Automatic Convexity

Charles A. Akemann

Department of Mathematics, University of California,
Santa Barbara, CA 93106, USA
akemann@math.ucsb.edu

Nik Weaver*

Department of Mathematics, Washington University,
St. Louis, MO 63130, USA
nweaver@math.wustl.edu

In many cases the convexity of the image of a linear map with range is $\mathcal{R}^n$ is automatic because of the facial structure of the domain of the map. We develop a four step procedure for proving this kind of "automatic convexity". To make this procedure more efficient, we prove two new theorems that identify the facial structure of the intersection of a convex set with a subspace in terms of the facial structure of the original set.

Let K be a convex set in a real linear space X and let H be a subspace of X that meets K. In Part 1 we show that the faces of $K \cap H$ have the form $F \cap H$ for a face F of K. Then we extend our intersection theorem to the case where X is a locally convex linear topological space, K and H are closed, and H has finite codimension in X. In Part 2 we use our procedure to "explain" the convexity of the numerical range (and some of its generalizations) of a complex matrix. In Part 3 we use the topological version of our intersection theorem to prove a version of Lyapunov's theorem with finitely many linear constraints. We also extend Samet's continuous lifting theorem to the same constrained siuation.

Historically there have been several theorems that concluded, unexpectedly, even mysteriously at first, that a certain set in $\mathcal{R}^n$ is convex. Perhaps the two best known examples are the convexity of the numerical range of an $n \times n$ complex matrix [9, 16] and Lyapunov's theorem on the convexity of the range of a vector measure [13]. In each of these cases the set in question is the image under some apparently non-linear map of a non-convex set. Each of these theorems has been generalized in many directions. Until the work of Lindenstrauss [12], Lyapunov's theorem remained a mystery with several complicated, yet incomplete, proofs (including Lyapunov's and a later proof by Halmos [7]) in the literature. See [1] for a discussion of Lyapunov's theorem and generalizations. As for the convexity of the numerical range, while the proofs in the literature have been complete, and they have gotten steadily simpler, the mystery of the appearance of convexity has remained (see [11, p. 78], [14] and [6, Sections 1.1 and 5.5]).

In [1] a number of automatic convexity theorems related to Lyapunov's Theorem were proved. The key to those theorems is given in [1, Theorem 1.6 and Corollary 1.7], which we restate here, correcting misprints, after introducing some notation.

Notation. K denotes a convex set in a real linear space X. For any distinct points $x, y \in X$ let (x, y) denote the line segment joining x and y, excluding the end points.

$E(K)$ denotes the set of extreme points of K. If K is not a singleton, the facial dimension [1, p. 10] of K is defined to be $inf\{dim(Q) : Q$ is a nonsingleton face of $K\}$. (Facial dimension ∞ is quite possible and especially interesting as we shall see in Part 3 of this paper.) For any subset F of K let $G(K, F)$ denote the smallest face of K containing F. In [1] this concept was defined and developed for singleton sets $F = \{v\}$, where the notation $G(K, v)$ was used.

[1, 1.6]. If K has facial dimension $> n$, Ψ is an affine map of K into $\mathcal{R}^n$, and $v \in K$, then every extreme point of $\Psi^{-1}(\Psi(v))$ is an extreme point of K.

[1, 1.7]. If X is a locally convex space, K is compact with facial dimension $> n$, and Ψ is a continuous affine map from K into $\mathcal{R}^n$, then $\Psi(E(K)) = \Psi(K)$.

The form of [1, 1.7] suggests the following procedure for proving that the image in $\mathcal{R}^n$ of certain kinds of maps are automatically convex. Let's assume that we have some set E and some function f that takes elements of E into $\mathcal{R}^n$. To prove that $f(E)$ is convex you can try the following procedure. We shall illustrate this in several cases in Parts 2 and 3 of this paper.

Automatic Convexity Procedure.

1. Find a suitable linear space X and linear map Ψ such that the elements E can be found in X (perhaps in a slightly different guise) and $f(e) = \Psi(e)$ for each $e \in E$.

2. Define $K = Conv(E)$ (or perhaps the closure of $Conv(E)$). Show that the extreme points of K lie in E.

3. Show that the facial dimension of K is more than the dimension of the range of Ψ, possibly using the intersection theorems in Part 1 below.

