Can. J. Math.Vol. **44** (2), 1992 pp. 280–297

C^*-CONVEXITY AND MATRICIAL RANGES

D. R. FARENICK

ABSTRACT. C^*-convex sets in matrix algebras are convex sets of matrices in which matrix-valued convex coefficients are admitted along with the usual scalar-valued convex coefficients. A Carathéodory-type theorem is developed for C^*-convex hulls of compact sets of matrices, and applications of this theorem are given to the theory of matricial ranges. If T is an element in a unital C^*-algebra $\mathcal{A}$, then for every $n \in \mathbb{N}$, the $n \times n$ matricial range $W^n(T)$ of T is a compact C^*-convex set of $n \times n$ matrices. The basic relation $W^1(T) = \operatorname{conv} \sigma(T)$ is well known to hold if T exhibits the normal-like quality of having the spectral radius of $\beta T + \mu 1$ coincide with the norm $\|\beta T + \mu 1\|$ for every pair of complex numbers β and μ. An extension of this relation to the matrix spaces is given by Theorem 2.6: $W^n(T)$ is the C^*-convex hull of the $n \times n$ matricial spectrum $\sigma^n(T)$ of T if, for every $B, M \in \mathcal{M}_n$, the norm of $T \otimes B + 1 \otimes M$ in $\mathcal{A} \otimes \mathcal{M}_n$ is the maximum value in $\{ \|\Lambda \otimes B + 1 \otimes M\| : \Lambda \in \sigma^n(T)\}$. The spatial matricial range of a Hilbert space operator is the analogue of the classical numerical range, although it can fail to be convex if $n > 1$. It is shown in §3 that if T has a normal dilation N with $\sigma(N) \subset \sigma(T)$, then the closure of the spatial matricial range of T is convex if and only if it is C^*-convex.

Introduction. In this paper we will be concerned with the structure of compact sets $\mathcal{K}$ in the C^*-algebra $\mathcal{M}_n$ of $n \times n$ complex matrices which possess the following convexity property:

$$\text{whenever } \Lambda_1, \ldots, \Lambda_p \in \mathcal{K} \text{ and } T_1, \ldots, T_p \in \mathcal{M}_n \text{ with } \sum_{i=1}^{p} T_i^* T_i = 1,$$
$$\text{then } \sum_{i=1}^{p} T_i^* \Lambda_i T_i \in \mathcal{K}.$$

Sets of this type have been called *matrix convex* [2], *matricially convex* [5], and because this form of convexity can be defined in general $*$-algebras other than $\mathcal{M}_n$, these sets have also been called C^**-convex* in [11], where the formal study of this generalized form of convexity was initiated by R. I. Loebl and V. I. Paulsen. In the conditions defining C^*-convexity, p can be any positive integer, the Λ_i's need not be distinct, and the C^*-*convex coefficients* T_i need not commute among themselves, nor with the Λ_i's. C^*-convex sets are, of course, plainly convex, and the class of C^*-convex sets in the 1-dimensional algebra $\mathbb{C}$ is precisely the class of convex sets. The unit ball in a (unital) C^*-algebra is a simple example of a C^*-convex set [11], however closed balls centred at non-scalar elements can fail to be C^*-convex. The matrix-valued analogue of the numerical range for elements of a unital C^*-algebra provides all examples of compact C^*-convex sets in $\mathcal{M}_n$ [11; Prop. 30].

Received by the editors July 16, 1990.
AMS subject classification: Primary: 47A12; secondary: 15A60, 46L05..

If T is an element in a unital C^*-algebra $\mathcal{A}$, then the $n \times n$ *matricial range* of T is the set

$$W^n(T) = \{\phi(T) : \phi \text{ is a unital completely positive map } C^*(T) \longrightarrow \mathcal{M}_n\},$$

where $C^*(T)$ is the unital C^*-algebra generated by T. The most basic fact concerning $W^n(T)$ is that it is a compact C*-convex unitary invariant of T [2;p.301]. If $\mathcal{A}$ is the algebra $\mathcal{B}(\mathcal{H})$ of bounded linear operators on a complex Hilbert space $\mathcal{H}$, then $W^n(T)$ has a path connected spatial component known as the $n \times n$ *spatial matricial range*:

$$W_s^n(T) = \{V^*TV : V \text{ is a linear isometry } \mathbb{C}^n \longrightarrow \mathcal{H}\}.$$

Although $W_s^1(T) = \{(Tx,x) : x \in \mathcal{H} \text{ and } \|x\| = (x,x)^{\frac{1}{2}} = 1\}$ is the classical numerical range of T, $W_s^n(T)$ need not be convex if $n > 1$. Basic references for the spatial and algebraic numerical ranges respectively are [8] and [20], and the theory of their matricial analogues is developed, for example, in [2], [4], [5], [7], [10], [12], [15], and [19].

The purpose of the first part of this paper is to prove a Carathéodory-type theorem for the C*-convex hull of a compact set in $\mathcal{M}_n$, and to give applications of this theorem to matricial range theory and to matrix theory. These results are motivated by the following well established fact (implicit, for example, in the proofs of [20;Thm. 8] and [13;2.10]): if T is an element in a unital C^*-algebra $\mathcal{A}$ such that the norm of every linear polynomial $\beta T+\mu 1$ in T coincides with the spectral radius of $\beta T+\mu 1$, then $W^1(T)$ is the convex hull of the spectrum $\sigma(T)$ of T. In § 2, the proper matricial extension of this norm—spectral radius relation is formulated, so that $W^n(T)$ is given by the C*-convex hull of the matricial spectrum $\sigma^n(T)$ of T. The formulation involves the comparison of the norm of every linear $n \times n$ matrix-polynomial in T to the maximum of the norms obtained in evaluating the matrix-polynomial over the $n \times n$ matricial spectrum of T, thereby sharpening, extending, and unifying some early results in this direction on matricial ranges. A key geometric idea is to introduce the notion of a matrix-valued disc which functions as a C*-convex generalization of a compact planar disc. As an outgrowth of these developments, the representation of contractive matrices as convex combinations of unitaries will be seen to extend to a statement that matrices with compact spectral sets $X \subset \mathbb{C}$ can be represented as C*-convex combinations of normal matrices N_i with $\sigma(N_i) \subset \partial X$.

The second part of this paper examines convexity and C*-convexity phenomena exhibited by the spatial matricial ranges of certain operators. In light of the fact that spatial matricial ranges need not be convex, it is somewhat surprising that they possess any interesting convexity properties at all. Special attention will be paid to the relation between convexity and C*-convexity, and between $W_s^n(T)$ and the spatial matricial spectrum $\sigma_s^n(T)$. Two other generalizations of the classical numerical range, the C-numerical range and the k-numerical range, have a role in obtaining these convexity theorems; in particular, a theorem of Y.-T. Poon on the k-numerical range for matrices is found to have a reasonably strong formulation in infinite dimensions.

The concepts that we will use from the theory of positive maps and from matrix-valued spectral theory are reviewed in § 1.

ACKNOWLEDGEMENT. The results presented here are based upon part of my doctoral dissertation at University of Toronto. I wish to thank my thesis advisor, Professor Chandler Davis, for his encouragement, support, and direction of this work. I would also like to thank Professor Chi-Kwong Li for sending me a preliminary version of [10].

1. **Positivity and matricial spectra.** Throughout, all C^*-algebras are assumed to possess a unit 1. When two elements S and T of a C^*-algebra $\mathcal{A}$ are unitarily equivalent (i.e., there is a unitary U with $U^*SU = T$), then this will be denoted by $S \simeq T$. Recall, from the theory of completely positive maps [13], that a continuous linear map $\phi : \mathcal{L} \longrightarrow \mathcal{B}$ of a self-adjoint linear manifold $\mathcal{L}$ in a C^*-algebra $\mathcal{A}$ into a C^*-algebra $\mathcal{B}$ is said to be *n-positive* if $\phi_n = \phi \otimes \mathrm{id}_n : \mathcal{L} \otimes \mathcal{M}_n \longrightarrow \mathcal{B} \otimes \mathcal{M}_n$, where id_n is the identity map on $\mathcal{M}_n$, is a positive map, and that ϕ is called *completely positive* if ϕ is n-positive for every $n \in \mathbb{N}$. For a given $T \in \mathcal{A}$, let $\mathrm{CP}^n(T)$ denote the space of all completely positive maps $\phi : C^*(T) \longrightarrow \mathcal{M}_n$ satisfying $\phi(1) = 1 \in \mathcal{M}_n$; in the BW-topology, the (convex) space $\mathrm{CP}^n(T)$ is compact [1],[13].

The matricial spectrum is developed in [5], [7], and [14], for example. The following discussion briefly reviews the parts of the theory that are pertinent in this paper. Generally, the Hilbert space $\mathcal{H}$ is assumed to be separable.

