STUDIA MATHEMATICA, T. LXI. (1977)

On the geometric properties of the joint spectrum of a family of self-adjoint operators

by

YU. SH. ABRAMOV (Leningrad)

Abstract. In this note we discuss certain geometric properties of the joint spectrum of a family od self-adjoint operators. It is shown that the conical points of the joint numerical range belong to the joint spectrum. Moreover, variational characterizations of eigenvalues are given in the case of a commuting family.

1. Introduction. Let $A = (A_1, \ldots, A_k)$ be a family of bounded linear operators in a Hilbert space H with the scalar product $(\cdot,\cdot)$ and the norm $\|\cdot\|$. We denote by $p_i(x)$ the Rayleigh functional for the operator A_i, $i = 1, \ldots, k$, i.e.,

$$p_i(x) = \frac{(A_i x, x)}{(x, x)}, \quad x \neq 0.$$

In recent years quite a large number of works has been devoted to the study of the following subset of the space C^k:

$$W = \{(p_1(x), \ldots, p_k(x)): x \neq 0\};$$

this set is an analogue of the numerical range of an operator (the case of $k = 1$). Preserving this terminology, we shall call W also the *numerical range of the family* A. As it has turned out, the properties of the numerical range are, in general, no more valid after passing to higher dimension $(k > 1)$; a positive fact is the convexity of W in the case of a commuting family of normal operators [1]. As far as is known to the author, no more essential properties of the numerical range of a family of operators have occurred in the literature.

The notion of the spectrum has also been extended to the case of a commuting family of operators, the resulting concept being called the *joint spectrum.* And, as it usually happens with generalizations, there exist various non-equivalent extensions. A detailed classification of joint spectra can be found in the paper by Z. Słodkowski and W. Żelazko [6], together with the bibliography of the subject. Here we distinguish the joint spectrum σ defined as the set of all points $\lambda = (\lambda^{(1)}, \ldots, \lambda^{(k)})$ such that the ideal generated by the operators $\{\lambda^{(i)} I - A_i\}$, $1 \leqslant i \leqslant k$, in the bicommutant of A, is proper. A. T. Dash [1] has shown that, in the case of a family of normal operators, the closure of the numerical range W coincides with the convex hull of the spectrum σ:

(1) $$\sigma \subset \overline{W} = \operatorname{conv} \sigma.$$

This fact will be useful to us in the sequel.

We shall focus our attention on the case where the family A consists of self-adjoint operators. Let us remark that already for $k = 3$ (without commutativity assumed) the set W need not be convex (see [4]). In the self-adjoint case the set W can be viewed in a natural way as a subset of the real Euclidean space $\boldsymbol{R}^k$.

A k-tuple $\lambda = (\lambda^{(1)}, \ldots, \lambda^{(k)})$ will be called an *eigenvalue of* A if there exists a non-zero eigenvector x corresponding to λ, i.e., such that

$$A_i x = \lambda^{(i)} x, \quad i = 1, \ldots, k.$$

If we introduce the *Rayleigh functional* (more precisely, the *Rayleigh operator*) for the family A by putting

$$p(x) = \big(p_1(x), \ldots, p_k(x)\big),$$

then we have $p(x) = \lambda$, where λ is an eigenvalue and x is the corresponding eigenvector. In particular, the eigenvalues belong to the range W.

In the case of a single operator it is well known that boundary points of the numerical range of a (self-adjoint) operator belong to its spectrum; consequently, in the finite-dimensional case, they are in fact eigenvalues. In Section 2 below we generalize this fact to families of operators; the role of boundary points will now be played by so-called *conical points*.

There exist well-known variational characterizations of the eigenvalues of self-adjoint operators; they are connected with the names of Rayleigh, Courant–Weil, Poincaré–Ritz. The natural ordering of the real line makes it possible to write those eigenvalues in the form of successive maxima and minima of the corresponding quadratic forms. Pushing the analogy further on, one is naturally led to the problem of extremal properties of the eigenvalues of the family A. The main difficulty, however, lies in the absence of a linear order in $\boldsymbol{R}^k$; we are not going to attack this difficulty now. We just assume that some partial ordering is induced by a cone (not every cone is equally good) and the variational problem will be formulated with respect to that ordering. In Section 4 we obtain certain principles of min-max type, providing means for the description of the eigenvalues of A not involving the eigenvectors. We do not want to restrict our considerations to compact operators; therefore we introduce the concept of the *approximative point spectrum* π for the family A. As in the one-dimensional case, the set $\sigma \setminus \pi$ consists of isolated eigenvalues of finite multiplicity, which we characterize in terms of variational principles.