4. Apply [1, 1.7] to get the desired convexity.

A knowledge of the facial structure of K is crucial to any application of [1, 1.7]. In Part 1 we prove two new theorems that describe the facial structure of the intersection of a convex set with certain subspaces in terms of the facial structure of the original convex set. These theorems will allow new applications of the automatic convexity procedure. In Part 2 of the present paper we discuss numerical range as an application of pure convexity theory in a way that (we believe) unravels the mystery and paves the way for more theorems having convexity as their conclusions. In Part 3 we further extend Lyapunov's convexity theorem and even the continuous lifting theorem of Samet [15]; again our methods open the way for many more results of the same type.

1.   The Intersection Theorems

Algebraic Intersection Theorem. *Given a subspace H in X and a point $x \in X$, let F be a face of $(H + x) \cap K$. Then $G(K, F) \cap H = F$.*

Proof. WLOG we can assume that $x = 0$. From [1, 1.1 and 1.2], $G(K, v)$ consists of all elements y of K such that there exists $\lambda > 0$ such that $(1 + \lambda)v - \lambda y \in K$. Let

$G = \bigcup \{G(K, v) : v \in F\}$. Claim $G(K, F) = G$. The inclusion $G \subset G(K, F)$ is clear from the face property, so we need only show that G is a face of K and that $F \subset G$.

If $x, y \in G$, then there exist $v, w \in F$ such that $x \in G(K, v)$ and $y \in G(K, w)$. We can assume a single λ such that $(1+\lambda)v - \lambda x \in K$ and $(1+\lambda)w - \lambda y \in K$. For any $\alpha \in (0, 1)$,

$$\alpha((1 + \lambda)v - \lambda x) + (1 - \alpha)((1 + \lambda)w - \lambda y) \in K.$$

Grouping the $(1 + \lambda)$ terms and the λ terms, we get

$$(1 + \lambda)(\alpha v + (1 - \alpha)w) - \lambda(\alpha x + (1 - \alpha)y) \in K.$$

Thus $(\alpha x + (1 - \alpha)y) \in G$. This shows that G is convex.

To show that G is a face of K, assume $x, y \in K$ such that $.5(x+y) \in G$. Then there exists $v \in F$ such that $.5(x + y) \in G(K, v)$. But $G(K, v)$ is a face of K, so $x, y \in G(K, v) \subset G$. Thus G is a face of K.

Finally we show that $G \cap H = F$. The inclusion $F \subset G \cap H$ is clear from the definition of G. If $y \in G \cap H$, then there is a $v \in F$ such that $y \in G(K, v) \cap H$. Thus there exists $\lambda > 0$ such that $(1 + \lambda)v - \lambda y \in K$. But $v \in F \subset H$ and $y \in H$, so $(1 + \lambda)v - \lambda y \in H$ since H is a subspace. Thus $(1 + \lambda)v - \lambda y \in H \cap K$. Since $v \in F, v = (1/(1 + \lambda))((1 + \lambda)v - \lambda y)) + (\lambda/(1 + \lambda))y$ and F is a face of $H \cap K$, we conclude that $y \in F$. $\square$

Comment. If F has a weak internal point v (in the sense of [1, p. 8]), then $G(K, F) = G(K, v)$. However, many interesting infinite dimensional convex sets do not have weak internal points, e.g. the state space of $C([0, 1])$ or most any other interesting C*-algebra.

Now we prove a topological version of this result. As will be clear from a subsequent example, we need to consider a restricted class of subspaces H in the topological situation.

Topological Intersection Theorem. *Assume now that K is a convex, closed set in a locally convex space X. Given a closed subspace H of finite co-dimension in X and a point $y \in X$, let F be a closed face of $(H + y) \cap K$. Then $G(K, F)$ is closed and $G(K, F) \cap H = F$.*

Proof. By a simple induction argument, it suffices to prove the theorem under the asumption that H is a closed hyperplane, and WLOG we can assume that $y = 0$. Let $f : X \to \mathcal{R}$ be a continuous linear functional such that $H = f^{-1}(0)$. We need only prove that $G(K, F)$ is closed, as $G(K, F) \cap H = F$ follows from the Algebraic Intersection Theorem.