If $T \in \mathcal{B}(\mathcal{H})$, then the $n \times n$ *spatial left matricial spectrum* of T is the set $\Pi^n_s(T)$ of matrices $\Lambda \in \mathcal{M}_n$ for which there exist isometries $V_k : \mathbb{C}^n \longrightarrow \mathcal{H}$ satisfying $0 = \lim_k \|TV_k - V_k\Lambda\|$. The set $\Pi^1_s(T)$ consists of the approximate eigenvalues of T, and in finite-dimensional spaces, $\Lambda \in \Pi^n_s(T)$ if and only if Λ is unitarily equivalent to an operator obtained by restricting T to one of its n-dimensional invariant subspaces. There is also a notion of reducing matricial spectra: a matrix $\Lambda \in \mathcal{M}_n$ is in the *spatial reducing matricial spectrum* $R^n_s(T)$ if there exist isometries $V_k : \mathbb{C}^n \longrightarrow \mathcal{H}$ with $0 = \lim_k \|TV_k - V_k\Lambda\|$ and $0 = \lim_k \|T^*V_k - V_k\Lambda^*\|$. The theory of the reducing spectrum $R^1_s(T)$, developed by N. Salinas in [17], will enter into our work frequently; the main feature of the reducing spectrum is that its properties are those of the spectrum of a normal operator. The $n \times n$ *reducing matricial spectrum* $R^n(T)$ is the set of all $\rho(T)$, where $\rho \in \mathrm{CP}^n(T)$ is a unital $*$-homomorphism $C^*(T) \longrightarrow \mathcal{M}_n$.

A matrix $\Lambda \in \mathcal{M}_n$ is an element of the *spatial matricial spectrum* $\sigma^n_s(T)$ of $T \in \mathcal{B}(\mathcal{H})$ if there exist isometries $V_k : \mathbb{C}^n \longrightarrow \mathcal{H}$ satisfying $0 = \lim_k \|V_k^* r(T)V_k - r(\Lambda)\|$ for every complex rational function r with poles off $\sigma(T)$. The simplest case occurs when $\mathcal{H}$ is finite-dimensional: a matrix $\Lambda \in \mathcal{M}_n$ is in $\sigma^n_s(T)$ if and only if Λ is unitarily equivalent to an operator obtained by compressing T to one of its n-dimensional semi-invariant subspaces. In all cases, $\sigma^n_s(T)$ is a non-empty compact set with $\sigma^1_s(T) = \sigma(T)$. The *matricial spectrum* $\sigma^n(T)$ of $T \in \mathcal{A}$ is the set of all $n \times n$ matrices $\phi(T)$ obtained from those $\phi \in \mathrm{CP}^n(T)$ which are homomorphisms when restricted to the algebra of rational functions in T.

Finally, recall that $\sigma_e(T)$, the *essential spectrum* of T, is the spectrum of the canonical image of T in the Calkin algebra $\mathcal{B}(\mathcal{H})$ modulo the compact operators, and that if N is normal, then $\sigma_e(N)$ is precisely all of $\sigma(N)$ except for isolated eigenvalues of finite multiplicity. (Here, "isolated" means that the element is isolated in the entire spectrum in the topological sense.) If $T \in \mathcal{B}(\mathcal{H})$, then $\lambda 1_n \in \sigma_s^n(T)$ for every $n \in \mathbb{N}$ whenever $\lambda \in \sigma_e(T)$; for this reason we say that $\lambda \in \sigma(T)$ has infinite multiplicity if $\lambda \in \sigma_e(T)$.

Some examples (see [7]): if T is a compact normal operator, if T is a self-adjoint operator (not necessarily compact), or if T is a unitary operator, then $\sigma_s^n(T) = R_s^n(T)$ for every $n \in \mathbb{N}$; in particular, if $\mathcal{H}$ is a finite-dimensional space, and if $T \in \mathcal{B}(\mathcal{H})$ is normal, then $\sigma_s^n(T) = R_s^n(T)$ for every $1 \leq n \leq \dim \mathcal{H}$. All of the spectral elements in these examples are normal. In general, if $T \in \mathcal{B}(\mathcal{H})$ is normal, then $\lambda_1 1_{k_1} \oplus \cdots \oplus \lambda_p 1_{k_p} \in R_s^n(T)$ if and only if each $\lambda_i \in \sigma(T)$ and $k_j \leq \dim\big(\ker(T - \lambda_j 1)\big)$ whenever λ_j is an isolated eigenvalue of finite multiplicity.

The following lemma will be used in § 3, and it will allow us to employ a self-adjoint operator with finite spectrum, rather than a general self-adjoint operator, in the proof of Theorem 3.1.

LEMMA 1. *If $T \in \mathcal{B}(\mathcal{H})$ is self-adjoint, and if $\epsilon > 0$, then there is a self-adjoint operator A with finite spectrum such that $\|T - A\| < \epsilon$ and $\sigma_s^n(A) \subset \sigma_s^n(T)$ for every $n \in \mathbb{N}$.*

PROOF. Because $\sigma(T)$ is compact, the Borel function $f : x \longmapsto x$ can be approximated in $L^\infty\big(\sigma(T)\big)$ to within $\frac{\epsilon}{2}$ by a step function

$$\phi(x) = \sum_{i=1}^{p} \zeta_i \chi_{\mathcal{V}_i}(x),$$

where $\mathcal{V}_1, \ldots, \mathcal{V}_p$ are finitely many mutually disjoint Borel subsets of $\sigma(T)$ with the property that $\mathcal{V}_i$ is a finite set if and only if $\mathcal{V}_i$ is a singleton set consisting of an isolated point of spectrum, and such that $\sigma(T) = \cup_{i=1}^{p} \mathcal{V}_i$. We will now construct a new step function ψ which is within $\frac{\epsilon}{2}$ of ϕ by replacing the points ζ_i with certain spectral values. Let E denote the spectral measure for T. If the rank of the projection $P_i = E(\mathcal{V}_i)$ is infinite, then there exists a $\lambda_i \in \mathcal{V}_i \cap \sigma_e(T)$; replace ζ_i by λ_i. If the rank of P_i is finite, then P_i is the projection onto a finite-dimensional invariant subspace, and so $\mathcal{V}_i$ is finite; thus, $\mathcal{V}_i$ is a singleton set $\{\zeta_i\}$, and ζ_i is an isolated eigenvalue of finite multiplicity, call it λ_i. Let $A = \psi(T)$; then $A = \sum_{i=1}^{p} \lambda_i P_i$ is a self-adjoint operator with finite spectrum, and

$$\|T - A\| = \|f - \psi\|_\infty \leq \|f - \phi\|_\infty + \|\phi - \psi\|_\infty < \epsilon.$$

The only thing left to verify is that $\sigma_s^n(A) \subset \sigma_s^n(T)$. By the choice $\lambda_1, \ldots, \lambda_p$, this amounts to verifying that the rank of the projections P_i are such that the multiplicities of the eigenvalues λ_i of A do not exceed the multiplicities of the spectral values λ_i of T. By the restrictions placed on the partition $\mathcal{V}_1, \ldots, \mathcal{V}_p$ (and, therefore, on the rank of each projection P_i), we see that indeed $\sigma_s^n(A) \subset \sigma_s^n(T)$, for every $n \in \mathbb{N}$. ∎

2. **Matrix-valued discs and a Carathéodory-type theorem.** A fundamental theorem of convex analysis is Carathéodory's Theorem which states that if $\mathcal{S}$ is a set in a real vector space of dimension n, then every element of the convex hull of $\mathcal{S}$ is expressible as a convex combination of elements of $\mathcal{S}$ involving at most $n+1$ terms. There are numerous consequences of this theorem, among which is the statement that the convex hull of a compact set is compact. The problem of whether the C*-convex hull of a compact set of $n \times n$ matrices is compact has been open for some time; a positive solution to this problem is found in Theorem 2.4, part of which can be viewed as an analogue of the Carathéodory theorem. Theorem 2.4 also contains a geometric characterization of the C*-convex hull of a compact set of matrices, where the geometric objects utilized are compact C*-convex sets known as matrix-valued discs.

To fix the notation and terminology, suppose that $\mathcal{S} \subset \mathcal{M}_n$. The *C*-convex hull* of $\mathcal{S}$ is the smallest C*-convex set in $\mathcal{M}_n$ containing $\mathcal{S}$, and it will be denoted by $\operatorname{mconv}(\mathcal{S})$; the convex hull of a set $\mathcal{S}$ will be denoted by $\operatorname{conv} \mathcal{S}$.

DEFINITION. If $B, M \in \mathcal{M}_n$ with $B \neq 0$, and if $r > 0$, then the *matrix-valued disc* $\mathrm{D}(B, M; r)$ induced by B and M of radius r is the set of all $\Lambda \in \mathcal{M}_n$ which satisfy the norm inequality $\|\Lambda \otimes B + 1 \otimes M\| \leq r$ in the C^*-algebra $\mathcal{M}_n \otimes \mathcal{M}_n$.

Observe that when $n = 1$, a matrix-valued disc is just a disc in the plane in the usual metric. To show that these generalized discs are C*-convex, we follow the procedure used to show that the unit ball in any unital C^*-algebra is C*-convex [13;p.64]: suppose that $\Lambda_1, \ldots, \Lambda_p$ are elements of the matrix-valued disc induced by $B, M \in \mathcal{M}_n$ of radius r, and suppose that $T_1, \ldots, T_p$ are C*-convex coefficients; then

$$\begin{aligned}\Big\|\Big(\sum_{j=1}^{p} T_j^* \Lambda_j T_j\Big) \otimes B + 1 \otimes M\Big\| &= \Big\|\sum_{j=1}^{p} (T_j \otimes 1)^* (\Lambda_j \otimes B + 1 \otimes M)(T_j \otimes 1)\Big\| \\ &= \|G^* X G\| \leq \|G\|^2 \|X\| = \|G^* G\| \, \|X\| \leq r,\end{aligned}$$

where X and G are operators on $\mathbb{C}^{np} \otimes \mathbb{C}^n$ such that X is the direct sum of $\Lambda_j \otimes B + 1 \otimes M$ for $j = 1, \ldots, p$ and the operator matrix G consists of zero operators except in the first column where each $(j, 1)$-entry is given by $T_j \otimes 1$. This proves that matrix-valued discs are C*-convex. Because it is plainly evident that $\mathrm{D}(B, M; r)$ is closed, the compactness of $\mathrm{D}(B, M; r)$ can be verified by showing the boundedness of $\mathrm{D}(B, M; r)$. If $\Lambda \in \mathrm{D}(B, M; r)$, then

$$r \geq \|\Lambda \otimes B + 1 \otimes M\| \geq \max\left\{\|b_{i,j}\Lambda + m_{i,j}1\| : 1 \leq i, j \leq n \text{ and } b_{i,j} \neq 0\right\},$$

and so $\mathrm{D}(B, M; r)$ is contained in the finite intersection of the compact (C*-convex) r-balls $\{\Omega : \|b_{i,j}\Lambda + m_{i,j}1\| \leq r\}$ for which $b_{i,j} \neq 0$. (A mildly disconcerting fact is that this intersection could be empty.) The geometric and affine properties of matrix-valued discs have not yet been studied in detail.