2. Conical points of the numerical range. In this and in the next section we consider a family $A = (A_1, \ldots, A_k)$ consisting of bounded self-adjoint operators, which are not assumed to commute.

For $\lambda = (\lambda^{(1)}, \ldots, \lambda^{(k)})$, $i = 1, \ldots, k$, we put

$$L_i(\lambda) = \lambda^{(i)} I - A_i.$$

The *spectrum of* A is defined as the set of all points $\lambda \in C^k$ for which there exists a sequence $\{x_n\}$ with $\|x_n\| = 1$ and such that

$$L_i(\lambda) x_n \to 0, \quad i = 1, \ldots, k.$$

If, additionally, the suquence $\{x_n\}$ converges weakly to zero (in symbols, $x_n \rightharpoonup 0$), we say that λ *belongs to the approximative point spectrum of* A.

The spectrum and the approximative point spectrum will be denoted by σ and π, respectively. Clearly, $\pi \subset \sigma$, and both these sets are compact subsets of $\boldsymbol{R}^k$. If $\lambda \in \sigma$, then for some sequence $\{x_n\}$ with $\|x_n\| = 1$ we have $(A_i x_n, x_n) \to \lambda^{(i)}$ for all i. Thus

$$\sigma \subset \overline{W},$$

where $\overline{W}$ denotes the closure of W.

Observe that, in the non-commutative case, we have taken as the spectrum what is usually called the *approximative spectrum*. This seems reasonable, since the operators in question are self-adjoint. In the commutative case σ coincides with the spectrum investigated in [1], [2], [4] and [6].

We now introduce the operator $\mathscr{D}_a(\lambda)$ depending on two parameters $a, \lambda \in \boldsymbol{R}^k$:

$$\mathscr{D}_a(\lambda) = \sum_{i=1}^{k} (\lambda^{(i)} - a^{(i)}) L_i(\lambda).$$

THEOREM 1. *Let $\gamma \in \overline{W}$ and assume that W lies inside m spheres with centres $a_1, \ldots, a_m$ and that γ belongs to all these spheres. Then the point* 0 *is in the joint spectrum of the family*

$$\mathscr{D} = \big(\mathscr{D}_{a_1}(\gamma), \ldots, \mathscr{D}_{a_m}(\gamma)\big).$$

If, moreover, $\gamma = p(x)$, then 0 *is an eigenvalue of $\mathscr{D}$ and x is a corresponding eigenvector.*

Proof. For $x \in H$, $\|x\| = 1$, we have

$$\big(\mathscr{D}_a(\gamma) x, x\big) = \sum_{i=1}^{k} (\gamma^{(i)} - a^{(i)}) \big(\gamma^{(i)} - p_i(x)\big)$$
$$= \|\gamma - a\|_k^2 + \sum_{i=1}^{k} (\gamma^{(i)} - a^{(i)}) \big(a^{(i)} - p_i(x)\big),$$

where $\|\cdot\|_k$ denotes the norm in $\boldsymbol{R}^k$.

Let r_ν be the radius of the sphere with centre $a_\nu = (a_\nu^{(1)}, \ldots, a_\nu^{(k)})$, $\nu = 1, \ldots, m$. Then

$$(\mathscr{D}_{a_\nu}(\gamma)x, x) \geqslant r_\nu^2 - \Big|\sum_{i=1}^{k}(\gamma^{(i)} - a_\nu^{(i)})(a_\nu^{(i)} - p_i(x))\Big|$$
$$\geqslant r_\nu^2 - \|\gamma - a_\nu\|_k \|a - p(x)\|_k$$
$$= r_\nu\big(r_\nu - \|a_\nu - p(x)\|_k\big).$$

