Linear Algebra and its Applications 703 (2024) 11–19

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Normal approximations of commuting square-summable matrix families

Alexandru Chirvasitu

Department of Mathematics, University at Buffalo, Buffalo, NY 14260-2900, USA

ARTICLE INFO

Article history:
Received 17 August 2024
Accepted 23 August 2024
Available online 30 August 2024
Submitted by P. Semrl

MSC:
15A24
15A27
47B15
47A12
15A18
15A42
47B07

Keywords:
Spectrum
Generalized eigenspace
Numerical radius
Normal operator
Compact operator
Quasi-nilpotent
Spectral radius
Hilbert-Schmidt norm
Hyers-Ulam stability
Upper-triangular
Superdiagonal
Cholesky factorization

ABSTRACT

For any square-summable commuting family $(A_i)_{i\in I}$ of complex $n \times n$ matrices there is a normal commuting family $(B_i)_i$ no farther from it, in squared normalized ℓ^2 distance, than the diameter of the numerical range of $\sum_i A_i^* A_i$. Specializing in one direction (limiting case of the inequality for finite I) this recovers a result of M. Fraas: if $\sum_{i=1}^{\ell} A_i^* A_i$ is a multiple of the identity for commuting $A_i \in M_n(\mathbb{C})$ then the A_i are normal; specializing in another (singleton I) retrieves the well-known fact that close-to-isometric matrices are close to isometries.

E-mail address: achirvas@buffalo.edu.

https://doi.org/10.1016/j.laa.2024.08.017

0. Introduction

The motivating result for this note is [5, Theorem 1]: commuting $n \times n$ matrices A_i, $1 \leqslant i \leqslant \ell$ with $\sum_{i=1}^{\ell} |A_i|^2 = 1$ are automatically normal, where $|A_i|^2 = A_i^* A_i$. The ingenious proof in [5] relies on the decomposition theory (e.g. [3, §§4 and 5]) of *completely positive* [1, Definition II.6.9.1] maps such as

$$M_n(\mathbb{C}) =: M_n \ni X \longmapsto \sum_{i=1}^{\ell} A_i X A_i^* \in M_n.$$

It seemed sensible, then, to seek for a more directly linear-algebraic proof and perhaps generalize the result in various ways in the process. For a positive integer n and $A = (a_{i,j})_{i,j=1}^n \in M_n$ write

$$|||A|||_2 := \frac{1}{\sqrt{n}} \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

for the normalized Hilbert-Schmidt norm on M_n and similarly for tuples $(A_i)_{i\in I}$ of matrices:

$$\left|\left|\left|(A_i)_{i\in I}\right|\right|\right|_2 := \left(\sum_{i\in I} |||A_i|||_2^2 \right)^{1/2}.$$

In the sequel, all infinite operator sums signify convergence in the *strong topology* of [1, §I.3.1], but see Remark 1.5; for operators on finite-dimensional vector spaces, i.e. plain matrices, this is simply the norm topology on M_n.

Recall that the *numerical range* [8, Chapter 22] of an operator A on a Hilbert space is

$$\mathbb{C} \supset W(A) := \{\langle \xi \mid A\xi \rangle \ \mid \|\xi\| = 1\}. \tag{0.1}$$

With all of this in place, one possible generalization [5, Theorem 1], to be proven below, reads as follows:

Theorem A. *If $A_i \in M_n$, $i \in I$ commute, there are normal commuting $B_i \in M_n$ with*

$$\left|\left|\left|(A_i - B_i)_{i\in I}\right|\right|\right|_2^2 \leqslant \operatorname{diam} W \left(\sum_{i\in I} A_i^* A_i \right) \tag{0.2}$$

(said diameter counting as infinite if $\sum_{i\in I} A_i^* A_i$ fails to converge).

The right-hand side of (0.2) should be regarded as a measure of how far $\sum_{i\in I} A_i^* A_i$ is from being a scalar multiple of the identity (henceforth 'being scalar', for short).

The following variation of the initial motivating result retains the normality context (as opposed to the *near*-normality of Theorem A) but allows for compact operators on infinite-dimensional Hilbert spaces. Recall ([16, pre Theorem 1.5.2], [18, §V.3, post Theorem 3.5]) that an operator on a Banach space is *quasi-nilpotent* if its spectrum is $\{0\}$ (equivalently: its *spectral radius* [16, §1.5] vanishes).