The work in this section is based primarily on two results: one is the description of the extremal structure of $\mathrm{CP}^n(T)$ when T is n-normal, and the other is the following lemma describing one way in which the matricial range can be characterized independently of positivity and algebraic considerations.

LEMMA 2.1. *If $T \in \mathcal{A}$ and $n \in \mathbb{N}$, then*

$$W^n(T) = \bigcap_{B,M \in \mathcal{M}_n} \left\{ \Lambda \in \mathcal{M}_n : \|\Lambda \otimes B + 1 \otimes M\| \leq \|T \otimes B + 1 \otimes M\| \right\}$$

PROOF. If $\Lambda \in W^n(T)$, then $\Lambda = \phi(T)$ for some unital completely positive map $\phi: C^*(T) \longrightarrow \mathcal{M}_n$. The induced map $\phi_n: C^*(T) \otimes \mathcal{M}_n \longrightarrow \mathcal{M}_n \otimes \mathcal{M}_n$ is a contraction, because unital completely positive maps are completely contractive. Hence, the norm inequality

$$\|\Lambda \otimes B + 1 \otimes M\| \leq \|T \otimes B + 1 \otimes M\|$$

holds for every $B, M \in \mathcal{M}_n$.

Conversely, if $\Lambda \in \mathcal{M}_n$ is such that for every $B, M \in \mathcal{M}_n$ the above norm inequality holds, then the map $\phi_n: \mathcal{L} \otimes \mathcal{M}_n \longrightarrow \mathcal{M}_n \otimes \mathcal{M}_n$ is a unital contraction, where $\mathcal{L}$ is the 2-dimensional subspace spanned by T and 1, and where $\phi: \mathcal{L} \longrightarrow \mathcal{M}_n$ is defined by $\phi(\alpha 1 + \beta T) = \alpha 1 + \beta \Lambda$. Therefore, ϕ_n has a positive extension $\tilde{\phi}_n: (\mathcal{L}^* + \mathcal{L}) \otimes \mathcal{M}_n \longrightarrow \mathcal{M}_n \otimes \mathcal{M}_n$, by [13;2.12], where the map $\tilde{\phi}: (\mathcal{L}^* + \mathcal{L}) \longrightarrow \mathcal{M}_n$ is given by

$$\tilde{\phi}(\alpha 1 + \beta T^* + \gamma T) = \alpha 1 + \beta \Lambda^* + \gamma \Lambda.$$

This means that $\tilde{\phi}$ is n-positive. A well known theorem of M.-D. Choi states that n-positive maps into $\mathcal{M}_n$ are completely positive [13;5.9], and so, via Arveson's Extension Theorem ([1],[13;6.5]), $\tilde{\phi}$ has a completely positive extension from $(\mathcal{L}^* + \mathcal{L})$ to all of $C^*(T)$. Hence, $\Lambda = \tilde{\phi}(T) \in W^n(T)$. ∎

COROLLARY 2.2. *If $n \in \mathbb{N}$, and if $S, T \in \mathcal{A}$ are such that $\|S \otimes B + 1 \otimes M\| \leq \|T \otimes B + 1 \otimes M\|$ for every $B, M \in \mathcal{M}_n$, then $W^n(S) \subset W^n(T)$.*

The following is a standard lemma, and it is stated here because it will be employed in more than just the next theorem. It states that every extreme point of $W^n(T)$ comes from an extreme point of $\mathrm{CP}^n(T)$.

LEMMA 2.3. *If $T \in \mathcal{A}$, and if $\Lambda \in W^n(T)$ is an extreme point of $W^n(T)$, then there is an extreme point ϕ of $\mathrm{CP}^n(T)$ such that $\Lambda = \phi(T)$.*

PROOF. Let $\mathcal{C}$ be the convex BW-compact subset of $\mathrm{CP}^n(T)$ of all $\phi \in \mathrm{CP}^n(T)$ satisfying $\Lambda = \phi(T)$. By the Krein-Milman Theorem, there is a $\phi \in \mathcal{C}$ which is an extreme point of $\mathcal{C}$. Therefore, to prove that ϕ is an extreme point of $\mathrm{CP}^n(T)$, it suffices to show that whenever $\phi = t\phi_1 + (1-t)\phi_2$ is a proper convex combination of $\phi_1, \phi_2 \in \mathrm{CP}^n(T)$, we must have $\phi_1, \phi_2 \in \mathcal{C}$. Thus, suppose that $\phi = t\phi_1 + (1-t)\phi_2$ is such a combination. Then $\Lambda = \phi(T) = t\phi_1(T) + (1-t)\phi_2(T)$; this is a representation of Λ as a proper convex combination of elements of $W^n(T)$. But since Λ is an extreme point of $W^n(T)$, it must be that $\Lambda = \phi_1(T) = \phi_2(T)$, and hence, $\phi_1, \phi_2 \in \mathcal{C}$. ∎

The proof of the Carathéodory-type theorem makes use of the notions of n-normality and hypoconvexity. An operator $T \in \mathcal{B}(\mathcal{H})$ is *n-normal* if T is unitarily equivalent to an $n \times n$ operator matrix consisting of commuting normal operators. An example, canonical up to compact perturbations of arbitrarily small norm [14], is the direct sum of a bounded set of $n \times n$ matrices.

A non-empty compact set $X \subset \mathcal{M}_n$ is said to be *hypoconvex* if X is invariant under unitary similarity transformations and if X is closed under the formation of sums of the type $\sum_{j=1}^{m} P_j X_j$, where $X_1, \ldots, X_m$ are arbitrary elements of X, and where $P_1, \ldots, P_m$ are projections with mutually orthogonal ranges and which satisfy $P_j X_j = X_j P_j$ for each $1 \le j \le m$ and $\sum_{j=1}^{m} P_j = 1$. (Consequently, $m \le n$.) The importance of hypoconvexity comes from the fact that if a compact subset $X \subset \mathcal{M}_n$ is to be the matricial spectrum of some operator, then X must be hypoconvex; this necessary condition is also sufficient as well [18]. The *hypoconvex hull* of a bounded set $\mathcal{S} \subset \mathcal{M}_n$ is the smallest hypoconvex set to contain $\mathcal{S}$. Plainly, the hypoconvex hull of a compact set is compact.

THEOREM 2.4. *If $\mathcal{S} \subset \mathcal{M}_n$ is compact, then* $\operatorname{mconv}(\mathcal{S})$ *is compact and there is a $p \le n^3(2n^2+1)$ such that every element of* $\operatorname{mconv}(\mathcal{S})$ *is a* C^**-convex combination of elements of $\mathcal{S}$ involving at most p terms. Moreover,* $\operatorname{mconv}(\mathcal{S})$ *is the intersection of all matrix-valued discs containing $\mathcal{S}$.*

PROOF. The idea is to identify the set

$$\bigcap_{B,M \in \mathcal{M}_n} \left\{ \Lambda \in \mathcal{M}_n : \|\Lambda \otimes B + 1 \otimes M\| \le \sup\{ \|S \otimes B + 1 \otimes M\| : S \in \mathcal{S}\} \right\}$$

with the n^{th} matricial range of some operator (by using Lemma 2.1), and then to show that this is precisely $\operatorname{mconv}(\mathcal{S})$.

Let $\{A_j\}_{j=1}^{\nu} \subset \mathcal{S}$ be a finite or a countable subset which is dense in $\mathcal{S}$, let $A = \oplus_{j=1}^{\nu} A_j$ act on an appropriate space $\mathcal{H}$, and let $T = A^{(\infty)}$, countably many copies of A. Each summand A_j in T appears with infinite multiplicity, and thus the work of Salinas on the reducing essential matricial spectra for n-normal operators can be applied to T. (Using T in place of A is not necessary though: see [7]).

For every $B, M \in \mathcal{M}_n$,

$$\|T \otimes B + 1 \otimes M\| = \sup\{ \|A_j \otimes B + 1 \otimes M\| : j \in \mathbb{N}\} = \sup\{ \|S \otimes B + 1 \otimes M\| : S \in \mathcal{S}\},$$

and therefore by Lemma 2.1,

$$\bigcap_{B,M \in \mathcal{M}_n} \left\{ \Lambda : \|\Lambda \otimes B + 1 \otimes M\| \le \sup\{ \|S \otimes B + 1 \otimes M\| : S \in \mathcal{S}\} \right\} = W^n(T).$$

Because $W^n(T)$ is C^*-convex and contains $\mathcal{S}$, $\operatorname{mconv}(\mathcal{S}) \subset W^n(T)$.

