BEYOND THE MINIMAX PRINCIPLE

CHANDLER DAVIS

ABSTRACT. Let A be selfadjoint and P an orthoprojector. Then for any real λ, the compression of A to the subspace $P\mathfrak{K}$ has spectral projector belonging to $]\lambda, \infty[$ no larger in dimensionality than does A. This generalizes the classical Poincaré inequality. Several other like generalizations are given of various versions of the "minimax principle".

For A a bounded selfadjoint operator on Hilbert space $\mathfrak{K}$, and for P a proper orthoprojector, with A written

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

according to the decomposition $\mathfrak{K} = P\mathfrak{K} \oplus (P\mathfrak{K})^\perp$, there is a familiar family of results relating the spectrum of A to that of its (selfadjoint) compression $A_{11} = PA|_{P\mathfrak{K}}$ [1, II, §1], [3, II, §§1, 2], [2]. Deservedly best known is what is sometimes called the Poincaré inequality. This will be stated for the case where A is positive compact, with positive eigenvalues (multiplicity counted) $\lambda_1 \geqslant \lambda_2 \geqslant \cdots$; then A_{11} is automatically also positive compact, with eigenvalues $\mu_1 \geqslant \mu_2 \geqslant \cdots$. The Poincaré inequality asserts that $\lambda_k \geqslant \mu_k$. This conclusion can be reformulated: the spectral projector of A_{11} belonging to $]\lambda, \infty[$ (any λ) has at most the dimensionality of the spectral projector of A belonging to the same interval. In this form it looks as though it ought to be true without the hypothesis of compactness, and the first business of this paper is that indeed it is (see Theorems 1 and 3).

Consider again the case where A is positive compact. Another well-known inequality says that, if $(P\mathfrak{K})^\perp$ is n-dimensional, then $\lambda_{n+k} \leqslant \mu_k$. This result is also reformulated below and proved for arbitrary selfadjoint A in Theorems 2 and 4.

The two inequalities for μ_k yield variational characterizations of the eigenvalues, leading the name "minimax principle" to be applied loosely to the whole circle of ideas.

In the compact case, results on perturbation follow from the Poincaré inequality. The Weyl perturbation theorem says that if H is selfadjoint compact then the positive eigenvalues of $A + H$ differ from those of A by no more than $\|H\|$, while the Weyl monotonicity theorem says that if in addition $H \geqslant 0$ then the eigenvalues of $A + H$ are no lower than those of A. The corresponding general versions here are inferred in the familiar way, see §2.

It is time to admit that, in this removing of finiteness restrictions from the

minimax principle, we grievously undermine it. The conclusions of the classical theorems are quite strong because of the circumstance that a positive compact operator is characterized up to unitary equivalence by assigning the dimensionalities of its spectral projectors belonging to $]\lambda, \infty[$ (all $\lambda > 0$) and $\{0\}$, so that the classical theorems say a good deal about the unitary-equivalence type of the operators involved. This is lost in the general situation. (For example, let A be any orthoprojector on $\mathcal{K}$ with $\dim(P\mathcal{K}) = \dim(P\mathcal{K})^{\perp} = \dim \mathcal{K}$, and let B be any selfadjoint operator on $\mathcal{K}$ with spectrum $\subseteq [0, 1]$. Then A cannot be distinguished from $A \oplus B$ just by dimensionalities of spectral projectors belonging to intervals $]-\infty, \lambda[$, $]-\infty, \lambda]$, $[\lambda, \infty[$, $]\lambda, \infty[$.) The gap is partially bridged by the Theorem of §3, especially part (ii), where the generality is less total and the conclusion has more import.

The reader is encouraged to think of $\mathcal{K}$ as a separable Hilbert space. However, results and proofs are left intact if higher dimensionalities are countenanced, and even for the dimension function of a semifinite von Neumann algebra if all isometries entering are also required to lie in the algebra.

1. Bounding spectral projectors.

NOTATION. For any selfadjoint A and real λ, let $\{A > \lambda\}$ denote the range of the spectral projector for A belonging to $]\lambda, \infty[$, and let $\dim\{A > \lambda\}$ denote its dimensionality (analogously for $\{A \geqslant \lambda\}$, etc.). Also, as usual, A^{+} and A^{-} denote the positive operators such that $A = A^{+} - A^{-}$ and $A^{+}A^{-} = 0$. Similar looking but unrelated is the symbol $\mathcal{R}^{-}$: the closure of the range of any operator B is written $\mathcal{R}^{-}(B)$.

LEMMA. *For any bounded Q and C,*
$$\dim(\mathcal{R}^{-}(QC)) \leqslant \dim(\mathcal{R}^{-}(C)), \qquad \dim(\mathcal{U}(CQ)) \geqslant \dim(\mathcal{U}(C)).$$
For any bounded B,
$$\mathcal{R}^{-}(B) = \mathcal{R}^{-}(BB^{*}), \qquad \mathcal{U}(B^{*}B) = \mathcal{U}(B).$$

This is elementary.