Since $p(x)$ lies inside the νth sphere, we have $\|a_\nu - p(x)\|_k \leqslant r_\nu$. It hence follows that the family $\mathscr{D}$ consists of positive operators. Since $\gamma \in \overline{W}$, there is a sequence $\{x_n\}$ with $\|x_n\| = 1$ and $p(x_n) \to \gamma$. Then

$$\big(\mathscr{D}_{a_\nu}(\gamma)x_n, x_n\big) = \sum_{i=1}^{k}(\gamma^{(i)} - a^{(i)})\big(\gamma^{(i)} - p_i(x_n)\big) \to 0.$$

The operators $\mathscr{D}_{a_\nu}(\gamma)$ being positive, we have

(2) $$\|\mathscr{D}_{a_\nu}(\gamma)x_n\|^2 \leqslant \|\mathscr{D}_{a_\nu}(\gamma)\|(\mathscr{D}_{a_\nu}(\gamma)x_n, x_n),$$

and this shows that 0 belongs to the joint spectrum of $\mathscr{D}$. If now $\gamma = p(x)$, then $\big(\mathscr{D}_{a_\nu}(\gamma)x, x\big) = 0$ and so, in view of (2), 0 is an eigenvalue of $\mathscr{D}$.

COROLLARY. *If $m = k$ and if the points $\gamma, a_1, \ldots, a_k$ are the vertices of a simplex in $\boldsymbol{R}^k$, then γ belongs to the spectrum σ of the family A.*

If, moreover, $\gamma = p(x)$, then γ is an eigenvalue of A and x is a corresponding eigenvector.

Proof. Since $\gamma, a_1, \ldots, a_k$ form a simplex, we have

$$\det(\gamma^{(j)} - a_i^{(j)}) = \det(\gamma^{(i)} - a_j^{(i)}) \neq 0.$$

Consequently the system of equations

$$\mathscr{D}_{a_\nu}(\gamma) = \sum_{i=1}^{k}(\gamma^{(i)} - a_\nu^{(i)})L_i(\gamma), \quad \nu = 1, \ldots, k,$$

can be solved with respect to the operators $L_i(\gamma)$. It remains to apply the fact that 0 is in the spectrum of $\mathscr{D}$. The second assertion follows from the last assertion of Theorem 1.

Let $F \subset \boldsymbol{R}^k$ and $\gamma \in F$. The point γ will be called a *conical point of F* whenever there exists a cone K (by a *cone* we mean a convex cone satisfying $K \cap -K = \{0\}$) such that the conical domain $\gamma - K$ contains F: $F \subset \gamma - K$.

Using the elementary facts of the geometry of $\boldsymbol{R}^k$, we derive from the last Corollary the following theorem.

THEOREM 2. *If γ is a conical point of $\overline{W}$, then γ belongs to the spectrum σ of the family A; in particular, if $\dim H < \infty$, then all conical points of W are eigenvalues of A.*

3. Three lemmas. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be any eigenvalues of A (not necessarily distinct) and let $x_1, x_2, \ldots, x_n$ be corresponding eigenvectors, linearly independent. We denote by $[x_1, \ldots, x_n]$ the *linear span of* $x_1, \ldots, x_n$.

LEMMA 1. *If $x \in [x_1, \ldots, x_n]$, $x \neq 0$, then*

$$p(x) \in \operatorname{conv}\{\lambda_1, \ldots, \lambda_n\}.$$

Proof. First observe that A maps the space $[x_1, \ldots, x_n]$ into itself and that the A_i's commute on that space

$$A_iA_jx = \sum_{\nu=1}^{n} c_\nu \lambda_\nu^{(i)}\lambda_\nu^{(j)}x_\nu.$$

Further, note that the spectrum of the restriction of A to $[x_1, \ldots, x_n]$ consists just of the eigenvalues $\lambda_1, \ldots, \lambda_n$. Now the lemma follows from (1).

LEMMA 2. *If E is a subspace of H of codimension $i \leqslant n-1$, then there is an $x \in E$, $x \neq 0$, such that*

$$p(x) \in \operatorname{conv}\{\lambda_1, \ldots, \lambda_n\}.$$

Proof. This follows from the fact that $E \cap [x_1, \ldots, x_n] \neq \{0\}$ and from the preceding lemma.