To prove the converse, we argue as in the proof of [5;3.9] and employ a few more facts about the extremal structure of $\mathrm{CP}^n(T)$. Because $W^n(T)$ is compact and convex, $W^n(T)$ is the convex hull of its extreme points; thus, it is enough to prove that every extreme point of $W^n(T)$ is in $\operatorname{mconv}(\mathcal{S})$. By Lemma 2.3, if $\Lambda \in W^n(T)$ is an extreme point of $W^n(T)$, then $\Lambda = \phi(T)$ for some extreme point ϕ of $\mathrm{CP}^n(T)$. W. B. Arveson studied the extreme points of spaces of completely positive mappings in a general setting [1;1.4.6], and R. R. Smith and J. D. Ward applied this to the special case $\mathrm{CP}^n(T)$ [19;6.4] : the result is that ϕ is of the form

$$\phi(X) = V^*(\pi_1(X) \oplus \cdots \oplus \pi_q(X))V,$$

where each π_i is an irreducible representation of $C^*(T)$ on $\mathcal{H}_i$, $V : \mathbb{C}^n \longrightarrow \oplus_{i=1}^q \mathcal{H}_i$ is an isometry, and $q \leq n^2$. With respect to the direct sum $\oplus_{i=1}^q \mathcal{H}_i$, V has an orthogonal decomposition defined by $V\xi = V_1\xi \oplus \cdots \oplus V_q\xi$. Because T is n-normal, irreducible representations of $C^*(T)$ take place on spaces of dimension of at most n; hence, each $\Omega_i = \pi_i(T)$ can be viewed as a $k_i \times k_i$ matrix ($1 \leq k_i \leq n$), and each $\mathcal{H}_i$ as $\mathbb{C}^{k_i}$. Decompose each A_j as a direct sum of irreducibles; then $T \simeq \oplus_\beta T_\beta$, where each T_β is a $k_\beta \times k_\beta$ irreducible matrix with $k_\beta \leq n$. Now since each $\Omega_i \in R^{k_i}(T)$ is irreducible, [18;1.11] implies that $\Omega_i \simeq \Lambda_i$ for some Λ_i which is a limit of $k_i \times k_i$ matrices Γ_i^α from $\{T_\beta\}_\beta$. To each Γ_i^α there corresponds a $(n-k_i) \times (n-k_i)$ matrix Θ_i^α such that $\Gamma_i^\alpha \oplus \Theta_i^\alpha \simeq A_{k(i,\alpha)}$ for some $A_{k(i,\alpha)} \in \{A_j\}_{j\in\mathbb{N}}$. By dropping down to a subsequence if necessary, we may assume the matrices Θ_i^α converge to a matrix Λ_i', so that the matrices $\Gamma_i^\alpha \oplus \Theta_i^\alpha$ converge to $\Lambda_i \oplus \Lambda_i'$. From the results of [18;§ 1], we see that each $\Lambda_i \oplus \Lambda_i'$ is in the hypoconvex hull of $\mathcal{S}$. For each $i = 1, \ldots, q$ there is a representation π_i' on a $(n-k_i)$-dimensional Hilbert space $\mathcal{H}_i'$ such that $\Lambda_i' \simeq \pi_i'(T)$. Therefore, if $\mathcal{H}_0 = \oplus_{i=1}^q(\mathcal{H}_i \oplus \mathcal{H}_i')$, and if $W : \mathbb{C}^n \longrightarrow \mathcal{H}_0$ is the isometry defined by

$$W\xi = (V_1\xi \oplus 0) \oplus \cdots \oplus (V_q\xi \oplus 0),$$

then

$$\Lambda = W^*\Big((\pi_1(T) \oplus \pi_1'(T)) \oplus \cdots \oplus (\pi_q(T) \oplus \pi_q'(T)) \Big)W = \sum_{i=1}^q \tilde{V}_i^*(\Lambda_i \oplus \Lambda_i')\tilde{V}_i,$$

where $\tilde{V}_i = V_i \oplus 0$ and $\sum_i \tilde{V}_i^*\tilde{V}_i = 1_n$. This represents Λ as a C*-convex combination of elements $\Lambda_i \oplus \Lambda_i'$ from the hypoconvex hull of $\mathcal{S}$. This proves that $W^n(T) = \mathrm{mconv}(\mathcal{S})$, since the hypoconvex hull of $\mathcal{S}$ is contained in the matricial convex hull of $\mathcal{S}$.

Finally, an arbitrary element of $\mathrm{mconv}(\mathcal{S})$ is a convex combination of at most $2n^2+1$ extreme points of $\mathrm{mconv}(\mathcal{S})$ (by Carathéodory's theorem). An extreme point of $\mathrm{mconv}(\mathcal{S})$ is a C^*-convex combination of no more than n^2 (possibly non-distinct) elements of the hypoconvex hull of $\mathcal{S}$ (by the preceding arguments and the aforementioned Smith-Ward theorem). And, an element of the hypoconvex hull of $\mathcal{S}$ is obtained in a C^*-convex combination of at most n elements of $\mathcal{S}$ (by definition). Hence, an arbitrary element of $\mathrm{mconv}(\mathcal{S})$ can be obtained in a C^*-convex combination requiring at most $n^3(2n^2+1)$ (possibly non-distinct) elements of $\mathcal{S}$. ∎

COMMENT. It is likely that the estimate $n^3(2n^2+1)$ is not sharp; the existence of such an upperbound, however, is important.

COROLLARY 2.5. *If $\mathcal{S} \subset \mathcal{M}_n$ is bounded, then* $\mathrm{mconv}(\mathcal{S}^-) = (\mathrm{mconv}\,\mathcal{S})^-$.

PROOF. Since $\mathcal{S} \subset \mathcal{S}^-$, we have $\mathrm{mconv}(\mathcal{S}) \subset \mathrm{mconv}(\mathcal{S}^-)$ and $(\mathrm{mconv}(\mathcal{S}))^- \subset \mathrm{mconv}(\mathcal{S}^-)$, because the right hand side of the final containment relation is closed (by

https://doi.org/10.4153/CJM-1992-019-1 Published online by Cambridge University Press

the theorem). Conversely, suppose that $\Lambda \in \mathrm{mconv}(\mathcal{S}^-)$, and let $\epsilon > 0$ be given. Then there are $S_1, \ldots, S_p \in \mathcal{S}^-$ and C*-convex coefficients $T_1, \ldots, T_p \in \mathcal{M}_n$ such that $\Lambda = \sum_{j=1}^p T_j^* S_j T_j$. Choose $A_1, \ldots, A_p \in \mathcal{S}$ such that $\|A_j - S_j\| < \frac{\epsilon}{p}$ for each j. Then $\Gamma = \sum_{j=1}^p T_j^* A_j T_j \in \mathrm{mconv}(\mathcal{S})$ and

$$\|\Lambda - \Gamma\| = \Big\|\sum_{j=1}^p T_j^*(S_j - A_j)T_j\Big\| \leq \sum_{j=1}^p \|T_j\|^2 \|S_j - A_j\| < (\frac{\epsilon}{p})p = \epsilon.$$

Because $\epsilon > 0$ is arbitrary, $\Lambda \in (\mathrm{mconv}(\mathcal{S}))^-$. ∎

We now apply our results on C*-convex hulls to matricial range theory and to matrix theory.

A polynomial of the form $p(z) = \sum_{j=0}^m z^j C_j$ with indeterminate z and coefficients $C_j \in \mathcal{M}_n$ is said be an $n \times n$ matrix-polynomial of degree m; matrix-polynomials of degree 1 are called linear matrix-polynomials. The evaluation of a matrix-polynomial p at an element T in a C^*-algebra $\mathcal{A}$ produces an element in $\mathcal{A} \otimes \mathcal{M}_n$: $p(T) = \sum_{j=0}^m T^j \otimes C_j$. It is well known that if $T \in \mathcal{A}$ is such that the norm of every linear (scalar) polynomial $\beta T + \mu 1$ in T coincides with the maximum of the numbers $|\beta \lambda + \mu|$ as λ varies throughout $\sigma(T)$, then $W^1(T) = \mathrm{conv}\, \sigma(T)$. Lemma 2.1 and Theorem 2.4 combine to show that if the norm of every linear $n \times n$ matrix-polynomial in T is the maximum of the norms obtained by evaluating the matrix-polynomial at each point of the matricial spectrum of T, then $W^n(T)$ is the C*-convex hull of $\sigma^n(T)$.

THEOREM 2.6. *If $T \in \mathcal{A}$ satisfies, for every $B, M \in \mathcal{M}_n$,*

$$\|T \otimes B + 1 \otimes M\| = \max\big\{\|\Lambda \otimes B + 1 \otimes M\| : \Lambda \in \sigma^n(T)\big\},$$

then $W^n(T) = \mathrm{mconv}\, \sigma^n(T)$.

The following example presents a class of operators which satisfy the matricial conditions of Theorem 2.6, and the example can be viewed as being the matricial analogue of the reason why the numerical ranges of, say, Toeplitz or subnormal operators are the convex hulls of their spectra: the key is that these operators have normal dilations (or extensions) which satisfy a spectral inclusion. As a consequence, we obtain, in the case $n = 1$, a slightly sharper form of [2;2.4.1] and a unification of some diverse results appearing in [15].

