1-2016

The von Neumann inequality for 3 × 3 matrices

Greg Knese
geknese@wustl.edu

THE VON NEUMANN INEQUALITY FOR 3×3 MATRICES

GREG KNESE

ABSTRACT. This note details how recent work of Kosiński on the three point Pick interpolation problem on the polydisc can be used to prove the von Neumann inequality for d-tuples of commuting 3×3 contractive matrices.

1. INTRODUCTION

The purpose of this note is to explain how recent results of Kosiński [14] provide a proof of the von Neumann inequality for d-tuples of 3×3 commuting contractive matrices.

Definition 1.1. A d-tuple of pairwise commuting contractive matrices or operators $T = (T_1, \dots, T_d)$ satisfies the *von Neumann inequality* if for every $p \in \mathbb{C}[z_1, \dots, z_d]$

$$\|p(T)\| \leq \sup_{z \in \mathbb{T}^d} |p(z)|.$$

Recall that contractive means the Hilbert space operator norm $\|T_j\| \leq 1$ and $\mathbb{T}^d$ is the unit d-torus in $\mathbb{C}^d$.

The von Neumann inequality holds for a single contractive operator—von Neumann's original result [16]—as well as for a pair of commuting contractions—a result of Andô [2]. For $d > 2$ there are known examples of d-tuples of commuting contractions for which the von Neumann inequality fails. Varopoulos [18] proved the existence of counterexamples with a probabalistic argument and later Kaijser and Varopoulos (see addendum to [18]) as well as Crabb and Davie [4] found explicit counterexamples. These counterexamples were all given by finite matrices. It turns out that the von Neumann inequality holds for d-tuples of 2×2 commuting contractive matrices; this result is essentially equivalent to

Date: September 17, 2015.

2010 *Mathematics Subject Classification.* 32E30, 47A63, 30E05.

Key words and phrases. von Neumann inequality, polydisc, polydisk, rational inner functions, Schur-Agler class, multi-variable operator theory, Schwarz lemma, Pick interpolation.

Partially supported by NSF grant DMS-1363239.

the Schwarz lemma on the polydisc. On the other hand, Holbrook [10] has found a 3-tuple of 4×4 matrices which fail the von Neumann inequality.

A great deal of effort has been expended to answer the following question: Does the von Neumann inequality hold for d-tuples of 3×3 commuting contractive matrices? For evidence of interest in this question see [3, 10, 15]. See also the thesis [8] which addressed some special cases of this question.

Recent work of Kosiński proves that the answer is yes.

Theorem 1.2. *The von Neumann inequality holds for d-tuples of 3×3 commuting contractive matrices.*

This is also interesting in light of the fact that there exists a 4-tuple of 3×3 commuting contractions that does not *dilate* to a commuting 4-tuple of unitary operators [3]. Possessing a unitary dilation is a much stronger property than satisfying a von Neumann inequality; in fact, it means that a von Neumann inequality holds for matrix-valued polynomials of all matrix sizes. Indeed, the 4-tuple of 3×3 matrices of [3] fails the von Neumann inequality for 2×1 matrix-valued polynomials. This leaves the following question unresolved:

> Does every 3-tuple T of 3×3 contractive matrices have a unitary dilation? Equivalently, for every such T and every matrix-valued polynomial $P = (p_{j,k}) \in \mathbb{C}^{M \times N}[z_1, z_2, z_3]$, do we have
> $$\|(p_{j,k}(T))_{j,k}\| \leq \sup\{\|P(z)\| : z \in \mathbb{D}^3\}?$$
> If not, what are the minimal matrix sizes $M \times N$ for which this does hold?

For some context, we point out that many of the von Neumann inequalities cited above have dilation counterparts: the Sz.-Nagy dilation theorem vis-à-vis von Neumann's inequality, Andô's dilation theorem (Andô's actual result in [2]), d-tuples of 2×2 commuting contractive matrices always dilate [6, 11]. The situation of a von Neumann inequality without a dilation theorem is not uncommon though, and it is closely related to the distinction between spectral sets and complete spectral sets. See [5] for more information.

In the next two sections, we point out some known reductions and then explain how Kosiński's work proves Theorem 1.2. In the third and final section, we explain some other operator theory related consequences of [14]. Namely, solvable three-point Pick interpolation problems on the polydisc can always be solved with a rational inner function in the Schur-Agler class.