Theorem B. *Let $A_i \in \mathcal{K}(H)$, $i \in I$ be commuting compact operators on a Hilbert space with*

$$\sum_{i\in I} A_i^* A_i = 1 \quad (\textit{strong convergence})\,. \tag{0.3}$$

There is an orthogonal decomposition $H = H_{qn} \oplus H_n$, invariant under all A_i, such that the restrictions $A_i|_{H_n}$ are normal and $A_i|_{H_{qn}}$ are quasi-nilpotent.

Acknowledgments

This work is partially supported by NSF grant DMS-2001128.

I am grateful for stimulating suggestions and comments from M. Fraas, B. Passer and L. Paunescu. An anonymous referee's comments have contributed much to the improvement of an initial draft.

1. Commuting operators square-summable to scalars

We denote the *generalized λ-eigenspace* [9, p.6-1] of an operator A on a Hilbert space by

$$K_\infty(\lambda; A) := \bigcup_{n\in\mathbb{Z}_{>0}} K_n(\lambda; A), \quad K_n(\lambda; A) := \ker(\lambda - A)^n, \tag{1.1}$$

with the subscript "1" on the symbol $K_1(\lambda; A)$ for the plain eigenspace occasionally omitted.

Recall [5, Theorem 1], stating that commuting matrices $A_{i=1}^t \in M_n$ such that

$$\sum_{i=1}^{t} A_i^* A_i = 1 \tag{1.2}$$

are automatically normal. That result turns out to be robust under deformation in the appropriate sense: roughly speaking, commuting families $(A_i)_{i\in I}$ of matrices that *almost* satisfy (1.2) are close to commuting normal families. To make sense of this, recall the *L^p-norm* $\|\cdot\|_p$ ([1, §I.8.7.3], [4, Definition XI.9.1]) defined on the ideal

$$\{\text{compact operators}\} =: \mathcal{K}(H) \trianglelefteq \mathcal{L}(H) := \{\text{bounded operators on a Hilbert space } H\}$$

of compact operators (if allowed to take infinite values):

$$\|T\|_p = \begin{cases} \left(\sum_{n\in\mathbb{Z}_{\geqslant 0}} \mu_n(T)^p\right)^{1/p} & \text{for } 1 \leqslant p < \infty \\ \mu_0(T) = \text{usual operator norm } \|T\| & \text{for } p = \infty \end{cases}$$

where

$$(\mu_0(T) \geqslant \mu_1(T) \geqslant \cdots) := \text{eigenvalues of } |T| := (T^*T)^{1/2}$$

rearranged non-increasingly (the *characteristic numbers* [4, §XI.9] or *s-numbers* [7, §II.2] of T). $\|\cdot\|_2$ is the familiar *Hilbert-Schmidt norm* [4, Definition XI.6.1]. For an operator $A \in \mathcal{L}(H)$ on a Hilbert space H, write

$$sw(A) := \operatorname{diam} W(A) = \sup_{z,z'\in W(A)} |z - z'| \quad (\textit{numerical spread} \text{ of } A),$$

where $W(A)$ is the numerical range (0.1) and $\langle - \mid - \rangle$ denotes the inner product in the ambient Hilbert space, linear in the second variable. Note that $sw(A)$ vanishes precisely for scalar operators, so in general it is a measure of the discrepancy from being scalar.

Remark 1.1. The numerical spread $sw(\cdot)$ is what most naturally fits the statement and proof of Theorem A, but note that for *normal* operators it is nothing but the diameter of the (convex hull of the) spectrum [8, Problem 216].

The term *spread* was in fact introduced for that quantity (diameter of the spectrum) in [15, §1] in the context of matrices. As for links between the two notions of spread (numerical and plain, again for matrices), see e.g. [2, §2]. ♦

This gives the necessary background for Theorem A above. Before going into the proof, note the following immediate consequence; it in turn recovers [5, Theorem 1] by restricting to *finite* families.