LEMMA 3. *If $\lambda \in \sigma \setminus \pi$, then λ is an isolated point of σ and an eigenvalue of A of finite multiplicity.*

If, moreover, the family A is commuting, then the converse is also true.

Proof. Let $\lambda \in \sigma \setminus \pi$. There exists a sequence $\{x_n\}$ on the unit sphere such that $L_i(\lambda)x_n \to 0$, $x_n \rightharpoonup x \neq 0$. Therefore $L_i(\lambda)x = 0$, $i = 1, \ldots, k$, and so λ is an eigenvalue. Its multiplicity is finite, otherwise the corresponding eigenspace would contain a sequence of unit vectors weakly convergent to zero.

Assume that λ is a cluster point of σ and let $\lambda_n \in \sigma$, $\lambda_n \to \lambda$. The spectrum π being closed, we may assume that $\lambda_n \in \sigma \setminus \pi$. As shown above, λ_n's are eigenvalues. Let x_n be corresponding eigenvectors, $\|x_n\| = 1$. We have

$$\|L_i(\lambda)x_n\| \leqslant |\lambda^{(i)} - \lambda_n^{(i)}|$$

and so $L_i(\lambda)x_n \to 0$, $i = 1, \ldots, k$. We may assume $x_n \rightharpoonup x$, so that $L_i(\lambda)x = 0$. Since

$$\big(L_i(\lambda)x, x_n\big) = \big(L_i(\lambda_n)x, x_n\big) = 0,$$

we see that $(\lambda^{(i)} - \lambda_n^{(i)})(x, x_n) = 0$ for $i = 1, \ldots, k$, whence $(x, x_n) = 0$. Letting $n \to \infty$ we obtain $x = 0$, contrary to the assumption that $\lambda \notin \pi$.

Now let A be a commuting family and let λ be an isolated point of σ which is an eigenvalue of A of finite multiplicity. Assume that $\lambda \in \pi$ and choose a sequence $\{x_n\}$ with $\|x_n\| = 1$, $x_n \rightharpoonup 0$ and $L_i(\lambda)x_n \to 0$, $i = 1, \ldots$ $\ldots, k$. Let M be a parallelepiped such that $M \cap \sigma = \{\lambda\}$, $M = M_1 \times \ldots$

$\ldots \times M_k$, where M_i's are one-dimensional intervals. Further, let $E_i(\cdot)$ and $E(\cdot)$ denote the spectral measures for the operator A_i and the family A, respectively. Then

$$E(M) = E_1(M_1) \ldots E_k(M_k).$$

As in the one-dimensional case (see [5]) it is not difficult to verify that $x_n - E_i(M_i)x_n \to 0$.

Now, we have the identity

$$I - E(M) = \sum_{i=1}^{k} B_i[I - E_i(M_i)],$$

where B_i are suitable bounded operators. It follows from this formula that $x_n - E(M)x_n \to 0$ and so

$$\lim \|E(M)x_n\| = 1.$$

Hence we infer that the projector $E(M)$ cannot have a finite rank, and this is a contradiction to the assumption that λ is an isolated eigenvalue of finite multiplicity and that M contains no other points of σ except λ.

4. Variational characterizations. In this section we assume that A is a commuting family.

Let K be a cone in $\mathbf{R}^k$ of the form

$$\{\lambda \in \mathbf{R}^k : f_i(\lambda) \geqslant 0,\ i = 1, \ldots, k\},$$

where f_i's are linearly independent functionals, so that K has exactly k edges and contains interior points (see [3]). Observe that, in the case $k = 2$, K is just an angle (with vertex at the origin) of angular measure $\varkappa$ with $0 < \varkappa < \pi$.

The cone K induces a partial order in $\mathbf{R}^k$. For $u, v \in \mathbf{R}^k$ we write $u \geqslant v$ if $u - v \in K$. If a set $F \subset \mathbf{R}^k$ is bounded above (relative to that ordering), then it has the least upper bound $\sup F$.