2. Reductions

Let $T = (T_1, \dots, T_d)$ be a d-tuple of 3×3 commuting contractive matrices. To start with, we can assume the matrices are strict contractions ($\|T_j\| < 1$, $j = 1, \dots, d$) and then the theorem will follow by continuity. Next, a d-tuple of 3×3 commuting matrices can always be perturbed to a simultaneously diagonalizable d-tuple of commuting matrices. The reference [12] points out several places where this is proven; see [7, 9, 15].

After adjusting our operators to be nicer, we will replace polynomials with functions that are less nice. If T is now a d-tuple of commuting simultaneously diagonalizable strictly contractive 3×3 matrices, then it suffices to prove

$$\|f(T)\| \leq 1$$

for all $f : \mathbb{D}^d \to \mathbb{D}$ holomorphic on the unit d-dimensional polydisc. This follows from the fact that such holomorphic functions can be approximated locally uniformly on $\mathbb{D}^d$ by polynomials $p \in \mathbb{C}[z_1, \dots, z_d]$ with supremum norm at most 1 on $\mathbb{D}^d$ (or $\mathbb{T}^d$ by the maximum principle). See Rudin [17, pg. 126].

Using bounded holomorphic functions makes it possible to apply Möbius transformations to the matrices $T_1, \dots, T_d$ in order to force one of the joint eigenvalues of T to be $0 \in \mathbb{C}^d$, while still maintaining all other properties of T. We can also apply a Möbius transformation to $f : \mathbb{D}^d \to \mathbb{D}$ and assume $f(0) = 0$.

Let $0, z, w \in \mathbb{D}^d$ be the joint eigenvalues of T with corresponding eigenvectors $e, u, v \in \mathbb{C}^3$. Then, $f(T)$ is the 3×3 matrix with eigenvalues $0, \sigma = f(z), \tau = f(w) \in \mathbb{D}$ and eigenvectors e, u, v. It now becomes of interest to understand all holomorphic functions $g : \mathbb{D}^d \to \mathbb{D}$ which solve the following interpolation problem:

$$\begin{aligned} 0 &\mapsto 0 \\ z &\mapsto \sigma \\ w &\mapsto \tau \end{aligned} \tag{2.1}$$

Theorem 1.2 will follow from the next result which is explained in the next section.

Theorem 2.1 (Kosiński). *If the interpolation problem* (2.1) *can be solved with* $g : \mathbb{D}^d \to \mathbb{D}$ *holomorphic, then there exist holomorphic* $F_1, F_2 : \mathbb{D}^2 \to \mathbb{D}$ *such that* (2.1) *can be solved with a function of the form*

$$F(z) = F_1(F_2(z_1, z_2), z_3) \tag{2.2}$$

after possibly permuting the variables.

Indeed, by Andô's inequality $S = F_2(T_1, T_2)$ is a contraction commuting with T_3 and therefore $F(T) = F_1(S, T_3)$ is a contraction equal to $f(T)$ as above. This proves Theorem 1.2 given Theorem 2.1.

3. The three point Pick problem on the polydisc

Theorem 2.1 follows from work in [14] after making some further reductions to put us in the most interesting situation (that of *extremal* and *non-degenerate* interpolation problems).

First, it is worth pointing out that the two point Pick problem on the polydisc is simple to analyze using one dimensional slices and the Schwarz lemma. It is possible to solve

$$0 \in \mathbb{D}^d \mapsto 0 \in \mathbb{D}$$
$$z \in \mathbb{D}^d \mapsto \sigma \in \mathbb{D}$$

with an analytic $f : \mathbb{D}^d \to \overline{\mathbb{D}}$ if and only if $|\sigma| \leq \max_j |z_j|$. However, even simple problems, such as $(0,0) \mapsto 0$ and $(\frac{1}{2}, \frac{1}{2}) \mapsto \frac{1}{2}$, will have many interpolants f due to the geometry of the polydisc. See Section 11.6 of the book [1].

For these reasons, it is useful to perturb the nodes $0, z = (z_1, \dots, z_d), w = (w_1, \dots, w_d)$ into a more generic position. Let $\rho(a, b) = \left|\frac{a-b}{1-\bar{a}b}\right|$ be the pseudo-hyperbolic distance on the unit disk. Perturb z, w so that all of the quantities below are distinct, yet T is still strictly contractive:

$$|z_1|, \dots, |z_d|, |w_1|, \dots, |w_d|, \rho(z_1, w_1), \dots, \rho(z_d, w_d).$$

Functions as in (2.2) form a normal family so we can approximate the less generic interpolation problems by the generic ones.