Suppose $\{x_t\}$ is a net in $G(K, F)$ such that $x_t \to x$; we must show $x \in G(K, F)$. Exchanging $-f$ for f if necessary and passing to a subnet, we can assume that $f(x_t) \geq 0$ for all t. If $f(x_t) = 0$ frequently, then we can pass to a subnet such that each $x_t \in G(K, F) \cap H = F$, and so $x \in F$ (and hence $x \in G(K, F)$) because F is closed.

Otherwise, pass to a subnet such that $f(x_t) > 0$ for all t. Let x_0 be any of the x_t and fix it. Since $f(x_0) > 0$, x_0 can't lie in F, so by [1, 1.1] $x_0 \in G(K, F)$ implies that there is a $y_0 \in G(K, F)$ such that the open line segment (x_0, y_0) intersects F. It follows by linearity of f that $f(y_0) < 0$. Now for each t, linearity of f implies that there is a unique point z_t in (x_t, y_0) such that $f(z_t) = 0$, i.e. $z_t \in H \cap (x_t, y_0)$. Explicitly, $z_t = r_t x_t + (1 - r_t)y_0$ where $r_t = -f(y_0)/(f(x_t) - f(y_0)) \in (0, 1)$ since $f(y_0) < 0$. Since x_t and y_0 are both in $G(K, F)$,

it follows from convexity of $G(K, F)$ that $z_t \in G(K, F)$. Hence $z_t \in G(K, F) \cap H = F$ by the Algebraic Intersection Theorem. Now $r = \lim r_t = -f(y_0)/(f(x) - f(y_0)) > 0$ because $x_t \to x$. Thus (z_t) converges; let $z = \lim z_t$, so $z \in F$. Then $z = rx + (1-r)y_0$, so if $r = 1$ then $x = z \in F$. If $r < 1$, then the line segment (x, y_0) contains z, which implies $x \in G(K, z) \subset G(K, F)$. Thus $G(K, F)$ is closed. $\qquad \square$

Example. In this example we show why it is necessary to restrict H to a subspace of finite co-dimension in the Topological Intersection Theorem.

We work in the Banach space $c_0(\mathbf{Z})$. Let h be the sequence with nth term

$$h_n = \begin{cases} 1/n & \text{if } n \geq 1 \\ 0 & \text{if } n \leq 0. \end{cases}$$

For each $n \in \mathbf{Z}$ let e^n be the sequence which is 1 at n and 0 elsewhere. Let K_1 be the closed convex hull of the vectors $k^n = \frac{1}{n}e^{-n} + h$ (for $n \geq 1$); let K_2 be the closed convex hull of the vectors $-\frac{1}{n}k^n$ (for $n \geq 1$); and let F be the set of sequences $a = (a_n)$ such that $a_n = 0$ for $n \leq 0$ and $0 \leq a_n \leq n^{-2}$ for $n \geq 1$.

K_1 and K_2 are each the closed convex hull of a convergent sequence of vectors in a Banach space, and hence are compact. F is compact because it is closed and totally bounded. Thus the convex hull K of K_1, K_2, and F is compact. (It is a continuous image of the compact set $K_1 \times K_2 \times F \times S$ where $S = \{(r, s, t) : r, s, t \geq 0 \text{ and } r + s + t = 1\}$.)

Explicitly, K_1 is the set of sequences (a_n) such that $a_0 = 0, a_n = h_n$ for $n \geq 1$, $a_n \geq 0$ for $n \leq 0$, and $\sum_{n=0}^{\infty} na_{-n} \leq 1$. K_2 is the set of sequences b_n such that $b_0 = 0, b_n \leq 0$ for $n \leq 0$, $\sum_{n=0}^{\infty} n^2 b_{-n} \geq -1$, and $b_n = \alpha h_n$ for $n \geq 1$ where $\alpha = \sum_{n=0}^{\infty} nb_{-n}$.

Observe that if $ra + sb$ is in the convex hull of K_1 and K_2 ($a \in K_1$, $b \in K_2$, $r + s = 1$) and $ra_n + sb_n = 0$ for all $n \leq 0$, then we must have

$$s \cdot \sum_{n=0}^{\infty} nb_{-n} = -r \cdot \sum_{n=0}^{\infty} na_{-n} \geq -r.$$

Thus $ra + sb = rh + s\alpha h$ where $\alpha \geq -r/s$, and thus $ra + sb = \beta h$ with $0 \leq \beta \leq 1$.