We introduce the following conditions

$$\sigma \subset (a - K) \cup (a + K), \tag{3}$$

$$\pi \cap (a + \operatorname{int} K) = \emptyset, \tag{4}$$

where a is a point in $\mathbf{R}^k$. According to Lemma 3 the portion of the spectrum contained within the conical domain $a + \operatorname{int} K$ is discrete, i. e. consists of isolated points, which are eigenvalues of finite multiplicity and which can cluster about boundary points only.

We shall say that the spectrum of A is *K-regular in* $a + \operatorname{int} K$ if the eigenvalues contained in $a + \operatorname{int} K$ can be arranged decreasingly (with respect to the order induced by K):

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n \geqslant \ldots, \tag{5}$$

each eigenvalue occurring as many times as its multiplicity indicates. Note that condition (5) implies that the sequence $\{\lambda_n\}$ can have just one limit point in the boundary of $a + \operatorname{int} K$.

From now on we assume that the spectrum σ of the family A satisfies conditions (3), (4) and (5). We are going to prove the analogues of the variational principles of Rayleigh, Courant–Weil and Poincaré–Ritz for the eigenvalues of A.

THEOREM 3. *There exists an orthonormal system of eigenvectors $x_1, \ldots, x_n, \ldots$ corresponding to the eigenvalues* (5) *and such that*

$$\sup_{x \perp x_1, \ldots, x_{n-1}} p(x) = \lambda_n, \tag{6}$$

the supremum being attained at the point x_n.

Proof. If $\lambda \in W$, then by (1) we have

$$\lambda = \sum_i a_i \mu_i + \sum_i \beta_i \lambda_i,$$

where

$$\mu_i \in \sigma \cap (a - K), \quad a_i, \beta_i \geqslant 0, \quad \sum_i (a_i + \beta_i) = 1.$$

Applying this representation and (5) we get $\lambda \leqslant \lambda_1$. If now x_1 is an eigenvector for λ_1, then $p(x_1) = \lambda_1$ and (6) is proved for $n = 1$.

We proceed by induction. Let $x_1, x_2, \ldots, x_n, \ldots$ be an orthonormal system of eigenvectors corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ Write

$$X_m = [x_1, \ldots, x_m], \quad E^m = H \ominus X_m.$$

Suppose that (6) is proved for $m = 1, \ldots, n-1$. We assume $\lambda_n \neq \lambda_{n-1}$ (the other case is trivial) and we restrict our attention to the subspace E^{n-1}. Since E^{n-1} is reductive for A, the restriction of A to E^{n-1} (denoted by $A \mid E^{n-1}$) has the numerical range $W(E^{n-1})$ such that

$$\overline{W(E^{n-1})} = \operatorname{conv} \sigma(A \mid E^{n-1})$$

and

$$\sigma(A \mid E^{n-1}) = \sigma \setminus \{\lambda_1, \ldots, \lambda_{n-1}\}.$$

Hence in view of (5) we obtain $p(x) \leqslant \lambda_n$ for $x \in E^{n-1}$. This proves (6) for $m = n$.

Let $\mathscr{E}_n$ ($\mathscr{E}^n$) denote the totality of all subspaces of H of dimension (codimension) n.

THEOREM 4. ***The following principle holds true:***

$$\inf_{E \in \mathscr{E}^{n-1}} \sup_{x \in E} p(x) = \lambda_n.$$

Proof. Let $E\in\mathscr{E}^{n-1}$; then by Lemma 2 we have $p(x)\in\operatorname{conv}\{\lambda_1,\ldots,\lambda_n\}$ for some $x\in E$, $x\neq 0$. Consequently $p(x)\geqslant\lambda_n$ and so

$$M(E)=\sup_{x\in E}p(x)\geqslant\lambda_n.$$

Hence it follows that λ_n is a lower bound for the set $\{M(E):\ E\in\mathscr{E}^{n-1}\}$. But, according to Theorem 3, we have $\lambda_n=M(E^{n-1})$ and thus λ_n is the greatest lower bound.