EXAMPLE 1. *If $T \in \mathcal{B}(\mathcal{H})$ has an n-normal dilation A such that $R^n(A) \subset \sigma^n(T)$, then*

$$W^{kn}(T) = \mathrm{mconv}\, \sigma^{kn}(T) = W^{kn}(A) \textit{ for all } k \in \mathbb{N}.$$

For the case $n = 1$, $W^k(T) = \mathrm{mconv}\ R^k(T)$ for every $k \in \mathbb{N}$, and $W^k(T)$ consists of all matrices of the form $\sum_{i=1}^p \lambda_i H_i$, where $p \leq k^3(2k^2+1)$, each $\lambda_i \in \sigma(T)$, each $H_i \geq 0$, and $\sum_{i=1}^p H_i = 1_k$.

PROOF. The central part of the proof concerns the computation of $\|A \otimes B + 1 \otimes M\|$ for $B, M \in \mathcal{M}_{kn}$ and n-normal A. This computation can be performed by employing the generalized Gelfand theories of [18;§2] or [12;Lemma 4.1], but a very transparent proof is given here. We will assume, however, that the space on which A acts is separable.

The Weyl-von Neumann Theorem for n-normal operators [14;3.5] states that for each $\epsilon > 0$ there exist operators S and K such that S is unitarily equivalent to a direct sum of $n \times n$ matrices, K is compact, $R^n(S) \subset R^n(A)$, $\|K\| < \epsilon$, and $A = S + K$. Therefore, if $B, M \in \mathcal{M}_{kn}$, then

$$\Big| \|A \otimes B + 1 \otimes M\| - \|S \otimes B + 1 \otimes M\| \Big| \leq \|K \otimes B\| < \epsilon \|B\|.$$

On the other hand, because S is unitarily equivalent to a block diagonal n-normal operator and because of [18;§1], $\|S \otimes B + 1 \otimes M\|$ is the maximum of the numbers $\|\Lambda \otimes B + 1_n \otimes M\|$ as Λ ranges through $R^n(S)$. As ϵ is arbitrary and $R^n(S) \subset R^n(A)$, we conclude that

$$(\dagger) \qquad \|A \otimes B + 1 \otimes M\| = \sup\Big\{ \|(\Lambda \otimes 1_k) \otimes B + 1_{kn} \otimes M\| : \Lambda \in R^n(A)\Big\}.$$

The characterization of $\operatorname{mconv}\{\Lambda \otimes 1_k : \Lambda \in R^n(A)\}$ given by Theorem 2.4 and (†) is

$$\bigcap_{B,M \in \mathcal{M}_{kn}} \{\Omega \in \mathcal{M}_{kn} : \|\Omega \otimes B + 1 \otimes M\| \leq \|A \otimes B + 1 \otimes M\| \}.$$

But this, by Lemma 2.1, is precisely $W^{kn}(A)$. Hence,

$$\begin{aligned} W^{kn}(T) \subset W^{kn}(A) &= \operatorname{mconv}\{\Lambda \otimes 1_k : \Lambda \in R^n(A)\} \\ &\subset \operatorname{mconv}\{\Lambda \otimes 1_k : \Lambda \in \sigma^n(T)\} \\ &\subset \operatorname{mconv} \sigma^{kn}(T) \subset W^{kn}(T). \end{aligned}$$

For the case $n = 1$, the hypothesis is that T has a normal dilation A such that $\sigma(A) \subset \sigma(T)$; hence, $W^1(T) = \operatorname{conv} \sigma(T)$. Consequently, every extreme point of $W^1(T)$ is both a spectral point of T and a boundary point of $W^1(T)$. By [1;3.1.2], to every extreme point $\zeta \in W^1(T)$ there corresponds a unique unital $*$-homomorphism (i.e., a character) $\rho : C^*(T) \longrightarrow \mathbb{C}$ such that $\rho(T) = \zeta$. Thus,

$$\begin{aligned} \operatorname{mconv}\big\{\rho(T) : \rho = \rho_1 \oplus \cdots \oplus \rho_k, \text{ where } \rho_i \text{ is a character of } C^*(T)\big\} \\ \subset \operatorname{mconv} R^k(T). \end{aligned}$$

The former set contains the C*-convex hull of $\{\lambda 1_k : \lambda \in \sigma(T)\}$ because extreme points of $W^1(T)$ correspond to characters of $C^*(T)$; hence,

$$W^k(T) \subset W^k(A) = \operatorname{mconv}\big\{\zeta 1_k : \zeta \in \sigma(A)\big\} \subset \operatorname{mconv} R^k(T) \subset W^k(T).$$

■

The second application of the ideas developed in this section concerns representations of matrices with respect to their spectral sets. Recall that a set $X \subset \mathbb{C}$ is a spectral set for an operator $T \in \mathcal{B}(\mathcal{H})$ if for every rational function r with poles off X,

$$\|r(T)\| \leq \sup\big\{|r(\zeta)| : \zeta \in X\big\}.$$

A classical theorem of von Neumann asserts that the closed unit disc $\mathbb{D}^-$ is a spectral set for every contractive operator on a Hilbert space. Combined with the fact that every matrix Λ in the closed unit ball of $\mathcal{M}_n$ is a convex combination of unitary matrices, we may say that every matrix which has the closed unit disc $\mathbb{D}^-$ as a spectral set can be expressed as a convex combination of normal matrices with spectra on $\partial\mathbb{D}^-$. It is natural to ask whether this is also true for arbitrary spectral sets, or for simply connected spectral sets: if $X \subset \mathbb{C}$ is a compact (and perhaps simply connected) spectral set for a matrix $\Lambda \in \mathcal{M}_n$, is Λ a convex combination of normals with spectra on ∂X? In the example below, C*-convex combinations are seen to produce sufficiently many matrices for a weaker form of the question to be answered affirmatively: in this case, no particular connectedness properties are required of the spectral set X.

EXAMPLE 2. *If $X \subset \mathbb{C}$ is a compact spectral set for a matrix $\Lambda \in \mathcal{M}_n$, then Λ is a* C**-convex combination of normal matrices having spectra on ∂X.*

PROOF. If $Y = \operatorname{conv} X$, then Y is also a spectral set for Λ. Because Y is convex and compact, the Berger-Foiaş-Lebow dilation theorem [13;4.4] asserts the existence of a (rational) normal dilation N_0 of Λ on a Hilbert space $\mathcal{H}$ such that $\sigma(N_0) \subset \partial Y$. Now let E be the closure of the set of extreme points of Y, and note that $E \subset \partial X$. Let $N \in \mathcal{B}(\mathcal{H})$ be a normal operator with spectrum E. Since $\sigma(N_0) \subset \operatorname{conv} E$, by Example 1 we have that

$$W^n(N_0) = \operatorname{mconv}\{\zeta 1_n : \zeta \in \sigma(N_0)\} \subset \operatorname{mconv}\{\zeta 1_n : \zeta \in E\} = W^n(N).$$

Hence, $\Lambda \in W^n(N_0) \subset W^n(N)$, and so Λ is a C*-convex combination of normals with spectra on E. ∎

3. **Convexity theorems for the spatial matricial range.** In this section, the C^*-algebras that are involved are concrete: operators on Hilbert spaces. We will assume throughout that all Hilbert spaces are separable, though no further restrictions on their dimension will be assumed.

We begin this section with a result concerning convexity rather than C*-convexity, and its proof requires the use of a generalized numerical range known as the C-numerical range. If $C \in \mathcal{B}(\mathcal{H})$ is a fixed trace class operator, then the *C-numerical range* of $A \in \mathcal{B}(\mathcal{H})$ is the set of all complex numbers of the form $\operatorname{tr}(CU^*AU)$, where U ranges over all unitary operators in $\mathcal{B}(\mathcal{H})$, and where 'tr' denotes the trace. (For an operator A in the ideal of trace class operators, the trace of A is $\operatorname{tr}(A) = \sum_i (A\phi_i, \phi_i)$, where $\{\phi_i\}_i$ is any orthonormal basis of $\mathcal{H}$.) Recall that a self-adjoint operator $T \in \mathcal{B}(\mathcal{H})$ is said to have a complete system of eigenvectors if $\mathcal{H}$ has an orthonormal basis consisting of eigenvectors of T.

THEOREM 3.1. *If $T \in \mathcal{B}(\mathcal{H})$ is a self-adjoint operator, then every element of $W_s^n(T)^-$ is a convex combination of elements of $\sigma_s^n(T)$.*

PROOF. Assume, initially, that T has finite spectrum; therefore, its spectrum consists solely of eigenvalues, and the corresponding system of eigenvectors is complete.

Moreover, the unitary orbit of T is closed [3;3.1], and consequently, the spatial matricial ranges of T are closed as well.

