MATH. SCAND. 75 (1994), 255–262

A REMARK ON PERTURBATIONS OF COMPACT OPERATORS

LARS HÖRMANDER and ANDERS MELIN

A generic linear transformation in $\mathbf{C}^N$ has only simple eigenvalues, for the discriminant of the characteristic equation is not identically 0. It is also well known that compact operators in a Banach space admit arbitrarily small perturbations with simple spectrum except at the origin. The purpose of this note is to discuss how far the multiplicity can be reduced by *perturbations of given finite rank.* The main result (Theorem 3) states that for perturbations of rank k which are generic in a certain sense the space of generalized eigenvectors belonging to any eigenvalue $\lambda \neq 0$ is reduced by removal of the largest k Jordan boxes for that eigenvalue while all new eigenvalues created are simple. No perturbation of rank k can lead to a larger reduction of the dimension. In particular, there is a perturbation of rank k making all eigenvalues $\neq 0$ simple if and only if the kernel of $T - \lambda$ is of dimension $\leqq k + 1$ and the kernel of $(T - \lambda I)^2$ is of dimension $\leqq 2k + 1$ for every $\lambda \neq 0$.

In the following lemma we introduce a convenient and systematic way of describing the multiplicity of generalized eigenvalues:

LEMMA 1. *Let T be a linear operator in a finite dimensional complex vector space V, and let $\lambda \in \mathbf{C}$. Then*

$$n(T, r, \lambda) = \begin{cases} \dim \operatorname{Ker}(T - \lambda I)^r, & r \geqq 1, \\ 0, & r = 0, \end{cases}$$

is a concave increasing sequence. The Legendre transform

$$\tilde{n}(T, k, \lambda) = \max_{r \geqq 0} (n(T, r, \lambda) - kr), \quad k \geqq 0, \tag{1}$$

is a convex, decreasing, non-negative sequences. $\tilde{n}(T, 0, \lambda) = \max_{r \geqq 0} n(T, r, \lambda)$ is the dimension of the space of generalized eigenvectors with eigenvalue λ, also called the algebraic multiplicity of λ. We have

Received October 27, 1993.

$$\tilde{n}(T, k, \lambda) = 0 \Leftrightarrow \dim \operatorname{Ker}(T - \lambda I) = n(T, 1, \lambda) \leqq k,$$

$$\tilde{n}(T, k, \lambda) \leqq 1 \Leftrightarrow \dim \operatorname{Ker}(T - \lambda I)^j = n(T, j, \lambda) \leqq jk + 1, \quad j = 1, 2.$$

If $n(T, 1, \lambda) > k$ *then the maximum is attained in* (1) *when*

$$n(T, r + 1, \lambda) - n(T, r, \lambda) \leqq k \leqq n(T, r, \lambda) - n(T, r - 1, \lambda).$$

One can recover $n(T, \cdot, \lambda)$ *from* $\tilde{n}(T, \cdot, \lambda)$ *by the inversion formula*

$$n(T, r, \lambda) = \min_{k \geqq 0} (\tilde{n}(T, k, \lambda) + kr), \quad r \geqq 0. \tag{2}$$

PROOF. To shorten notation we may assume that $\lambda = 0$ and write $n(r) = n(T, r, 0)$, $\tilde{n}(k) = \tilde{n}(T, k, 0)$. That $n(\cdot)$ is concave means that

$$n(r + 1) - n(r) \leqq n(r) - n(r - 1), \quad r \geqq 1.$$

With the convention $T^0 = I$ the map

$$\operatorname{Ker} T^{r+1}/\operatorname{Ker} T^r \xrightarrow{T} \operatorname{Ker} T^r/\operatorname{Ker} T^{r-1}$$

is injective for $r \geqq 1$, for if $x \in \operatorname{Ker} T^{r+1}$ and $Tx \in \operatorname{Ker} T^{r-1}$ then $x \in \operatorname{Ker} T^r$. This proves the concavity. Hence $n(r) \leqq kr$ for all $r \geqq 0$, if this is true when $r = 1$, which proves that $\tilde{n}(k) = 0$ if and only if $n(1) \leqq k$. Similarly $\tilde{n}(k) \leqq 1$ implies $n(r) \leqq kr + 1$. If $n(1) > k$ it follows that $n(1) = k + 1$ and that $n(2) \leqq 2k + 1$; by the concavity these conditions imply $n(r) \leqq rk + 1$, $r \geqq 1$, hence $\tilde{n}(k) = 1$. If $n(1) > k$ the maximum of $n(r) - kr$ is assumed for some $r \geqq 1$, and maximality at r means precisely that

$$n(r - 1) - k(r - 1) \leqq n(r) - kr, \quad n(r + 1) - k(r + 1) \leqq n(r) - kr,$$

that is, $n(r + 1) - n(r) \leqq k \leqq n(r) - n(r - 1)$.