THEOREM 5. *The following principle holds true*:

$$\sup_{E\in\mathscr{E}_n}\inf_{x\in E}p(x)=\lambda_n.$$

Proof. Let $E\in\mathscr{E}_n$; then there is an element $y\neq 0$ in $E\cap E^{n-1}$ and, by Theorem 3, we have

$$m(E)=\inf_{x\in E}p(x)\leqslant p(y)\leqslant\lambda_n.$$

Hence it follows that λ_n is an upper bound for the set $\{m(E):\ E\in\mathscr{E}_n\}$. But, according to Lemma 1, we have $m(X_n)=\lambda_n$ and thus λ_n is the least upper bound.

Remarks. 1. If the family $A=(A_1,\ldots,A_k)$ consists of compact operators, then it is easy to see that the approximative point spectrum π contains no other points but the origin, i.e., $\pi=\{0\}$.

2. Assume, for simplicity, that H is finite dimensional and let $A=(A_1,A_2)$, where A_1 has a simple spectrum:

$$\lambda_1^{(1)}>\lambda_2^{(1)}>\ldots>\lambda_n^{(1)},$$

and A_2 has the norm satisfying the estimate

$$\|A_2\|\leqslant\tfrac{1}{2}\tan\tfrac{1}{2}\varkappa\min\,(\lambda_i^{(1)}-\lambda_{i+1}^{(1)})$$

with some $\varkappa$, $0<\varkappa<\pi$. Then the eigenvalues $\lambda_i=(\lambda_i^{(1)},\lambda_i^{(2)})$ of the family A can be arranged decreasingly in the first component; and this is precisely the arrangement induced by the cone $K=\{z:\ |\arg z|\leqslant\tfrac{1}{2}\varkappa\}$.

References

[1] A. T. Dash, *Joint numerical range*, Glasnik Mat. 7, 1 (1972), pp. 75–81.
[2] — *Joint spectra*, Studia Math. 45 (1973), pp. 225–237.
[3] N. M. Day, *Normed Linear Spaces*, Berlin 1973.
[4] P. Halmos, *A Hilbert Space Problem Book*, Toronto–London 1967.
[5] F. Riesz and B. Sz.-Nagy, *Leçons d'Analyse Fonctionnelle*, Budapest 1965.
[6] Z. Słodkowski and W. Żelazko, *On joint spectra of commuting families of operators*, Studia Math. 50 (1974), pp. 127–148.

Received December 5, 1975 (1100)

STUDIA MATHEMATICA, T. LXI. (1977)

Some properties of functions with bounded mean oscillation

by

UMBERTO NERI (College Park, Md.)*

Abstract. Functions with bounded mean oscillation (BMO) have been shown to be of great interest in several areas of analysis and probability. In the first part of this paper, we examine the basic properties of these functions, giving a new proof of the John–Nirenberg inequality and proving the completeness of the function space. In the second part, we discuss various examples and remarks which have arisen recently, and we give another characterization of the harmonic functions in a half-space with boundary values in BMO.

Introduction. Nearly 15 years ago Fritz John and Louis Nirenberg introduced in [6] the class of functions with bounded mean oscillation, in view of its apparent interest in real analysis as well as in partial differential equations. Ten years later, Charles Fefferman [3] gave new impetus to this subject by discovering, in his famous duality theorem, the important link between BMO and harmonic analysis in several real variables. Thus, he set the stage, in his joint work with Elias Stein [4], to several new developments and applications. For instance, references [1], [2], [5], [7], [8], [9] and [10] show a part of this outgrowth in various branches of analysis, whereas the works of Burkholder, Gundy, A. Garsia and others exemplify the new developments taking place in probability theory, stimulated by the revival of interest in BMO.

§ 1. BMO revisited. Let us consider locally integrable functions f on $\boldsymbol{R}^n$ and "regular sets" Q (such as balls, or cubes with sides parallel to the axes), and denote by f_Q the integral average

$$f_Q=|Q|^{-1}\int_Q f(x)\,dx$$

or, mean-value of f on Q. We call the function

$$Q\ni x\mapsto|f(x)-f_Q|$$

* The author gives thanks to Lars Gårding and the Institut Mittag-Leffler for the support and hospitality received in May 1975.