The convex hull of a compact set $\mathcal{S}$ in $\mathcal{M}_n$ is the intersection of all closed half-spaces containing $\mathcal{S}$; since $\operatorname{conv}\sigma_s^n(T) \subset \operatorname{conv} W_s^n(T)$, to prove the theorem, it is enough to show that every closed half-space $\mathcal{E}$ in $\mathcal{M}_n$ which contains $\sigma_s^n(T)$ also contains $W_s^n(T)$. Denote by g_C the linear functional on $\mathcal{M}_n$ defined by $g_C(\Gamma) = \operatorname{tr}(C\Gamma)$, where $C \in \mathcal{M}_n$, and let $\Re(g_C)$ denote the real part of g_C. Every half-space $\mathcal{E}$ in $\mathcal{M}_n$ is determined by some pair $(\alpha, C) \in \mathbb{R} \times \mathcal{M}_n$ via

$$\mathcal{E} = \{\Gamma \in \mathcal{M}_n : \Re\big(g_C(\Gamma)\big) \leq \alpha\}.$$

Thus, suppose that $(\alpha, C) \in \mathbb{R} \times \mathcal{M}_n$ induce a half-space $\mathcal{E}$ such that $\sigma_s^n(T) \subset \mathcal{E}$. To prove $W_s^n(T) \subset \mathcal{E}$, it is necessary to establish the inequality

$$\Re\big(\operatorname{tr}(CV^*TV)\big) \leq \alpha, \text{ for every isometry } V\colon \mathbb{C}^n \longrightarrow \mathcal{H}.$$

Embed $\mathbb{C}^n$ into $\mathcal{H}$ so that $\mathcal{H} = \mathbb{C}^n \oplus \mathcal{H}_0$. With respect to this decomposition, consider the finite-rank (and trace class) operator $\tilde{C} = C \oplus 0$. The set $g_C\big(W_s^n(T)\big)$ can be identified with the $\tilde{C}$-numerical range of T by way of the equality $\operatorname{tr}(CV^*TV) = \operatorname{tr}(\tilde{C}U^*TU)$, whenever V is an isometric map $\mathbb{C}^n \longrightarrow \mathcal{H}$; here, $U \in \mathcal{B}(\mathcal{H})$ is any unitary extension of V (i.e., matricially, $U = [V, W]\colon \mathcal{H} \longrightarrow \mathcal{H}$, for some W). Because T is a self-adjoint operator with a complete system of eigenvectors, the $\tilde{C}$-numerical range of T is given by

$$g_C\big(W_s^n(T)\big) = \Big\{\sum_{j=1}^{\nu} \lambda_j(\tilde{C}u_j, u_j) : \{u_j\}_{j=1}^{\nu} \text{ is an orthonormal basis of } \mathcal{H}\Big\},$$

where $\{\lambda_j\}_{j=1}^{\nu}$ is a list of the eigenvalues of T, repeated according to multiplicity. This characterization of $g_C\big(W_s^n(T)\big)$ is obtained by observing that the $\tilde{C}$-numerical range of T is the T-numerical range of $\tilde{C}$, and in turn, this is the D-numerical range of $\tilde{C}$, where D is the diagonal operator with eigenvalues $\{\lambda_i\}_{i=1}^{\nu}$. Similarly, applying g_C to the unitary orbit of any $\Lambda \in \sigma_s^n(T)$ produces the C-numerical range of Λ:

$$(\dagger) \qquad \Big\{\sum_{j=1}^{n} \lambda_{i_j}(Cx_j, x_j) : x_1, \ldots, x_n \in \mathbb{C}^n \text{ are orthonormal}\Big\},$$

where $\sigma(\Lambda) = \{\lambda_{i_1}, \ldots, \lambda_{i_n}\} \subset \sigma(T)$. Because the spatial matricial range $W_s^n(T)$ is closed and path connected, the set $\Re g_C\big(W_s^n(T)\big)$ must be a compact interval in $\mathbb{R}$. By using our preceding arguments and the fact that the spectrum of T is real, this compact interval is the D-numerical range of $\Re(\tilde{C})$:

$$\Re g_C\big(W_s^n(T)\big) = \Big\{\sum_{j=1}^{\nu} \lambda_j(\Re(\tilde{C})u_j, u_j) : \{u_j\}_{j=1}^{\nu} \text{ is an orthonormal basis of } \mathcal{H}\Big\}.$$

Let $\omega = \operatorname{tr}(DU^*\Re(\tilde{C})U)$ be the maximum value of this compact interval. The arguments given by Li in [9;2.1], which show that D and $U^*\Re(\tilde{C})U$ commute if $\mathcal{H}$ is finite-dimensional, are dimension independent (they depend only on the existence of a trace) and, therefore, are valid here as well. Hence, D and $U^*\Re(\tilde{C})U$ are simultaneously diagonalizable; thus,

$$\omega = \sum_{j=1}^{\nu} \lambda_{\gamma(j)}\zeta_j,$$

where $\{\zeta_j\}_{j=1}^{\nu} = \sigma\big(\Re(\tilde{C})\big) = \sigma\big(\Re(C)\big) \cup \{0\}$ and γ is a bijection on the set with cardinal number ν.

Without loss of generality, assume $\sigma\big(\Re(C)\big) = \{\zeta_1, \ldots, \zeta_n\}$, and consider $\mathcal{H}$ as $\mathbb{C}^n \oplus \mathcal{H}_0$. Let $\{u_j\}_{j=1}^{\nu}$ be an orthonormal basis of $\mathcal{H}$ such that $\{u_1, \ldots, u_n\}$ both spans $\mathbb{C}^n$ (as a subspace of $\mathcal{H}$) and diagonalizes $\Re(C)$. Therefore, ω can be expressed as

$$\omega = \sum_{j=1}^{\nu} \lambda_{\gamma(j)}(\Re(\tilde{C})u_j, u_j) = \Re\Big(\sum_{j=1}^{n} \lambda_{\gamma(j)}(Cu_j, u_j)\Big).$$

But the right hand side of this expression is, by equation (†), precisely $\Re\big(g_C(\Lambda)\big)$ for some $\Lambda \in \sigma_s^n(T)$ with spectrum $\sigma(\Lambda) = \{\lambda_{\gamma(j)} : 1 \leq j \leq n\}$. Consequently, for all isometries $V\colon \mathbb{C}^n \longrightarrow \mathcal{H}$,

$$\Re\big(\mathrm{tr}(CV^*TV)\big) \leq \omega = \Re\big(\mathrm{tr}(C\Lambda)\big) \leq \alpha;$$

that is, the half-space $\mathcal{E}$ contains $W_s^n(T)$.

To complete the proof, we now consider the situation for T having infinite spectrum. Suppose that $\Lambda \in W_s^n(T)^-$, and suppose that $\epsilon > 0$ is arbitrary. By Lemma 1, there is a self-adjoint operator A with finite spectrum such that A approximates T to within $\frac{\epsilon}{2}$ and such that $\sigma_s^n(A) \subset \sigma_s^n(T)$. By hypothesis, there is an isometry $V\colon \mathbb{C}^n \longrightarrow \mathcal{H}$ such that $\|\Lambda - V^*TV\| < \frac{\epsilon}{2}$; so, $\|\Lambda - V^*AV\| < \epsilon$. Because A has finite spectrum, V^*AV is a convex combination of elements of $\sigma_s^n(A)$, and therefore, from $\sigma_s^n(A) \subset \sigma_s^n(T)$, Λ is within ϵ of $\mathrm{conv}\,\sigma_s^n(T)$. ∎

It should be noted that Theorem 3.1 is false if "self-adjoint" is replaced by "normal" in its hypothesis: the nilpotent Jordan matrix of order n is an element of the $n \times n$ spatial matricial range of the cyclic shift operator U on $\mathbb{C}^{n+1}$, but it does not lie within the convex hull of $\sigma_s^n(U)$ [7].

The next objective is to establish Theorem 3.2. The first step is to extend a theorem of Poon concerning a vector-valued range to a wider class of operators and spaces. An extension of this form in finite dimensions has been given by Li and it will appear in forthcoming work; the presentation of the proof given below is based somewhat on his. For $n \in \mathbb{N}$ not exceeding the dimension of $\mathcal{H}$, let $D_s^n(T)$ denote the set of n-tuples in $\mathbb{C}^n$ of the form $\big((T\phi_1, \phi_1), \ldots, (T\phi_n, \phi_n)\big)$, where $\phi_1, \ldots, \phi_n$ are n orthonormal vectors from $\mathcal{H}$; equivalently, $D_s^n(T)$ is the set of n-tuples obtained from the diagonals of the elements of $W_s^n(T)$.

THE DIAGONAL LEMMA OF Y.-T. POON. *If $T \in \mathcal{B}(\mathcal{H})$ is an operator which is not a linear function of a self-adjoint operator, but which does satisfy* $W^1(T) = \mathrm{conv}\,\sigma(T)$,

https://doi.org/10.4153/CJM-1992-019-1 Published online by Cambridge University Press

then the closure of the set $D^n_s(T)$ cannot be convex if T has an isolated eigenvalue of multiplicity $k < n$ among the extreme points of $W^1_s(T)$.

PROOF. The proof is modelled on the original [16]: we show that, under the hypothesis, the closure of $D^{k+1}_s(T)$ is not convex.

Suppose that $\lambda \in \sigma(T)$ is an isolated eigenvalue of multiplicity $k < n$, and that λ is an extreme point of $W^1_s(T)$. Because λ is eigenvalue on the boundary of the numerical range of T, λ is a reducing eigenvalue [8; Satz 2(i)]. Moreover, because λ is isolated in $\sigma(T)$, T is unitarily equivalent to $\lambda 1_k \oplus B$, where B is an operator on $\mathcal{H}_0$, the orthogonal complement of $\ker(T - \lambda 1)$, and where $\lambda \notin \sigma(B)$. Hence, by the hypothesis $W^1(T) = \operatorname{conv}\sigma(T)$, $\lambda \notin W^1(B)$, and so $W^1(T)$ is the convex hull of (the disjoint union) $\{\lambda\} \cup W^1(B)$. There exist, therefore, extreme points γ_1, γ_2 of $W^1(T)$, which are elements of $\sigma(T)$ as well, such that

(i) γ_1, γ_2 are adjacent to λ (i.e., they are the nearest extreme points to λ).

Because $\sigma(T)$ is not contained within a line, it is also true that

(ii) $\gamma_1 \neq \gamma_2$.