Now let H be the set of sequences (a_n) such that $a_n = 0$ for all $n \leq 0$. This is a closed subspace of $c_0(\mathbf{Z})$. K_1 intersects H in the point h and F is contained in H, so $K \cap H$ contains the convex hull C of h and F. Moreover, any element of K — that is, any convex combination $ra + sb + tc$ with $a \in K_1$, $b \in K_2$, and $c \in F$ — which lies in H must satisfy $ra + sb \in H$; then by the last paragraph, $ra + sb = \beta h$ where $0 \leq \beta \leq r + s$, so we have

$$ra + sb + tc = \beta h + tc = (1 - t)\frac{\beta}{1 - t}h + tc$$

where $\beta/(1 - t) = \beta/(r + s) \leq 1$. Since C contains h and 0, it contains $(\beta/(1 - t))h$, and therefore it contains $ra + sb + tc$. We have shown that $C = K \cap H$.

Next we claim F is a closed face of C. It is closed because it is compact. It is a face because if $a, b \in C$ and neither belongs to F then $\lim a_n/h_n$ and $\lim b_n/h_n$ both exist and are strictly positive, so the same is true of $(a + b)/2$, which implies $(a + b)/2 \notin F$. This proves the claim.

Finally, we claim that any closed face G of K that contains F must contain h. For $0 \in F$, and 0 lies in the line segment joining k^n and $-\frac{1}{n}k^n$, which both belong to K, so $k^n \in G$. Since G is closed and $k^n \to h$, it follows that $h \in G$. This proves the final claim and shows that $G \cap H \neq F$. $\qquad\square$

2. Applications to Numerical Range

Notation. Let M_n denote the set of $n \times n$ complex matrices and U_n the set of unitary matrices in M_n. Let τ denote the trace on M_n and 1 the identity matrix. For $a, b \in M_n$ write $a \geq b$ if $a - b$ is positive semi-definite. Define $K = \{a \in M_n : 0 \leq a \leq 1\}$. When we need to specify a norm on M_n we shall always take the operator norm, i.e. $\|a\| = sup\{\|a\eta\| : \eta$ is a unit vector in $\mathcal{C}^n\}$. The k-numerical range of an $n \times n$ matrix b is $W_k(b) = \{(1/k) \sum_1^k (bx_j, x_j) :$ the x_j are orthonormal$\}$. When no confusion can develop we identify the complex plane with $\mathcal{R}^2$.

Let's illustrate our four step method by proving the convexity of the k-numerical range of $b \in M_n$. This was first shown by Berger [5]. A more accessible proof based on the convexity of the ordinary numerical range can be found in [8, Problem 167]. The first step is to linearize the function that produces the points in the k-numerical range. The definition of $W_k(b)$ calls for calculating a complex number for each k-tuple of orthonormal vectors in $\mathcal{C}^n$. Replace such a k-tuple $\{x_j\}_1^k$ with the orthogonal projection p of their span. Then $\tau(pb) = \sum_1^k (bx_j, x_j)$, thus we can see that $kW_k(b) = \{\tau(bq) : q$ is a projection of rank $k\}$, so it suffices to show that the latter set is convex. Setting $E = \{q \in M_n : q$ is a projection of rank $k\}$ completes the first step.

For the second step we define $Q_k = Conv(E)$. Clearly $Q_k \subset K$. Since $\tau(q) = k$ for each $q \in E$, this suggests that we consider Q_k as a subset of $\{a \in K : \tau(a) = k\}$. In Proposition 2.1 below we show that $E(Q_k) = E$.

For the third step we need to determine the facial structure of Q_k. It seems sensible to start with K. This is probably classical, but a readable (and more general) account appears as [3, 2.2] where faces of K are shown to have the form $F = \{x \in K : p \leq x \leq q\}$, where q, p are self-adjoint projections in M_n. This can be rewritten in terms of the difference $q - p = r$ as $F = p + rKr$. A face of this form is an extreme point exactly when $r = 0$, and then the extreme point is just the projection p, i.e. the extreme points of K are exactly the projections. Since the analysis of the facial structure of Q_k uses the intersection theorem from Part 1, we state the facts as a proposition.

Proposition 2.1. *For $1 \leq k < n$, $Q_k = \{a \in K : \tau(a) = k\}$. The facial dimension of Q_k is 3. Further, the extreme points of Q_k are exactly the projections of rank k.*

Proof. We already noted that $Q_k \subset \{a \in K : \tau(a) = k\}$. If we show that the right hand side has exactly the projections of rank k as its extreme point set, then equality will follow.