THEOREM 1. *For any selfadjoint A and any orthoprojector P, we have*
(i) $\dim\{PA|_{P\mathcal{K}} > \lambda\} \leqslant \dim\{A > \lambda\}$,
(ii) $\dim\{PA|_{P\mathcal{K}} < \lambda\} \leqslant \dim\{A < \lambda\}$.

PROOF OF (i). Assume without loss of generality that $\lambda = 0$. We are to prove that the subspace $\{PA|_{P\mathcal{K}} > 0\} = \mathcal{R}^{-}((PAP)^{+})$ has dimensionality no greater than that of $\{A > 0\} = \mathcal{R}^{-}(A^{+})$. Let Q denote the orthoprojector onto $\mathcal{R}^{-}((PAP)^{+})$. Then
$$(PAP)^{+} = QPAPQ = QAQ \leqslant QA^{+}Q, \tag{1}$$
and, both operators being positive, $\mathcal{R}^{-}(QA^{+}Q) \supseteq \mathcal{R}^{-}((PAP)^{+})$. But, using the Lemma,
$$\dim(\mathcal{R}^{-}(QA^{+}Q)) = \dim(\mathcal{R}^{-}(QA^{+1/2}))$$
$$\leqslant \dim(\mathcal{R}^{-}(A^{+1/2})) = \dim(\mathcal{R}^{-}(A^{+})),$$
which completes the proof.

THEOREM 2. *For any selfadjoint A and any orthoprojector P, we have*
(i) $\dim\{PA|_{P\mathcal{K}} \geqslant \lambda\} + \dim(P\mathcal{K})^{\perp} \geqslant \dim\{A \geqslant \lambda\}$,
(ii) $\dim\{PA|_{P\mathcal{K}} \leqslant \lambda\} + \dim(P\mathcal{K})^{\perp} \geqslant \dim\{A \leqslant \lambda\}$.

PROOF OF (ii). Assume without loss of generality that $\lambda = 0$. We are to prove that the subspace $\{PA|_{P\mathcal{K}} \leqslant 0\} \oplus (P\mathcal{K})^{\perp} = \mathcal{R}((PAP)^{+})$ has dimensionality no less than that of $\{A \leqslant 0\} = \mathcal{R}(A^{+})$. Let Q be as above; this time, (1) avails to show that $\mathcal{R}(QA^{+}Q) \subseteq \mathcal{R}((PAP)^{+})$. Using the Lemma,

$$\dim(\mathcal{R}(QA^{+}Q)) = \dim(\mathcal{R}(A^{+1/2}Q)) \geqslant \dim(\mathcal{R}(A^{+1/2})) = \dim(\mathcal{R}(A^{+})),$$

which completes the proof.

The remaining two theorems, which add nothing in the finite-dimensional case, in the general case are distinct and even entail a new difficulty.

THEOREM 3. *For any selfadjoint A and any orthoprojector P, we have*
(i) $\dim\{PA|_{P\mathcal{K}} \geqslant \lambda\} \leqslant \dim\{A \geqslant \lambda\}$,
(ii) $\dim\{PA|_{P\mathcal{K}} \leqslant \lambda\} \leqslant \dim\{A \leqslant \lambda\}$.

PROOF OF (i). Assume without loss of generality that $\lambda = 0$. We are to prove that the subspace

$$\{PA|_{P\mathcal{K}} \geqslant 0\} = \bigcap_{\lambda < 0} \mathcal{R}^{-}((P(A - \lambda)P)^{+})$$

has dimensionality no greater than that of

$$\{A \geqslant 0\} = \bigcap_{\lambda < 0} \mathcal{R}^{-}((A - \lambda)^{+}).$$

Let Q_{λ} denote the orthoprojector onto $\mathcal{R}^{-}((P(A - \lambda)P)^{+})$. The relation (1) extended to variable λ yields

$$\bigcap_{\lambda < 0} \mathcal{R}^{-}(Q(A - \lambda)^{+}Q) \supseteq \bigcap_{\lambda < 0} \mathcal{R}^{-}((P(A - \lambda)P)^{+}).$$

At the very last step, where the elementary lemma served in Theorem 1, we now need the fact that

$$\dim\left(\bigcap_{\lambda < 0} \mathcal{R}^{-}(Q_{\lambda}(A - \lambda)^{+1/2})\right) \leqslant \dim\left(\bigcap_{\lambda < 0} \mathcal{R}^{-}((A - \lambda)^{+1/2})\right).$$

This is, however, no serious obstacle: any collection $\{u_{i}\}_{i \in \mathcal{I}}$ of vectors in the space on the left, such that the numbers $\text{dist}(u_{i}, \text{span}(u_{j}: j \neq i))$ are bounded away from 0, readily yields, for arbitrary $\lambda < 0$, a collection $\{v_{i}\}_{i \in \mathcal{I}}$ in $\mathcal{R}((A - \lambda)^{+1/2})$ with the like property. This completes the proof.