The proof of the theorem consists of showing that the convexity of the closure of $D^{k+1}_s(T)$ implies the contradictory statement $\gamma_1 = \gamma_2$.

At this point, we know that $\gamma_1, \gamma_2 \in \sigma(T)$; however, it is of added benefit for our arguments if the points are elements of the spatial numerical range. Therefore, assume, for the moment, that $W^n_s(T)$ is closed, so that for each $1 \leq j \leq n$ the sets $W^j_s(T)$ and $D^j_s(T)$ are closed (note that this also implies that $W^1_s(T) = W^1(T)$). Because $W^1_s(T)$ is closed there exist unit vectors $x_1, x_2 \in \mathcal{H}$ such that $\gamma_i = (Tx_i, x_i)$ for $i = 1, 2$; moreover, because λ is a reducing eigenvalue, x_1 and x_2 can be taken to be orthogonal to $\ker(T - \lambda 1)$. Therefore, the $(k+1)$-tuples

$$\alpha = \big(\lambda, \lambda, \ldots, \lambda, \gamma_2, \lambda\big) \text{ and } \beta = \big(\lambda, \lambda, \ldots, \lambda, \lambda, \gamma_1\big) \in D^{k+1}_s(T).$$

The remaining arguments follow those of [16;Thm. 1]. To show that $D^{k+1}_s(T)$ is not convex, it suffices to show that the assumption $\xi = \frac{1}{2}(\alpha + \beta) \in D^{k+1}_s(T)$ leads to a contradiction of $\gamma_1 \neq \gamma_2$. To argue this, assume that ξ is an element of $D^{k+1}_s(T)$; thus, there exist orthonormal vectors $\psi_1, \ldots, \psi_{k+1} \in \mathcal{H}$ such that $\xi_i = \lambda = (T\psi_i, \psi_i)$ for $i = 1, \ldots, k-1$, $\xi_k = \frac{1}{2}(\lambda + \gamma_2) = (T\psi_k, \psi_k)$, and $\xi_{k+1} = \frac{1}{2}(\lambda + \gamma_1) = (T\psi_{k+1}, \psi_{k+1})$. By the extremal theory for numerical ranges (e.g. [8]), the vectors $\psi_1, \ldots, \psi_{k-1}$ must be eigenvectors of T corresponding to λ. With respect to the decomposition $\mathcal{H} = \ker(T - \lambda 1) \oplus \mathcal{H}_0$, express the vectors ψ_k and ψ_{k+1} as $\psi_k = \theta_k \oplus \omega_k$ and $\psi_{k+1} = \theta_{k+1} \oplus \omega_{k+1}$. Thus,

$$\begin{aligned} \xi_k = (T\psi_k, \psi_k) &= \lambda \|\theta_k\|^2 + (T\omega_k, \omega_k) \\ &= \lambda \|\theta_k\|^2 + \|\omega_k\|^2 (T\tilde{\omega}_k, \tilde{\omega}_k) \end{aligned}$$

and

$$\begin{aligned} \xi_{k+1} = (T\psi_{k+1}, \psi_{k+1}) &= \lambda \|\theta_{k+1}\|^2 + (T\omega_{k+1}, \omega_{k+1}) \\ &= \lambda \|\theta_{k+1}\|^2 + \|\omega_{k+1}\|^2 (T\tilde{\omega}_{k+1}, \tilde{\omega}_{k+1}), \end{aligned}$$

where $\tilde{x}$ denotes $\|x\|^{-1}x$ for non-zero $x \in \mathcal{H}$. The first equation expresses ξ_k as a convex combination of λ and $(T\tilde{\psi}_k, \tilde{\psi}_k)$; however, ξ_k is on the line segment connecting the adjacent extreme points λ and γ_2, and so $\|\theta_k\|^2 = \|\omega_k\|^2 = \frac{1}{2}$. Similarly, $\|\theta_{k+1}\|^2 = \|\omega_{k+1}\|^2 = \frac{1}{2}$. The vectors θ_k, θ_{k+1} are elements of the k-dimensional T-invariant subspace $\ker(T - \lambda 1)$, and both are orthogonal to $\psi_1, \ldots, \psi_{k-1}$; hence, $\tilde{\theta}_k = \zeta\tilde{\theta}_{k+1}$ for some ζ of unit modulus. Moreover, from $0 = (\psi_k, \psi_{k+1})$, we conclude that $(\omega_k, \omega_{k+1}) = -(\theta_k, \theta_{k+1})$, and thus,

$$\frac{1}{2} = \|\omega_k\| \, \|\omega_{k+1}\| \geq |(\omega_k, \omega_{k+1})| = \|\theta_k\| \, \|\theta_{k+1}\| \, |(\tilde{\theta}_k, \tilde{\theta}_{k+1})| = \frac{1}{2}|\zeta| = \frac{1}{2}.$$

But this is equality in the Cauchy-Schwarz inequality, and so $\omega_k = \rho\omega_{k+1}$ for some $|\rho| = 1$. In the equation

$$\frac{1}{2}\gamma_1 = (T\omega_k, \omega_k) = |\rho|(T\omega_{k+1}, \omega_{k+1}) = \frac{1}{2}\gamma_2,$$

we have a contradiction, for this implies that $\gamma_1 = \gamma_2$.

All that it is left is to consider the case where $W_s^n(T) = W_s^n(\lambda 1_k \oplus B)$ is not closed. In this case, replace B with $\pi(B)$ for some unital $*$-representation π of $C^*(B)$ for which $W_s^n(\pi(B))$ is closed and $W_s^n(B)^- \subset W_s^n(\pi(B))$ (such a representation is constructed in [4; pp. 150–155]). Observe that λ is an isolated eigenvalue of multiplicity k in $\sigma(\lambda 1_k \oplus \pi(B))$. An application of the preceding arguments to the operator $\lambda 1_k \oplus \pi(B)$ shows that no convex combination of elements of (the closed set) $D_s^{k+1}(\lambda 1_k \oplus \pi(B))$ can produce ξ; but since $\xi \in D_s^{k+1}(T)^- \subset D_s^{k+1}\big(\lambda 1_k \oplus \pi(B)\big)$, neither can any convex combination of elements of $D_s^{k+1}(T)^-$ produce ξ. ■

The preceding lemma is a statement concerning the convexity of the diagonals of $W_s^n(T)^-$, based on the numerical range of T sharing a property common to normal operators. To say something about the convexity of $W_s^n(T)^-$ itself, we will require T to behave very much like a normal operator behaves with respect to all of its matricial ranges; consequently, we will consider operators T which possess a normal dilation N such that $\sigma(N) \subset \sigma(T)$.

THEOREM 3.2. *If $T \in \mathcal{B}(\mathcal{H})$ has a normal dilation N such that $\sigma(N) \subset \sigma(T)$, then the following statements are equivalent:*

(i) $W_s^n(T)$ is dense in a convex set.

(ii) $W_s^n(T)$ is dense in a C^*-convex set.

(iii) Every isolated eigenvalue of T among the extreme points of $W_s^1(T)$ is of multiplicity no less than n.

PROOF. (i) $\Rightarrow$ (iii) The assumption that $W_s^n(T)^-$ is convex implies that $D_s(T)^-$ is convex as well; hence, if T is not a linear function of a self-adjoint operator, then Poon's Diagonal Lemma shows that the closure of $W_s^n(T)$ is convex only if (iii) holds. If T is a linear function of a self-adjoint operator, then there is no loss in assuming that $T \geq 0$, and in proving (iii) only for $\lambda = \|T\|$. In the finite-dimensional case, the equivalence of (i) and (iii) is covered (for self-adjoint operators) by [10;2.1], and so we assume that

$\mathcal{H}$ is of infinite dimension. Because (iii) is a statement about extreme points of $W^1(T)$ which are not elements of $\sigma_e(T)$, we assume, in addition, that λ is an isolated eigenvalue.

Because $T^* = T$ and because $W_s^n(T)^-$ is convex, we have $W_s^n(T)^- = \operatorname{conv} \sigma_s^n(T)$, by Theorem 3.1. Bauer's Maximum Principle states that the maximum value of the (convex) function $\operatorname{tr} : W_s^n(T)^- \longrightarrow \mathbb{R}$ is attained at an extreme point of $W_s^n(T)^-$; hence, there is a $\Lambda \in \sigma_s^n(T)$ such that $\operatorname{tr}(\Lambda) \geq \operatorname{tr}(\Omega)$ for every $\Omega \in W_s^n(T)^-$. Because $\|T\| \notin \sigma_e(T)$, an $n \times n$ spectral element of maximal trace must be unitarily equivalent to a matrix of the form

$$\lambda_1 1_{k_1} \oplus \lambda_2 1_{k_2} \oplus \cdots \oplus \lambda_p 1_{k_p},$$

where $\|T\| = \lambda_1 > \lambda_2 > \cdots > \lambda_p$, and where each λ_i is an isolated eigenvalue of multiplicity k_i, for $1 \leq i < p$ (λ_p could be an element of $\sigma_e(T)$). Let $\alpha = \frac{1}{n} \operatorname{tr}(\Lambda)$; there exist unitaries $U_1, \ldots, U_n \in \mathcal{M}_n$ such that

$$\alpha 1_n = \sum_{i=1}^{n} \frac{1}{n} U_i^* \Lambda U_i.$$

Hence, $\alpha 1_n \in W_s^n(T)^-$, by the convexity of $W_s^n(T)^-$. If $\alpha \leq \lambda_p$, then $\alpha = \lambda_1 = \cdots = \lambda_p$ because in addition to $\alpha \leq \lambda_i$ for every $1 \leq i \leq p$, α is a convex combination of $\lambda_1, \ldots, \lambda_p$. Therefore, the inequality $\alpha \leq \lambda_p$ is enough to imply (iii), since this inequality implies $\lambda_1 1_n \in W_s^n(T)$, and so there exist n orthonormal vectors $x_1, \ldots, x_n \in \mathcal{H}$ with $\lambda_1 = \|T\| = (Tx_i, x_i)$ for each $1 \leq i \leq n$; hence, $Tx_i = \lambda_1 x_i$ for every $1 \leq i \leq n$, thereby implying that $\ker(T - \lambda_1 1)$ is at least n-dimensional.