Note that if we intersect K with the hyperplane $H = \{x : \tau(x) = k\}$, then we get exactly $\{a \in K : \tau(a) = k\}$. Using the notation developed just above the statement of the proposition, let a face F of K have the form $F = p + rKr$. By the Algebraic Intersection Theorem the typical face of $\{a \in K : \tau(a) = k\}$ is $F \cap H$.

WLOG suppose $F \cap H$ is nonempty. If $\tau(p) = k$ then $F \cap H = \{p\}$ has dimension 0. The same is true if $\tau(q) = k$. Otherwise, $\tau(p) < k < \tau(q)$, which implies $\tau(r) \geq 2$. Then rKr has real dimension at least 4, hence so does the face F, and H contains an interior point of F, so in this case $F \cap H$ has dimension at least 3. Thus we have shown that the set of extreme points of $\{a \in K : \tau(a) = k\}$ is exactly the set of projections of rank k, thereby completing the proof of $Q_k = \{a \in K : \tau(a) = k\}$. We also have shown that Q_k has no faces of dimension 1 or 2, hence its facial dimension is at least 3. Faces of dimension exactly 3 occur when $\tau(r) = 2$. □

We complete step 4 with the following proposition.

Proposition 2.2. *If $b \in M_n$, then the k-numerical range of b is convex.*

Proof. The linear map $\Psi(a) = \tau(ab)$ takes Q_k into $\mathcal{C}$ and its range is exactly the k-numerical range of b. Since the facial dimension of Q_k is 3 and the extreme points are projections of rank k, [1, 1.7] gives the desired convexity. □

As another example of this method, we prove the convexity of the c-numerical range for a self-adjoint element c of M_n. For any $c \in M_n$ the c-numerical range of a matrix $a \in M_n$ is defined to be $W_c(a) = \{\tau(cuau^*), u \in U_n\}$. It is easy to check that the k-numerical range $W(a)$ is obtained from this definition when c is taken to be a self-adjoint projection matrix of rank k (for the ordinary numerical range simply take $k = 1$). It is known that the c-numerical range is convex when c is self-adjoint [6, sect. 5.5]. In the next proposition we show "why" this is true.

Proposition 2.3. *If $c \in M_n$ is self adjoint, then the c-numerical range of b is convex for all $b \in M_n$.*

Proof. In this formulation the first step of the automatic convexity procedure is straightforward. Fix an element $b \in M_n$. Define $E = \{u^*cu : u \in U_n\}$. Let $M = Conv(E)$. Note that M is compact since E is compact [17, 2.2.6]. The set of extreme points of M is exactly the set E since M contains extreme points [17, 2.6.16], these lie in E [17, 2.6.4], and any point of E can be mapped onto any other by a linear isometry of M onto itself (namely $u^*cu \rightarrow v^*u(u^*cu)u^*v = v^*cv$). For any $a \in M_n$ define $\Psi(a) = \tau(ab)$. Then $\Psi(E)$ is exactly the c-numerical range of b. To complete the proof using [1, 1.7], we need only show that the facial dimension of M is at least 3. This is done in the following lemma. □

Since we have to borrow from matrix theory for the proof of the next lemma, for comparison and convenience we use the notation of [6, Section 5.5]. Because of the change to the notation of [6, Section 5.5] what we called c in the previous proposition is now C, while c stands for the real vector consisting of the eigenvalues of C.

Lemma 2.4. *Fix a self-adjoint matrix $C \in M_n$. Then the facial dimension of $M = Conv(\{U^*CU : U \in U_n\})$ is at least 3.*

Proof. Let $\alpha, \beta \in \mathcal{R}^k$. We say that α is *obtained from β by pinching* if all components of α and β agree except for two, α_i and α_j, which satisfy $\alpha_i = \lambda\beta_i + (1 - \lambda)\beta_j$ and $\alpha_j = (1 - \lambda)\beta_i + \lambda\beta_j$ for some $\lambda \in [0, 1]$. We require the following fact: the positive vector

α is obtained from the positive vector β by a finite number of pinchings if and only if

$$\sum_{i=1}^{k} \alpha_i \le \sum_{i=1}^{k} \beta_i$$

for $1 \le k \le n$, with equality when $k = n$. Write $\alpha \prec \beta$ for this relation.