Corollary 1.2. *Commuting matrices $\{A_i\}_{i\in I} \subset M_n$ with $\sum_{i\in I} A_i^* A_i$ scalar are all normal and hence generate a commutative C^*-algebra.*

Proof. The first statement is an immediate consequence of Theorem A (since the right-hand side of (1.3) is now assumed to vanish). The second claim then follows from the *Putnam-Fuglede theorem* [8, Problem 192]: commutation with a normal operator entails commutation with its adjoint. □

Remark 1.3. Theorem A is an instance of *Hyers-Ulam(-Rassias) stability*: almost-linear operators between Banach spaces are close to linear operators [12, Theorems 1.1 and

1.2], surjective almost-isometries on Hilbert spaces are close to surjective isometries [12, Theorem 13.4], almost-homogeneous functions between Banach spaces are close to homogeneous functions [12, Theorem 5.11], etc. ♦

Proof of Theorem A. Being commuting, the A_i are simultaneously upper-triangular [10, Theorem 2.3.3] with respect to some orthonormal basis $(e_j)_{j=1}^n$. The sought-after B_i will be the respective diagonals of the matrices A_i:

$$B_i := \operatorname{diag}\left(\lambda_{i,j} := \langle e_j \mid A_i e_j\rangle,\ 1 \leqslant j \leqslant n\right), \quad \forall i.$$

We will verify (0.2) in the rescaled form

$$\sum_{i\in I} \|A_i - B_i\|_2^2 \leqslant n \cdot sw\left(\sum_{i\in I} A_i^* A_i\right). \tag{1.3}$$

To that end, set $T := \sum_{i\in I} A_i^* A_i$ and note first that

$$\sum_{i\in I} \|A_i e_j\|^2 = \sum_{i\in I} \langle e_j \mid A_i^* A_i e_j\rangle = \langle e_j \mid T e_j\rangle, \quad \forall 1 \leqslant j \leqslant n. \tag{1.4}$$

We claim next that for every $1 \leqslant j \leqslant n$ we have

$$\bigcap_{i\in I} K(\lambda_{i,j}; A_i) \neq \{0\}. \tag{1.5}$$

Momentarily taking this for granted, for each fixed j we can simultaneously upper-triangularize the A_i with respect to a new orthonormal basis with a vector e'_j in (1.5) listed first, so that

$$\sum_{i\in I} \|B_i e_j\|^2 = \sum_{i\in I} \|A_i e'_j\|^2 = \langle e'_j \mid T e'_j\rangle, \quad \forall 1 \leqslant j \leqslant n. \tag{1.6}$$

(1.4) and (1.6) are at most $sw(T)$ apart by the latter's definition, hence the conclusion upon summing over $1 \leqslant j \leqslant n$: because the matrices B_i are the respective diagonals of the A_is,

$$\sum_{i\in I} \|A_i - B_i\|_2^2 = \sum_{\substack{i\in I\\ 1\leqslant j\leqslant n}} \left(\|A_i e_j\|^2 - \|B_i e_j\|^2\right) \leq n \cdot sw(T).$$

It remains to settle (1.5). Since A_i commute and thus preserve each other's eigenspaces, that assertion is equivalent to the non-trivial intersection of the *generalized* eigenspaces $K_\infty(\lambda_{i,j}; A_i)$. Were that intersection trivial, the 1-dimensional module $A_i \mapsto \lambda_{i,j}$ of the commutative algebra $\mathcal{A}$ generated by the A_i would not appear as a subquotient in a *Jordan-Hölder filtration* [17, Proposition III.3.7] of either of the two $\mathcal{A}$-modules

$$K_\infty(\lambda_{i_0,j}; A_{i_0}) \quad \text{and} \quad V/K_\infty(\lambda_{i_0,j}; A_{i_0}), \quad V := \text{ambient space } \mathbb{C}^n$$

for a fixed index i_0, so would not appear in such a filtration at all. This is at odds with the original triangularization with respect to (e_j), hence the contradiction. □

Remarks 1.4.

(1) The commutativity of the family $(A_i)_{i\in I}$ of Corollary 1.2 cannot be relaxed to simultaneous unitary upper-triangularizability (as the proof, appealing crucially to that commutativity, suggests): *every* positive operator on $\mathbb{C}^n$ is expressible as T^*T for upper-triangular T (the celebrated *Cholesky factorization* [10, Corollary 7.2.9]), so it is enough to decompose $1 \in M_n$ as a sum of non-diagonal positive operators, say

$$1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

and express each summand as $A_i^*A_i$ for upper-triangular A_i.