We claim: $\alpha \leq \lambda_p$. Let E denote the spectral measure for T, and for use in the argument, let G denote the spectral measure for $\alpha 1_n$. Let B be a compression of T and let F be the spectral measure for B. By the generalized minimax principle,

$$\operatorname{rank} F\big((\lambda, \infty)\big) \leq \operatorname{rank} E\big((\lambda, \infty)\big) \text{ for every } \lambda \in \mathbb{R} \ [6;1(\mathrm{i})].$$

In particular, if $\lambda \geq \lambda_p$, then

$$\operatorname{rank} F\big((\lambda, \infty)\big) \leq \sum_{i<p} k_i = n - k_p < n.$$

The element $\alpha 1_n \in W_s^n(T)^-$ is the limit of a sequence of compressions of T, and therefore, by the lower semicontinuity of rank, we have that

$$n = \operatorname{rank} G\big(\{\alpha\}\big) = \operatorname{rank} G\big((\alpha - \epsilon, \infty)\big) \leq \operatorname{rank} E\big((\alpha - \epsilon, \infty)\big)$$

for every $\epsilon > 0$. Hence, we must have $\alpha - \epsilon < \lambda_p$ for every $\epsilon > 0$, because $\operatorname{rank} E\big((\lambda_p, \infty)\big) < n$.

(iii) $\Rightarrow$ (ii) It is sufficient to prove that the closure of $W_s^n(T)$ is the compact C*-convex set $\mathcal{K}$ of matrices of the form $\sum_{i=1}^{p} \lambda_i H_i$, where $p \leq n^3(2n^2+1)$, each $\lambda_i \in \sigma(T)$, each $H_i \geq 0$ in $\mathcal{M}_n$, and $\sum_{i=1}^{p} H_i = 1$. Because T has a normal dilation N with $\sigma(N) \subset \sigma(T)$, $\mathcal{K}$ contains the closure of $W_s^n(T)$, by Example 1 of § 2.

Conversely, suppose that $\Lambda = \sum_{i=1}^{p} \lambda_i H_i \in \mathcal{K}$, and suppose that $\epsilon > 0$; we are to show that Λ is within ϵ of some element of $W_s^n(T)$. Assume, to begin with, that

$\lambda_1, \ldots, \lambda_p$ are extreme points of $W^1(T)$. Because each $\lambda_i \in \partial W^1(T) \cap \sigma(T)$, the points $\lambda_1, \ldots, \lambda_p \in R_s^1(T)$ [8; Satz 2]. (It is the properties of $R_s^1(T)$ which allow us, in the absence of a spectral theorem for T, to treat $\lambda_1, \ldots, \lambda_p$ just as though they were spectral points of a normal operator.) An element of $R_s^1(T)$ is either a reducing isolated eigenvalue or a reducing essential eigenvalue [17; 6.1]. Thus, there exist mutually orthogonal projections $Q_1, \ldots, Q_p \in \mathcal{B}(\mathcal{H})$ such that for every $1 \leq i \leq p$ the operators $(T - \lambda_i 1)Q_i$ and $(T^* - \lambda_i^* 1)Q_i$ are compact and are of norm no greater than $p^{-2}\epsilon$ [17;4.1,4.2]; moreover, by condition (iii), the dimension of each subspace $\mathcal{H}_i = Q_i(\mathcal{H})$ is no less than n. These facts imply that for each $1 \leq i \leq p$ there is an isometry $V_i : \mathbb{C}^n \longrightarrow \mathcal{H}_i$ satisfying $\|(T - \lambda_i 1)V_i\| < p^{-2}\epsilon$. The matrix Λ is within ϵ of $V^*TV \in W_s^n(T)$, where V is the isometry defined by $V = \sum_j V_j H_j^{\frac{1}{2}}$, since

$$\begin{aligned} \|V^*TV - \Lambda\| &= \Big\|\sum_j \sum_i H_i^{\frac{1}{2}} V_i^* T V_j H_j^{\frac{1}{2}} - \sum_j \sum_i H_i^{\frac{1}{2}} V_i^* \lambda_j V_j H_j^{\frac{1}{2}}\Big\| \\ &\leq \sum_j \sum_i \|H_i^{\frac{1}{2}}\| \, \|H_j^{\frac{1}{2}}\| \, \|V_i^*(TV_j - \lambda_j V_j)\| \\ &< p^2\Big(\frac{\epsilon}{p^2}\Big) = \epsilon. \end{aligned}$$

This proves that Λ is in the closure of $W_s^n(T)$.

Now suppose that Ω is an arbitrary element of $\mathcal{K}$, and that Ω is of the form $\Omega = \sum_{j=1}^p \omega_j H_j$. Each ω_j is an element of the convex hull of the extreme points of $W^1(T)$; furthermore, an application of Carathéodory's Theorem produces, for each $1 \leq j \leq p$, the convex combination $\omega_j = \sum_{i=1}^3 t_i^j \lambda_i^j$, where each $t_i^j \in [0, 1]$ with $\sum_{i=1}^3 t_i^j = 1$, and where the points λ_i^j are extreme points of $W^1(T)$. Let $K_i^j = t_i^j H_j$; then each

$$K_i^j \geq 0, \quad \sum_{j=1}^{p} \sum_{i=1}^{3} K_i^j = 1_n, \quad \text{and} \quad \Omega = \sum_{j=1}^{p} \sum_{i=1}^{3} \lambda_i^j K_i^j.$$

In this form, we see that $\Omega \in W_s^n(T)^-$, by the preceding paragraph.

(ii) $\Rightarrow$ (i) This is obvious. ■

Whether it is true, for general operators T, that $W_s^n(T)^-$ cannot be convex without already being C^*-convex is not known. It seems, though, that there ought to be some sort of analogue of Theorem 3.2 for the $kn \times kn$ spatial matricial ranges of n-normal operators.

REFERENCES

1. W. B. Arveson, *Subalgebras of C^*-algebras*, Acta Math. **123**(1969), 141–224.

2. ———, *Subalgebras of C^*-algebras II*, Acta Math. **129**(1972), 271–308.

3. E. Azoff and Ch. Davis, *On distances between unitary orbits of self-adjoint operators*, Acta Sci. Math. (Szeged) **47**(1984), 419–439.

4. F. Bonsall and J. Duncan, *Numerical Ranges II*. London Math. Soc. Lecture Note Series, no. 10, Cambridge University Press, Cambridge, 1972.

5. J. Bunce and N. Salinas, *Completely positive maps on C^*-algebras and the left matricial spectrum of an operator*, Duke Math. J. **43**(1976), 747–774.

6. Ch. Davis, *Beyond the minimax principle*, Proc. Amer. Math. Soc. **81**(1981), 401–405.

7. D. R. Farenick, *The matricial spectrum and range, and C^*-convex sets*. Dissertation, University of Toronto, 1990.

8. S. Hildebrandt, *Über den numerischen Wertebereich eines Operators*, Math. Ann. **163**(1966), 230–247.

9. C.-K. Li, *The C-convex matrices*, Linear and Multilinear Algebra **21**(1987), 249–257.

10. C.-K. Li and N.-K. Tsing, *On the k^{th} matrix numerical range*, Linear and Multilinear Algebra, **28**(1991), 229–239.

11. R. I. Loebl and V. I. Paulsen, *Some remarks on C^*-convexity*, Linear Algebra Appl. **35**(1981), 63-78.

12. F. J. Narcowich and J. D. Ward, *Support functions for matrix ranges: analogues of Lumer's formula*, J. Operator Theory **7**(1982), 25–49.

13. V. I. Paulsen, *Completely bounded maps and dilations*. Pitman Research Notes 146, Longman Scientific and Technical, New York, 1986.

14. C. Pearcy and N. Salinas, *The reducing essential matricial spectra of an operator*, Duke Math. J. **42**(1975), 423–434.

15. F. M. Pollack, *Properties of the matrix range of an operator*, Indiana Univ. Math. J. **22**(1973), 419–427.

16. Y.-T. Poon, *The generalized k-numerical range*, Linear and Multilinear Algebra **9**(1980), 181–186.

17. N. Salinas, *Reducing essential eigenvalues*, Duke Math. J. **40**(1973), 561–580.

18. ———, *Hypoconvexity and essentially n-normal operators*, Trans. Amer. Math. Soc. **256**(1979), 325–357.

19. R. R. Smith and J. D. Ward, *Matrix ranges for Hilbert space operators*, Amer. J. Math. **102**(1980), 1031–1081.

20. J. G. Stampfli and J. P. Williams, *Growth conditions and the numerical range in a Banach algebra*, Tôhoku Math. J. **20**(1968), 417–424.

Centre de Recherches Mathématiques
Université de Montréal
Montréal, Québec H3C 3J7