Since adding a scalar multiple of the identity matrix to C only shifts the C-numerical range, WLOG we can let C be the positive diagonal matrix with diagonal c, denoted $C = [c]$, where c is arranged in decreasing order. Let $M' = \{U^*[b]U : U \text{ is unitary and } b \prec c\}$. We shall show that $M' = M$. Note that M' is the set of positive matrices B whose ordered eigenvalue list b satisfies $b \prec c$. Observe that the sum of the first k eigenvalues of B equals $\sup\{\tau(BP) : P \text{ is a rank } k \text{ projection}\}$ [2, Lemma 1.3]. Thus, M' is the set of positive matrices B such that $\tau(B) = \tau(C)$ and

$$\tau(BP) \le \sum_{i=1}^{k} c_i$$

for $1 \le k \le n$ and every rank k projection P. It easily follows that M' is closed and convex.

Next, we claim that the extreme points of M' are precisely the matrices of the form $U^*[c]U$ for U a unitary matrix. To see this, let $B = U^*[b]U \in M'$ and suppose B is not of the form $U^*[c]U$. Then $[b]$ is obtained from $[c]$ by a finite, nonempty sequence of pinchings. It follows that $[b]$ is obtained from some $[a]$ by a single pinching, where $a \prec c$. That is, $b_i = ta_i + (1-t)a_j$ and $b_j = (1-t)a_i + ta_j$ for some $t \in (0,1)$, where $a_i \ne a_j$, and all other components of a and b agree. Let a' be the real vector obtained from a by switching the i and j components. Then $A = U^*[a]U$ and $A' = U^*[a']U$ are both in M', and $B = tA + (1-t)A'$. So B is not an extreme point. Thus, every extreme point of M' must be of the form $U^*[c]U$. Thus $M' = M$ by [17, 2.6.16].

Finally, we claim that the facial dimension of M is at least 3. To see this, let $B = U^*[b]U \in M$ and suppose B is not an extreme point. Define A and A' as in the last paragraph. Then, with respect to a basis which diagonalizes B, we have

$$A_{[ij]} = \begin{bmatrix} a_i & 0 \\ 0 & a_j \end{bmatrix} \quad \text{and} \quad A'_{[ij]} = \begin{bmatrix} a_j & 0 \\ 0 & a_i \end{bmatrix},$$

where we use the subscript $[ij]$ to indicate restriction to the (i,i), (i,j), (j,i), and (j,j) entries. (Recall that A and A' agree elsewhere.) Define new matrices A_1, A_1', A_2, and A_2' by setting

$$(A_1)_{[ij]} = \begin{bmatrix} ta_i + (1-t)a_j & (t-t^2)^{1/2}(a_i - a_j) \\ (t-t^2)^{1/2}(a_i - a_j) & (1-t)a_i + ta_j \end{bmatrix}$$

$$(A_1')_{[ij]} = \begin{bmatrix} ta_i + (1-t)a_j & -(t-t^2)^{1/2}(a_i - a_j) \\ -(t-t^2)^{1/2}(a_i - a_j) & (1-t)a_i + ta_j \end{bmatrix}$$

$$(A_2)_{[ij]} = \begin{bmatrix} ta_i + (1-t)a_j & i(t-t^2)^{1/2}(a_i - a_j) \\ -i(t-t^2)^{1/2}(a_i - a_j) & (1-t)a_i + ta_j \end{bmatrix}$$

$$(A_2')_{[ij]} = \begin{bmatrix} ta_i + (1-t)a_j & -i(t-t^2)^{1/2}(a_i - a_j) \\ i(t-t^2)^{1/2}(a_i - a_j) & (1-t)a_i + ta_j \end{bmatrix}$$

and letting them agree with A and A' elsewhere. It is clear that each of these matrices is self-adjoint, and as the 2×2 parts all have the same trace and determinant, they all have the same eigenvalues (namely, a_i and a_j). Thus they all belong to M. But $B = (A_1 + A_1')/2 = (A_2 + A_2')/2$, and the affine space spanned by A, A', A_1, A_1', A_2, and A_2' is three-dimensional, so the smallest face containing B has dimension at least 3. This proves the final claim. $\square$