THEOREM 4. *For any selfadjoint A and any orthoprojector P, we have*
(i) $\dim\{PA|_{P\mathcal{K}} > \lambda\} + \dim(P\mathcal{K})^{\perp} \geqslant \dim\{A > \lambda\}$,
(ii) $\dim\{PA|_{P\mathcal{K}} < \lambda\} + \dim(P\mathcal{K})^{\perp} \geqslant \dim\{A < \lambda\}$.

Proof analogous.

2. Perturbation of spectrum.

THEOREM 5. *For A selfadjoint and $H \geqslant 0$,*
(i) $\dim\{A + H > \lambda\} \geqslant \dim\{A > \lambda\}$,
(ii) $\dim\{A + H \geqslant \lambda\} \geqslant \dim\{A \geqslant \lambda\}$.

PROOF. For (i), let P be the orthoprojector onto the subspace $\{A > \lambda\}$. Then clearly

$$P\mathcal{H} = \{A > \lambda\} = \{PA|_{P\mathcal{H}} > \lambda\} = \{P(A + H)|_{P\mathcal{H}} > \lambda\}$$

(using $H \geqslant 0$). Application of Theorem 1(i) to $A + H$ completes the proof. Part (ii) follows from Theorem 3(i) in the same way.

THEOREM 6. *For A and H selfadjoint,*
(i) $\dim\{A + H > \lambda + \|H\|\} \leqslant \dim\{A > \lambda\}$,
(ii) $\dim\{A + H < \lambda - \|H\|\} \leqslant \dim\{A < \lambda\}$,
(iii) $\dim\{A + H \geqslant \lambda + \|H\|\} \leqslant \dim\{A \geqslant \lambda\}$,
(iv) $\dim\{A + H \leqslant \lambda - \|H\|\} \leqslant \dim\{A \leqslant \lambda\}$.

Consider (i), for example. Appeal twice to Theorem 5(i):

$$\dim\{A + H > \lambda + \|H\|\} \leqslant \dim\{A + H^+ > \lambda + \|H\|\}$$
$$= \dim\{A + H^+ - \|H\| > \lambda\} \leqslant \dim\{A > \lambda\}.$$

Inequalities in the other direction can be obtained by applying Theorems 2 and 4 in the same way Theorems 1 and 3 were applied here. Details are left to the reader.

Here is a slightly less immediate variant of Theorem 6.

THEOREM 7. *For A and H selfadjoint and $\lambda > 0$,*
(i) $\dim\{\mu - \lambda < A < \mu + \lambda\} \leqslant \dim\{\mu - \lambda - \|H\| < A + H < \mu + \lambda + \|H\|\}$,
(ii) $\dim\{\mu - \lambda \leqslant A \leqslant \mu + \lambda\} \leqslant \dim\{\mu - \lambda - \|H\| \leqslant A + H \leqslant \mu + \lambda + \|H\|\}$.

LEMMA. $(A + H)^+ + (A + H)^- \leqslant A^+ + H^+ + A^- + H^-$.

This is elementary.

PROOF OF THEOREM 7. Without loss of generality, let $\mu = 0$. Consider, for example, inequality (i). On the left, note that we have $\{A^+ + A^- < \lambda\} = \{-\lambda < A < \lambda\}$; and similarly on the right. Thus it is required to prove that $\dim\{A^+ + A^- < \lambda\} \leqslant \dim\{(A + H)^+ + (A + H)^- < \lambda + \|H\|\}$.

Now $\dim\{A^+ + A^- < \lambda\} \leqslant \dim\{A^+ + A^- + H^+ + H^- < \lambda + \|H\|\}$ by virtue of Theorem 6(ii). An application of the Lemma completes the proof.

3. Better description of compressions.

THEOREM 8. *Let A be selfadjoint, let P be an orthoprojector and let E be the orthoprojector onto $\{A > 0\}$. Then (i) $PA|_{P\mathcal{H}} \leqslant WA^+W^*$ where W is some partial isometry into $P\mathcal{H}$; and (ii) in order that it be possible in (i) to choose W an isometry of $E\mathcal{H}$ onto $P\mathcal{H}$, it is sufficient that $(P\mathcal{H}) \cap (E\mathcal{H})^\perp$ and $(P\mathcal{H})^\perp \cap (E\mathcal{H})$ both be zero.*

This kind of bound on the compression $PA|_{P\mathcal{K}}$ is stronger than those of §1; in particular, (i) here implies the bound there on $\dim\{PA|_{P\mathcal{K}} > 0\}$. In the familiar situation–say, A positive compact and P of finite rank–we can apply the results of §1 and deduce the conclusion of Theorem 8(ii) without any special assumption on the relative position of $P\mathcal{K}$ and $E\mathcal{K}$. Once we leave this situation, as by allowing A to have an interval of continuous spectrum, Theorem 8 is asserting much more.

