermitian and $S_{21} = -S_{12}^\star$. We have that S_{12} is non-zero for otherwise S and X (and therefore also A and B) have a non-trivial invariant subspace which is impossible. Applying Lemma 6 to $R = S_{12}$ we construct the skew-hermitian matrix

$$W = \begin{pmatrix} 0 & \mathrm{i}G\\ \mathrm{i}G^\star & 0\end{pmatrix}$$

and it follows that

$$[S, W] = Z = \begin{pmatrix} Z_{11} & Z_{12}\\ -Z_{12}^\star & Z_{22}\end{pmatrix},$$

where $Z_{11} = E_1ZE_1$ and $Z_{22} = E_2ZE_2$ are non-zero rank one skew-hermitian matrices.

Now define $P = \psi(\mathrm{ad}_S)W$, where $\psi(z) = \frac{z}{1-\mathrm{e}^z}$ has a power series in z with radius of convergence 2π. It follows that

$$P - APA^{-1} = P - \mathrm{e}^S P \mathrm{e}^{-S} = \left(I - \mathrm{e}^{\mathrm{ad}_S}\right)P = \mathrm{ad}_S W = Z,$$

using the standard relation between the adjoint representations of $GL(n, \mathbb{C})$ and its Lie algebra, see Varadarajan [9, Theorem 2.13.2] for details. For this choice of P, the quadratic term is purely tangential and the linear term is non-zero since

$$\begin{aligned}\mathrm{tr}(T^{-1}[P, A]) &= \mathrm{tr}((I - tX)^{-1}Z) = \mathrm{tr}((c_1E_1 + c_2E_2)Z)\\ &= (c_1 - c_2)\mathrm{tr}(E_1Z) = (c_1 - c_2)\mathrm{tr}(Z_{11}) \neq 0\end{aligned}$$

since $c_1 \neq c_2$ and $\mathrm{tr}(Z_{11}) \neq 0$. Thus there is a curve segment lying entirely in Δ with zero curvature at $\det(A - B)$. $\square$

Proof of Theorem 2. We prove the result by strong induction on n. For $n = 1$ and $n = 2$ the result is easy to verify by direct calculation.

Let $n \geqslant 3$ and $t, s \geqslant 0$. We suppose that $|a_1| = |a_2| = \cdots = |a_n| = s$ and $|b_1| = |b_2| = \cdots = |b_n| = t$. We consider Conjecture 1 in this case. If $t = 0$ or $s = 0$ the result is obvious. If $t = s$, then the result of [1] applies (or we may perturb t and s and apply a limiting argument). Rescaling the problem and interchanging A and B if necessary, we can suppose that $s = 1$ and that $0 < t < 1$. It suffices to suppose that the sets $\{-1, a_1, \ldots, a_n\}$ and $\{b_1, \ldots, b_n\}$ have distinct elements since the general case can be obtained from this one by perturbation and continuity.

Now suppose that $\Delta \not\subseteq \Delta_0$. Then it follows from Lemma 4 that there is an extreme point z of $\mathrm{co}(\Delta)$ which is not almost flat and such that $z \notin \Delta_0$. Let A and B be the corresponding matrices. Applying Propositions 7–9, we can conclude that A and B possess a common non-trivial invariant linear subspace. The orthogonal complement is also simultaneously invariant. The fact that the eigenvalue sets consist of distinct elements allows the matrices (or rather the corresponding operators) to be decomposed simultaneously on spaces of lower dimension. This allows a contradiction to be established from the strong induction hypothesis. Hence $\Delta \subseteq \Delta_0$ as required. $\square$

Acknowledgements

The author wishes to thank NSERC for financial support (Grant RGPIN 8548-06) and the referees for helpful comments.

References

[1] N. Bebiano, J. da Providência, Some remarks on a conjecture of de Oliveira, Linear Algebra Appl. 102 (1988) 241–246.

[2] N. Bebiano, A. Kovačec, J. da Providência, The validity of the Marcus–de Oliveira conjecture for essentially Hermitian matrices, in: Second Conference of the International Linear Algebra Society (ILAS) (Lisbon, 1992), Linear Algebra Appl., 197–198, 1994, pp. 411–427.

[3] Miroslav Fiedler, Bounds for the determinant of the sum of hermitian matrices, Proc. Amer. Math. Soc. 30 (1971) 27–31.

[4] S. Lang, Linear Algebra, second ed., Addison-Wesley, 1971.

[5] M. Marcus, Derivations, Plücker relations and the numerical range, Indiana Univ. Math. J. 22 (1973) 1137–1149.

[6] G.N. de Oliveira, Research problem: normal matrices, Linear and Multilinear Algebra 12 (1982) 153–154.

[7] J. da Providência, N. Bebiano, Matrices satisfying a conjecture of G.N. de Oliveira on determinants, Linear Algebra Appl. 78 (1986) 187–198.
[8] F.A. Valentine, Convex Sets, McGraw-Hill, New York, 1964.
[9] V.S. Varadarajan, Lie Groups, Lie Algebras and their Representations, Prentice-Hall, 1974.