To prove the inversion formula (2) for the Legendre transform we note that (1) implies that

$$n(r) - kr \leqq \tilde{n}(k), \quad \text{for all} \quad k \geqq 0, \quad \text{hence} \quad n(r) \leqq \min_{k \geqq 0} (\tilde{n}(k) + kr).$$

On the other hand, with $k = n(r + 1) - n(r)$ the concavity of $n(\cdot)$ gives

$$n(s) \leqq n(r) + k(s - r), \quad s \geqq 0, \quad \text{hence} \quad \tilde{n}(k) + kr \leqq n(r).$$

The lemma is proved.

To explain the properties of $\tilde{n}(T, k, \lambda)$ we consider, still with $\lambda = 0$, a partial Jordan decomposition $V = V_1 \oplus V_2$ where $TV_j \subset V_j$, and the matrix of the restriction T_1 of T to $V_1 = \mathbf{C}^m$ is a standard $m \times m$ Jordan box

$$T_1 = \begin{pmatrix} 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \tag{3}$$

while V_2 is the direct sum of the generalized eigenspaces belonging to eigenvalues $\neq 0$ and T invariant spaces of dimension $\leqq m$ where T is nilpotent. Let T_2 be the restriction of T to V_2 and define $n_2(r) = \dim \operatorname{Ker} T_2^r$ when $r > 0$, $n_2(0) = 0$. Then $n(r) = r + n_2(r)$ when $r \leqq m$, and we obtain if $k \geqq 1$

$$\tilde{n}(k) = \max_{r \geqq 0} (n(r) - kr) = \max_{r \geqq 0} (n_2(r) - (k-1)r) = \tilde{n}(T_2, k-1, 0), \tag{4}$$

for the maxima are attained when $r \leqq m$ because $n(r)$ and $n_2(r)$ are constant when $r \geqq m$. Since $\tilde{n}(T, 0, 0)$ is the dimension of the space of generalized eigenvectors with eigenvalue 0 we conclude by repeated use of (4) that *$\tilde{n}(k)$ is the dimension of the space remaining in the Jordan decomposition of the space of generalized eigenvectors with eigenvalue* 0 *if k boxes of highest possible dimension are dropped.* This simple result is the reason why it is useful to introduce the Legendre transform $\tilde{n}$. Note also that

$$2n(r) - n(r-1) - n(r+1) = \delta_{rm} + 2n_2(r) - n_2(r-1) - n_2(r+1)$$

is the number of $r \times r$ boxes in the Jordan decomposition of T when $r \geqq 1$.

Denote by $\mathscr{L}(V)$ the space of linear operators in V and let $\mathscr{L}_k(V)$ be the closed subset consisting of operators of rank $\leqq k$. It is clear that the set $\mathscr{L}_k(V) \setminus \mathscr{L}_{k-1}(V)$ of operators of rank exactly k is an open dense subset of $\mathscr{L}_k(V)$ if $1 \leqq k \leqq \dim V$. More generally, if Q is a polynomial in $\mathscr{L}(V)$, that is, $Q(T)$ is a polynomial in the matrix elements of $T \in \mathscr{L}(V)$ with respect to some chosen basis, then either Q vanishes identically in $\mathscr{L}_k(V)$ or else $\{S \in \mathscr{L}_k(V); Q(S) \neq 0\}$ is an open dense subset of $\mathscr{L}_k(V)$. That it is open is obvious. If it is not dense we can find $S_0 \in \mathscr{L}_k(V) \setminus \mathscr{L}_{k-1}(V)$ such that $Q = 0$ in a neighborhood of S_0 in $\mathscr{L}_k(V)$. Thus $Q(AS_0B) = 0$ for all $A, B \in \mathscr{L}(V)$ in a neighborhood of the identity, which implies that this is true for all $A, B \in \mathscr{L}(V)$. Since every $S \in \mathscr{L}_k(V)$ is of the form AS_0B it follows that $Q(S) = 0$, $S \in \mathscr{L}_k(V)$.