The interpolation problem 2.1 is said to be *extremal* if it cannot be solved with a holomorphic function g satisfying $\sup_{\mathbb{D}^d} |g| < 1$. There is no harm in multiplying σ, τ by $r > 1$ if necessary to force the problem to be extremal.

The interpolation problem 2.1 is said to be *non-degenerate* if no two-point subproblem is extremal. If a two point subproblem is extremal (the degenerate case) then one of the following holds

$$|\sigma| = \max_{j=1,\dots,d} |z_j| \text{ or } |\tau| = \max_{j=1,\dots,d} |w_j| \text{ or } \rho(\sigma, \tau) = \max_{j=1,\dots,d} \rho(z_j, w_j).$$

By our genericity assumption, whichever maximum occurs above will occur at a unique j and this forces the solution function to be unique and to depend on one variable. We provide some details in Lemma 3.1 at the end of this section.

Thus, in the degenerate case we can certainly solve with a function of the form (2.2) and we may now assume our interpolation problem is non-degenerate.

To finish, we may quote appropriate results from [14]. Lemma 3 of [14] states that for $d = 3$, if the interpolation problem (2.1) is extremal, non-degenerate, and strictly 3-dimensional, then it can be solved with a function of the form (2.2). Strictly 3-dimensional means the problem can be solved with a function depending on 3 variables but not with a function only depending on 2 variables. We may certainly assume that we are in the strictly 3-dimensional case since otherwise there is nothing to prove.

For $d > 3$, Lemma 5 of [14] states that if (2.1) is extremal and non-degenerate, then after permuting variables if necessary the problem can be solved with a function of the form (2.2). This completes our explanation of Kosiński's Theorem 2.1.

We used the following fact above.

Lemma 3.1. *Suppose $f : \mathbb{D}^d \to \mathbb{D}$ is holomorphic, $f(0) = 0$ and there exists $w \in \mathbb{D}^d$ such that $f(w) = w_1$ and $|w_1| > |w_j|$ for all $j \neq 1$. Then, $f(z) \equiv z_1$ for all $z \in \mathbb{D}^d$.*

Proof. For $\zeta \in \mathbb{D}$, let $h(\zeta) = f(\frac{\zeta}{w_1} w)$. Then, $h(0) = 0, h(w_1) = w_1$ and therefore by the classical Schwarz lemma, $h(\zeta) \equiv \zeta$. This implies

$$\sum_{j=1}^{d} \frac{\partial f}{\partial z_j}(0) \frac{w_j}{w_1} = 1.$$

By [17, pg. 179], $\sum_{j=1}^{d} \left| \frac{\partial f}{\partial z_j}(0) \right| \leq 1$. Since $|w_1| > |w_j|$ for $j \neq 1$, this can only happen if $\frac{\partial f}{\partial z_1}(0) = 1$. This implies $f(\zeta, 0, \ldots, 0) \equiv \zeta$. By Lemma 3.2 of [13], this implies $f(z) \equiv z_1$. □

4. The Schur-Agler class

The Schur-Agler class on the polydisc $\mathbb{D}^d$ consists of holomorphic functions $f : \mathbb{D}^d \to \overline{\mathbb{D}}$ such that

$$\|f(T)\| \leq 1$$

for all d-tuples T of commuting strictly contractive operators. Functions of the form (2.2) are certainly in the Schur-Agler class; indeed the argument after the statement of Theorem 2.1 proves this.

Thus, three-point Pick interpolation problems on the polydisc can be solved with functions in the Schur-Agler class. However, the Schur-Agler class has a well-known interpolation theorem due to Agler; see

[1, Thm 11.49, Thm 11.90]. Combining these observations we get the following.

Theorem 4.1. *Given $z_1, z_2, z_3 \in \mathbb{D}^d$ and $t_1, t_2, t_3 \in \mathbb{D}$, there exists a holomorphic function $f : \mathbb{D}^d \to \mathbb{D}$ satisfying $f(z_j) = t_j$ for $j = 1, 2, 3$ if and only if there exist positive semi-definite 3×3 matrices $\Gamma^1, \Gamma^2, \dots, \Gamma^d$ such that*

$$1 - t_j\overline{t_k} = \sum_{n=1}^{d}(1 - z_j^n\overline{z_k^n})\Gamma^n_{j,k}$$

for $j, k = 1, 2, 3$. Here superscripts are used to denote components of $z_j \in \mathbb{D}^d$.