3.   Applications to Lyapunov Type Theorems

Let $(X, \mathcal{M})$ be a measurable space. A vector measure is an n-tuple $(\mu_1, \ldots, \mu_n) = \mu$ of real-valued measures on $(X, \mathcal{M})$. Lyapunov's Theorem [13] states that, if μ is nonatomic, then the range of μ is a convex, compact set in $\mathcal{R}^n$. Following the 4 step plan for proving convexity (and often compactness in the same stroke, as is the case here) one observes that

$$\int \chi_A d\mu = \mu(A) = (\int \chi_A d\mu_1, \ldots, \int \chi_A d\mu_n) = (\int \chi_A f_1 d\nu, \ldots, \int \chi_A f_n d\nu),$$

where χ_A is the characteristic function of the set A, $\nu = \sum_1^n |\mu_i|$ is a finite, positive measure, and f_i is the Radon-Nikodym derivative of μ_i with respect to ν for each i. This formulation suggested the definition of the map $\Psi : L^\infty(X, \mathcal{M}, \nu) \to \mathcal{R}^n$ by $\Psi(g) = (\int g f_1 d\nu, \ldots, \int g f_n d\nu)$. Moving to step 2 in the plan, we note that if E is viewed as the set of characteristic functions in $L^\infty(X, \mathcal{M}, \nu)$, then the closed convex hull K of E in the weak* topology is exactly the set of positive functions of norm no more than 1. The facial dimension of K is shown in [3] to be ∞, so [1, 1.7], the weak* compactness of K and the weak* continuity of Ψ complete the proof of Lyapunov's Theorem.

As with the numerical range situation discussed earlier in this paper, once the problem was put into the correct notation, the convexity was automatic from facial structure considerations and [1, 1.7]. Of course [1] contained many results that could be viewed as generalizations of Lyapunov's Theorem. Now let's combine these results with the Topological Intersection Theorem to show how even more theorems of the Lyapunov type are true using our 4 step method. In the next Theorem we extend [1, 2.5], which is itself an extension of Lyapunov's theorem to a non-commutative situation.

Theorem 3.1. *Suppose that N is a non-atomic von Neuman algebra and $\{f_1, \ldots, f_n\}$ and $\{g_1, \ldots, g_k\}$ are self-adjoint, normal linear functionals on N. Let $z_1, \ldots, z_n \in \mathcal{R}$ and define*

$$K = \{a \in N : \|a\| \leq 1, a \geq 0, f_j(a) = z_j, j = 1, \ldots, n\}.$$

Let N_{sa} denote the set of self adjoint elements of N. Define $\Psi : N_{sa} \to \mathcal{R}^k$ by $\Psi(a) = (g_1(a), \ldots, g_k(a))$. Then $E(K) = \{p : p$ is a projection in $K\}$ and $\Psi(K) = \Psi(E(K))$.

If N is abelian, then there is a continuous map $\Phi : \Psi(K) \to E(K)$ that is a right inverse for Ψ.

Proof. If K is void, the theorem is trivially true, so asume not. If N_1^+ denotes the positive part of the unit ball of N, then the facial dimension of N_1^+ is ∞ by [3, 2.2]. Since K is the intersection of N_1^+ with a subspace of finite codimension, the Topological Intersection Theorem applies to show that the faces of K are either extreme points of N_1^+ or else infinite dimensional faces. Since the extreme points of N_1^+ are exactly the

projections of N by [3, 2.2] , we get $E(K) = \{p : p$ is a projection in $K\}$. The conclusion $\Psi(K) = \Psi(E(K))$ follows from [1, 1.7].

Now assume that N is abelian and let N_{sa} be its self-adjoint part. Define $\Psi' : N_{sa} \to \mathcal{R}^{n+k}$ by the formula

$$\Psi'(a) = (\Psi(a), f_1(a), \ldots, f_n(a)).$$

By Lyapunov's Theorem $\Psi'(N_1^+)$ is compact and convex. By [15] there is a continuous right inverse $\Phi' : \Psi(K) \to E(K)$ for Ψ'. Now let

$$S = \{(\Psi(a), f_1(a), \ldots, f_n(a)) \in \Psi'(N_1^+) : f_j(a) = z_j, j = 1, 2, \ldots, n\}.$$

Clearly $\Psi'^{-1}(S) = \{a \in N_1^+ : f_j(a) = z_j, j = 1, 2, \ldots, n\} = K$. Thus the restriction of Φ' to K is the desired lifting if we identify the first k coordinates of $\mathcal{R}^{k+n}$ with $\mathcal{R}^k$. $\square$

We present two corollaries of Theorem 3.1. The first is a version of Lyapunov's theorem with linear constraints, and could possibly have applications in control theory along the lines of the classical Lyapunov theorem [10]. The second gives a von Neumann algebra version of the convexity of the k-numerical range (where here $k = z$).