THEOREM 2. *Let V be a finite dimensional vector space over* $\mathbf{C}$, *let $T \in \mathscr{L}(V)$ and $S \in \mathscr{L}_k(V)$ where $k \geqq 1$. Then*

$$\tilde{n}(T + S, \nu, \lambda) \geqq \tilde{n}(T, \nu + k, \lambda), \quad \nu \geqq 0. \tag{5}$$

For fixed T there is an open dense subset Γ_T of $\mathscr{L}_k(V)$ such that for every $S \in \Gamma_T$

(i) *there is equality in* (5) *if λ is an eigenvalue of T,*

(ii) *all other eigenvalues of $T + S$ are simple.*

PROOF. Set $T_\lambda = T - \lambda I$, and let $N = \operatorname{Ker} S$. We have

$$\operatorname{Ker}(T_\lambda + S)^r \supset N_r = \{x \in \operatorname{Ker} T_\lambda^r;\ x \in N,\ T_\lambda x \in N, \ldots, T_\lambda^{r-1} x \in N\}.$$

In fact, if $x \in N_r$ then $(T_\lambda + S)x = T_\lambda x$, $(T_\lambda + S)^2 x = T_\lambda^2 x, \ldots, (T_\lambda + S)^r x = T_\lambda^r = 0$. Hence

$$\dim \operatorname{Ker}(T_\lambda + S)^r \geqq \dim \operatorname{Ker} T_\lambda^r - r \operatorname{codim} N \geqq \dim \operatorname{Ker} T_\lambda^r - rk,$$

$$\dim \operatorname{Ker}(T_\lambda + S)^r - vr \geqq \dim \operatorname{Ker} T_\lambda^r - (v + k)r,$$

which proves inequality (5).

If $k \geqq d = \dim V$ then equality in (5) means that $\tilde{n}(T + S, v, \lambda) = 0$ for every v, that is, that λ is not an eigenvalue of $T + S$. This condition is independent of k, and $\mathscr{L}_k(V) = \mathscr{L}_d(V)$ when $k \geqq d$, so we may assume that $k \leqq d$ in the remaining part of the proof.

Assuming at first that $\lambda = 0$ and that T is nilpotent we shall now prove that for some S of rank k there is equality in (5) when $\lambda = 0$ while the eigenvalues $\neq 0$ of $T + S$ are simple. As above we choose a partial Jordan decomposition $V = V_1 \oplus V_2$ such that $TV_j \subset V_j$, the matrix of the restriction T_1 of T to $V_1 = \mathbf{C}^m$ is the $m \times m$ Jordan box (3), and the restriction T_2 of T to V_2 is the direct sum of operators with such matrices of size at most $m \times m$. Assuming as we may that the statement is already proved for spaces of lower dimension we can choose an operator S_2 in V_2 of rank $k - 1$ such that the eigenvalues $\neq 0$ of $T_2 + S_2$ are simple and

$$\tilde{n}(T_2 + S_2, v, 0) = \tilde{n}(T_2, v + k - 1, 0) = \tilde{n}(T, v + k, 0),$$

where the second equality follows from (4). If $m = 1$ we choose any S_1 different from the eigenvalues of $T_2 + S_2$. If $m > 1$ we define S_1 by the matrix with the element ε in the lower left corner and all others equal to 0. Then the characteristic equation of $T_1 + S_1$ is $\lambda^m - \varepsilon = 0$ so the roots are distinct and different from the eigenvalues of $T_2 + S_2$ for suitable ε. Set $S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}$. Then the eigenvalues $\neq 0$ of $T + S$ are simple and

$$\tilde{n}(T + S, v, 0) = \tilde{n}(T_2 + S_2, v, 0) = \tilde{n}(T, v + k, 0),$$

so there is equality in (5) when $\lambda = 0$.

In the general case let $\lambda_1, \ldots, \lambda_J$ be the different eigenvalues of T, and let V_j be the corresponding spaces of generalized eigenvectors. The statements (i) and (ii) in Theorem 2 consist of two parts:

(i) If λ is an eigenvalue of $T + S$ and $\lambda \neq \lambda_j$, $j = 1, \ldots, J$, then λ is a simple eigenvalue.

(ii) (5) is valid with equality when $\lambda = \lambda_j$, $j = 1, \ldots, J$.