(2) The dependence on n in (1.3) vanishes upon substituting the *normalized* Hilbert-Schmidt norm $\frac{1}{\sqrt{n}}\|\cdot\|_2$ on M_n for $\|\cdot\|_2$, as is customary in the literature on almost-commutative matrices ([6, §1], [11, §2], etc.). ♦

It is noted in [5, §4] that the *unilateral shift* [8, Problem 82]

$$e_n \xmapsto{S} e_{n+1}, \quad n \in \mathbb{Z}_{\geqslant 0}$$

on a Hilbert space with orthonormal basis $(e_n)_{n\in\mathbb{Z}_{\geqslant 0}}$ is a non-self-adjoint singleton giving a counterexample to Corollary 1.2 in infinite-dimensional spaces. Corollary 1.2 does, however, suggest a more hopeful infinite-dimensional variant: the $A_i \in \mathcal{L}(H)$ might be commuting *compact* [1, Definition I.8.1.1] operators on a Hilbert space H, with the convergence of Corollary 1.2 valid in the *strong* [1, Definition I.3.1.1] topology on $\mathcal{L}(H)$. In that context, Theorem B above requires little more than has already been noted.

Remark 1.5. Which of the six standard weaker-than-norm topologies [1, §I.3.1] on $\mathcal{L}(H)$ (weak, σ-weak, strong, σ-strong, strong* and σ-strong*) is explicitly mentioned in (0.3) is a matter of taste: per [8, Problem 120] (phrased in terms of plain sequences but applicable in the present generality), for bounded non-decreasing *nets* [19, Definition 11.2] of positive operators those topologies induce the same notion of convergence. ♦

There is a theory of upper-triangularization for single compact operators: the term for what we would here call 'upper-triangular' is *superdiagonal* in [16, §4.3]; other sources [4, §XI.10] work with *sub*diagonal operators instead. That material extends straightforwardly to commuting families of compact operators: the central result driving the theory, namely [16, Theorem 4.2.1] the fact that compact operators have non-trivial invariant subspaces, is now well-known ([13, Theorem], [14]) for the *commutant* of a non-zero compact operator. We take all of this for granted, along with the requisite background on compact-operator spectral theory.

Recall [16, Theorem 1.8.1], in particular, that for compact $A \in \mathcal{L}(H)$ the generalized eigenspaces (1.1) attached to $\lambda \in \mathbb{C}^\times := \mathbb{C} \setminus \{0\}$ are finite-dimensional. In particular, the same goes for the (plain) eigenspaces $K(\lambda; A) := K_1(\lambda; A)$.

Proof of Theorem B. We isolate a single operator $A := A_{i_0}$ and fix a non-zero $\lambda \in \sigma(A)$. There is [16, Theorem 1.8.1] a direct-sum decomposition

$$H = K_\infty(\lambda; A) \oplus R_\infty(\lambda; A)$$

('R' for 'range') where, by analogy to (1.1),

$$R_\infty(\lambda; A) := \bigcap_{n \in \mathbb{Z}_{>0}} R_n(\lambda; A), \quad R_n(\lambda; A) := \operatorname{im}(\lambda - A)^n = (\lambda - A)^n H.$$

In an appropriate orthonormal basis for H, compatible with the orthogonal decomposition $H = R_\infty(\lambda; A) \oplus R_\infty(\lambda; A)^\perp$, we have

$$A = \begin{pmatrix} A' & \bullet \\ 0 & T \end{pmatrix}$$

with T finite (of width $\dim K_\infty(\lambda; A) < \infty$), upper triangular, with diagonal entries λ. The argument employed in the proof of Corollary 1.2 will then show that the $\bullet$ block vanishes. This is sufficient to ensure that

- for non-zero $\lambda \in \sigma(A)$ the generalized eigenspaces are in fact eigenspaces:

$$K_\infty(\lambda; A) = K(\lambda; A), \quad \forall \lambda \in \sigma(A)^\times := \sigma(A) \setminus \{0\};$$

- and those eigenspaces are mutually orthogonal for distinct λ:

$$K(\lambda; A) \perp K(\lambda'; A), \quad \forall \lambda \neq \lambda' \in \sigma(A)^\times;$$

- and finally, said eigenspaces are all orthogonal to the largest A-invariant subspace where A is quasi-nilpotent:

$$K(\lambda; A) \perp R_\infty(\sigma(A)^\times; A) := \bigcap_{\mu \in \sigma(A)^\times} R_\infty(\mu; A). \tag{1.7}$$