Corollary 3.2. *Let $(X, \mathcal{M})$ be a measurable space, let $\mu = (\mu_1, \ldots, \mu_k)$ be a vector measure on $(X, \mathcal{M})$, let $\nu_1, \ldots, \nu_n$ be measures which are absolutely continuous with respect to $\nu = |\mu_1| + \cdots + |\mu_k|$, and let $z_1, \ldots, z_n \in \mathcal{R}$. Then the set $\{\mu(A) : \nu_j(A) = z_j, j = 1, 2, \ldots, n\}$ is compact and convex.*

Proof. We translate into the language of Theorem 3.1 by letting $N = L^\infty(X, \nu)$ and letting the f_i and g_j be the Radon-Nikodym derivatives of the μ_i and ν_j with respect to ν. Then $\{\mu(A) : \nu_j(A) = z_j, j = 1, 2, \ldots, n\} = \Psi(E(K)) = \Psi(K)$, and $\Psi(K)$ is clearly compact and convex. $\square$

Corollary 3.3. *Let N be a non-atomic von Neumann algebra with normal tracial state τ and let $b \in N$ and $z \in [0, 1]$. Then the set*

$$\{\tau(pb) : p \text{ is a projection and } \tau(p) = z\}$$

is a compact and convex subset of $\mathcal{C}$.

Proof. In Theorem 3.1, take $n = k = 1$, $f_1 = \tau$, $g_1 = \tau(\cdot b)$, and $z_1 = z$. $\square$

References

[1] C. Akemann, J. Anderson: Lyapunov Theorems for Operator Algebras, Memoirs of the Amer. Math. Soc. No. 458, v. 94, Providence, November (1991).

[2] C. Akemann, J. Anderson, N. Weaver: A geometric spectral theory for n-tuples of self-adjoint operators in finite von Neumann algebras, J. Funct. Anal. 165 (1999) 258–292.

[3] C. Akemann, G. K. Pedersen: Facial structure in operator algebra theory, Proc. London Math. Soc. 64(3) (1992) 418–448.

[4] C. Akemann, G. Shell, N. Weaver: Locally nonconical convexity, J. Convex Analysis 8 (2001) 1–21.

[5] C. Berger: Normal Dilations, Ph. D. Dissertation, Cornell (1963).

[6]   K. Gustafson, D. Rao: Numerical Range: The Field of Values of Operators and Matrices, Springer, New York (1997).

[7]   P. Halmos: On the set of values of a finite measure, Bull. Amer. Math. Soc. 53 (1947) 138–141.

[8]   P. Halmos: A Hilbert Space Problem Book, Van Nostrand, New York (1967).

[9]   F. Hausdorff: Das algebraische Analogon zu einem Satz von Fejér, Math. Z. 2 (1918) 187–197.

[10] H. Hermes, J. P. LaSalle: Functional Analysis and Time Optimal Control, Academic Press, New York (1969).

[11] R. Horn, C. Johnson: Topics in Matrix Analysis, Cambridge University Press, Cambridge (1991).

[12] J. Lindenstrauss: A short proof of Lyapunov's convexity theorem, J. Math. Mech. 15 (1966) 971–972.

[13] A. Lyapunov: On completely additive vector functions, Bull. Akad. Sci. USSR 4 (1940) 465–478 (Russian).

[14] Y. Poon: Generalized numerical ranges, joint positive definiteness and multiple eigenvalues, Proc. Amer. Math. Soc. 125(6) (1997) 1625–1634.

[15] D. Samet: Continuous selections for vector measures, Math. Oper. Res. 12(3) (1987) 536–543.

[16] O. Toeplitz: Der Wertevorrat einer Bilinearform, Math. Z. 3 (1919) 314–316.

[17] R. Webster: Convexity, Oxford University Press, Oxford (1994).