This result can be used to prove that every solvable 3-point Pick problem can be solved with a rational inner function in the Schur-Agler class. The paper [14] already proves this in the case of non-degenerate extremal problems but then the above general machinery can be used to show that every solvable problem, whether extremal or not, has a rational inner solution.

References

[1] Jim Agler and John E. McCarthy, *Pick interpolation and Hilbert function spaces*, Graduate Studies in Mathematics, vol. 44, American Mathematical Society, Providence, RI, 2002. MR1882259 (2003b:47001)

[2] T. Andô, *On a pair of commutative contractions*, Acta Sci. Math. (Szeged) **24** (1963), 88–90. MR0155193 (27 #5132)

[3] Man-Duen Choi and Kenneth R. Davidson, *A 3×3 dilation counterexample*, Bull. Lond. Math. Soc. **45** (2013), no. 3, 511–519, DOI 10.1112/blms/bds109. MR3065020

[4] M. J. Crabb and A. M. Davie, *von Neumann's inequality for Hilbert space operators*, Bull. London Math. Soc. **7** (1975), 49–50. MR0365179 (51 #1432)

[5] Michael A. Dritschel and Scott McCullough, *The failure of rational dilation on a triply connected domain*, J. Amer. Math. Soc. **18** (2005), no. 4, 873–918, DOI 10.1090/S0894-0347-05-00491-1. MR2163865 (2008i:47024)

[6] S. W. Drury, *Remarks on von Neumann's inequality*, Banach spaces, harmonic analysis, and probability theory (Storrs, Conn., 1980/1981), Lecture Notes in Math., vol. 995, Springer, Berlin, 1983, pp. 14–32, DOI 10.1007/BFb0061886. MR717226 (85b:47006)

[7] Murray Gerstenhaber, *On dominance and varieties of commuting matrices*, Ann. of Math. (2) **73** (1961), 324–348. MR0132079 (24 #A1926)

[8] Rajeev Gupta, *The Carathéodory-Fejér interpolation problems and the von-Neumann inequality*, Indian Institute of Science, Bangalore, 2015.

[9] Robert M. Guralnick, *A note on commuting pairs of matrices*, Linear and Multilinear Algebra **31** (1992), no. 1-4, 71–75, DOI 10.1080/03081089208818123. MR1199042 (94c:15021)

[10] John A. Holbrook, *Schur norms and the multivariate von Neumann inequality*, Recent advances in operator theory and related topics (Szeged, 1999), Oper. Theory Adv. Appl., vol. 127, Birkhäuser, Basel, 2001, pp. 375–386. MR1902811 (2003e:47016)

[11] ______, *Inequalities of von Neumann type for small matrices*, Function spaces (Edwardsville, IL, 1990), Lecture Notes in Pure and Appl. Math., vol. 136, Dekker, New York, 1992, pp. 189–193. MR1152347 (93d:47015)

[12] John Holbrook and Matjaž Omladič, *Approximating commuting operators*, Linear Algebra Appl. **327** (2001), no. 1-3, 131–149, DOI 10.1016/S0024-3795(00)00286-X. MR1823346 (2002a:15018)

[13] Greg Knese, *A refined Agler decomposition and geometric applications*, Indiana Univ. Math. J. **60** (2011), no. 6, 1831–1841, DOI 10.1512/iumj.2011.60.4512. MR3008252

[14] Łukasz Kosiński, *Three-point Nevanlinna Pick problem in the polydisc*, to appear in Proc. Lond. Math. Soc.

[15] B. A. Lotto, *von Neumann's inequality for commuting, diagonalizable contractions. I*, Proc. Amer. Math. Soc. **120** (1994), no. 3, 889–895, DOI 10.2307/2160483. MR1169881 (94e:47012)

[16] Johann von Neumann, *Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes*, Math. Nachr. **4** (1951), 258–281 (German). MR0043386 (13,254a)

[17] Walter Rudin, *Function theory in polydiscs*, W. A. Benjamin, Inc., New York-Amsterdam, 1969. MR0255841 (41 #501)

[18] N. Th. Varopoulos, *On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory*, J. Functional Analysis **16** (1974), 83–100. MR0355642 (50 #8116)

Washington University in St. Louis, Department of Mathematics, St. Louis, Missouri 63130

E-mail address: geknese@math.wustl.edu