In short: A is an orthogonal direct sum of a quasi-nilpotent compact operator and a normal compact operator, operating respectively on the space $R_\infty(\sigma(A)^\times; A)$ of (1.7) and its orthogonal complement. Finally, setting

$$H_{qn} := \bigcap_i R_\infty(\sigma(A_i)^\times; A_i)$$

will do. □

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Alexandru Chirvasitu reports financial support was provided by NSF grant DMS-2001128. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] B. Blackadar, Operator algebras, in: Theory of C^*-Algebras and Von Neumann Algebras, Operator Algebras and Non-commutative Geometry, III, in: Encyclopaedia of Mathematical Sciences, vol. 122, Springer-Verlag, Berlin, 2006.
[2] Mao-Ting Chien, Chi-Kwong Li, Hiroshi Nakazato, The diameter and width of higher rank numerical ranges, Linear Multilinear Algebra 69 (5) (2021) 871–887.
[3] Giuseppe Ilario Cirillo, Francesco Ticozzi, Decompositions of Hilbert spaces, stability analysis and convergence probabilities for discrete-time quantum dynamical semigroups, J. Phys. A, Math. Theor. 48 (8) (2015) 28, Id/No 085302.
[4] Nelson Dunford, Jacob T. Schwartz, Linear Operators. Part II: Spectral Theory. Self-Adjoint Operators in Hilbert Space, Pure and Applied Mathematics, vol. 7, Interscience Publishers, a Division of John Wiley and Sons, New York and London, 1963, pp. 859–1923. With the assistance of William G. Bade and Robert G. Bartle.
[5] Martin Fraas, Commuting Kraus operators are normal, http://arxiv.org/abs/2308.05450v1, 2023.
[6] Lev Glebsky, Almost commuting matrices with respect to normalized Hilbert-Schmidt norm, http://arxiv.org/abs/1002.3082v1, 2010.
[7] I.C. Gohberg, M.G. Kreĭn, Introduction to the Theory of Linear Nonselfadjoint Operators, Translations of Mathematical Monographs, vol. 18, American Mathematical Society, Providence, R.I., 1969. Translated from the Russian by A. Feinstein.
[8] Paul R. Halmos, A Hilbert Space Problem Book, 2nd ed., rev. and enl, Grad. Texts Math., vol. 19, Springer, Cham, 1982.
[9] Leslie Hogben (Ed.), 2nd enlarged ed. edition, Handbook of Linear Algebra, Discrete Math. Appl. (Boca Raton), Chapman & Hall/CRC, Boca Raton, FL, 2014.
[10] Roger A. Horn, Charles R. Johnson, Matrix Analysis, 2nd ed. edition, Cambridge University Press, Cambridge, 2013.
[11] Adrian Ioana, Stability for product groups and property (τ), J. Funct. Anal. 279 (9) (2020) 108729.
[12] Soon-Mo Jung, Hyers-Ulam-Rassias Stability of Functional Equations in Nonlinear Analysis, Springer Optim. Appl., vol. 48, Springer, Berlin, 2011.
[13] V.I. Lomonosov, Invariant subspaces for the family of operators which commute with a completely continuous operator, Funct. Anal. Appl. 7 (1974) 213–214.
[14] A.J. Michaels, Hilden's simple proof of Lomonosov's invariant subspace theorem, Adv. Math. 25 (1977) 56–58.
[15] L. Mirsky, The spread of a matrix, Mathematika 3 (1956) 127–130.
[16] John R. Ringrose, Compact Non-self-Adjoint Operators, Van Nostrand Reinhold Mathematical Studies, vol. 35, Van Nostrand Reinhold Company. VI, London, 1971, 238 p.

[17] Bo Stenström, Rings of Quotients. An Introduction to Methods of Ring Theory, Grundlehren Math. Wiss., vol. 217, Springer, Cham, 1975.

[18] Angus E. Taylor, David C. Lay, Introduction to Functional Analysis, 2nd ed., Robert E. Krieger Publishing Company. XI, Malabar, Florida, 1986, 467 p. (Reprint of the orig. 1980, publ. by John Wiley & Sons, Inc., New York etc.).

[19] Stephen Willard, General Topology, Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1970 original [Addison-Wesley, Reading, MA, MR0264581].