The hypothesis in (ii) that $(P\mathcal{K}) \cap (E\mathcal{K})^{\perp}$ and $(P\mathcal{K})^{\perp} \cap (E\mathcal{K})$ both be zero is at any rate a natural one. One important context in which these questions are raised is the approximation of spectral subspaces (such as our $E\mathcal{K}$) by more accessible subspaces. If $P\mathcal{K}$ were a computed subspace intended to be approximately the same as $E\mathcal{K}$, it might well be verifiable that $\|P - E\| < 1$, hence that $(P\mathcal{K}) \cap (E\mathcal{K})^{\perp} = (P\mathcal{K})^{\perp} \cap (E\mathcal{K}) = 0$ (and incidentally $\dim P\mathcal{K} = \dim E\mathcal{K}$). The next weaker hypothesis to try in (ii) would seemingly be $\dim((P\mathcal{K}) \cap (E\mathcal{K})^{\perp}) = \dim((P\mathcal{K})^{\perp} \cap (E\mathcal{K}))$, but that would be too weak for the following proof.

PROOF OF THEOREM 8. Clearly $PAP = PA^{+}P - PA^{-}P \leqslant PA^{+}P$. Consider the polar resolution $PA^{+1/2} = WH$ with $H \geqslant 0$ and W a partial isometry. By definition, W takes $\mathcal{R}^{-}(A^{+1/2}P) \subseteq E\mathcal{K}$ into $\mathcal{R}^{-}(PA^{+1/2}) \subseteq P\mathcal{K}$, so we may write the inequality as $PA|_{P\mathcal{K}} \leqslant WH^{2}W^{*}$, with $\mathcal{R}(W) \subseteq P\mathcal{K}$.

For (i), it remains to prove that $H^{2} \leqslant A^{+}$. But $H^{2} = A^{+1/2}PA^{+1/2}$ and $P \leqslant 1$.

As to (ii), for it to be possible that W be extended to an isometry on $E\mathcal{K}$ onto $P\mathcal{K}$, it is necessary and sufficient that

$$\dim((E\mathcal{K}) \cap \mathcal{R}(PA^{+1/2})) = \dim((P\mathcal{K}) \cap \mathcal{R}(A^{+1/2}P)).$$

The definition chosen for E guarantees that $A^{+1/2}$ is one-to-one, and the special hypothesis $(P\mathcal{K}) \cap (E\mathcal{K})^{\perp} = (P\mathcal{K})^{\perp} \cap (E\mathcal{K}) = 0$ guarantees that PE and EP are one-to-one on their respective spaces; this makes the null-spaces of $PA^{+1/2}$ and its adjoint both zero. The proof is complete.

A final remark is made here delimiting the scope of applicability of this theorem. If only A and P are given, we might seek λ such that, for some orthoprojector E with $\{A > \lambda\} \subseteq E\mathcal{K} \subseteq \{A \geqslant \lambda\}$, we could use such an argument. In the finite-dimensional case, all we need of λ is that $\dim\{A > \lambda\} \leqslant \dim(P\mathcal{K}) \leqslant \dim\{A \geqslant \lambda\}$; this is always realizable, and the information sought is obtained. In the general case, snags abound. Here is an example to exhibit one of the difficulties. Let A be $\mathrm{diag}(-1, -\frac{1}{2}, -\frac{1}{3}, \ldots)$, so that $\{A \geqslant 0\}$ is zero, but for any $\lambda < 0$, $\{A > \lambda\}$ is infinite-dimensional; and let P be of finite rank. Then for no λ does $A - \lambda$ behave as in the above proof.

References

1. I. C. Gohberg and M. G. Kreĭn, *Introduction to the theory of linear non-selfadjoint operators*, Izdat. "Nauka", Moscow, 1965; English transl., Transl. Math. Monographs, vol. 18, Amer. Math. Soc., Providence, R. I., 1969.

2. H. F. Weinberger, *Variational methods for eigenvalue approximation*, CBMS Regional Conf. in Appl. Math., No. 15, SIAM, Philadelphia, Pa., 1974.

3. A. Weinstein and W. Stenger, *Methods of intermediate problems for eigenvalues: theory and ramifications*, Academic Press, New York, 1972.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, TORONTO M5S 1A1, CANADA