In particular, for $\nu = 0$ the condition (ii) means that

(iii) $\nu_j = \tilde{n}(T, k, \lambda_j)$,

if ν_j is the dimension of the space of generalized eigenvectors of $T + S$ with eigenvalue λ_j. The conditions (i) and (iii) mean precisely that the polynomial

$$Q_S(\lambda) = \det(\lambda I - (T + S)) / \prod_{\nu_j > 1} (\lambda - \lambda_j)^{\mu_j - 1}, \quad \mu_j = \tilde{n}(T, k, \lambda_j),$$

has only simple zeros, that is, $D(S) \neq 0$ if $D(S)$ is the discriminant of Q_S, which is a polynomial in S. We shall prove below that $D(S)$ does not vanish identically in $\mathscr{L}_k(V)$. The full condition (ii) means in addition that S shall avoid a proper algebraic subvariety of $\mathscr{L}_k(V)$. In fact, inequality in (5) at λ_j means that

$$\dim \operatorname{Ker}(T + S - \lambda_j I)^r - \nu r > \tilde{n}(T, \nu + k, \lambda_j) \tag{6}$$

for some ν and r. Now if W is a linear map in V, a condition of the form $\dim \operatorname{Ker} W > \alpha$ means that rank $W < \dim V - \alpha$, which is equivalent to the vanishing of all minors of size $\dim V - \alpha$ in the matrix of W. Thus there are polynomials $Q_{\nu,r,j,1}(S), \dots, Q_{\nu,r,j,t}(S)$ such that (6) holds if and only if they all vanish. For arbitrary ν, r, j we can choose $\tau = \tau(\nu, r, j)$ so that $Q_{\nu,r,j,\tau}$ does not vanish identically in $\mathscr{L}_k(V)$, for we have proved that there is a perturbation of rank $\leqq k$ of the restriction T_j of T to V_j such that there is equality in (5) at λ_j. Let Q be the product of the polynomials $Q_{\nu,r,j,\tau(\nu,r,j)}$ taken over all ν, r, j such that $\nu r < \dim V$ and $0 < r < \dim V$. Then Q does not vanish identicallt in $\mathscr{L}_k(V)$, and when $S \in \mathscr{L}_k(V)$, $Q(S) \neq 0$, we have equality in (5) for all eigenvalues of T.

It remains to prove that there is some $S \in \mathscr{L}_k(V)$ such that $D(S) \neq 0$, that is, such that $T + S$ has only simple eigenvalues apart from the points λ_j which have algebraic multiplicity μ_j. Choose closed disjoint discs $D_1, \dots, D_J$ with centers $\lambda_1, \dots, \lambda_J$. The number of zeros of $\det(\lambda I - (T + S))$ in D_j is constant for small S, and the projection on the direct sum N_j of the corresponding generalized eigenspaces along the direct sum of the other spaces N_i is given by

$$P_j(S) = \frac{1}{2\pi i} \int_{\partial D_j} (zI - (T + S))^{-1} \, dz,$$

which is an analytic function of (the matrix elements of) S. We have $P_j(0) V_i = 0$ when $i \neq j$, and $P_j(0)$ restricted to V_j is the identity. The map Φ defined for small S by

$$\Phi(S) : V = \bigoplus_1^J V_j \ni (x_1, \dots, x_J) \mapsto \sum_1^J P_j(S) x_j \in V,$$

is equal to the identity when $S = 0$, so it is invertible when S is small enough. Now

$$\Phi(S)^{-1}(T + S)\Phi(S) = T + \tilde{S}(S),$$

where $\tilde{S}(S)$ depends analytically on S for small S and $\tilde{S}(0) = 0$. Here T restricts to an operator T_j in V_j, and $\tilde{S}(S)$ restricts to an operator $\tilde{S}_j(S)$ in V_j. In fact, $\Phi(S)$ defines a bijection of V_j on N_j, which is invariant under $T + S$. When the discriminant of $(\lambda - \lambda_j)^{1-\mu_j}\det(\lambda I_j - (T_j + \tilde{S}_j(S)))$ does not vanish, the eigenvalues in D_j are simple apart from λ_j which has algebraic multiplicity μ_j. Taking for S an operator mapping V_j to V_j and V_i to 0 for $i \neq j$, we know that the discriminant is not identically zero for small $S \in \mathscr{L}_k(V)$. Hence we can choose $S \in \mathscr{L}_k(V)$ small such that this discriminant does not vanish for any $j = 1, \ldots, J$. This implies that $T + S$ has only simple eigenvalues apart from the points λ_j, which have algebraic multiplicity μ_j, which completes the proof.

REMARK. We have proved more than stated: The set Γ_T in Theorem 2 can be taken as the complement in $\mathscr{L}_k(V)$ of a proper algebraic subset.

Let B be a Banach space and denote by $\mathscr{L}(B)$ the space of continuous linear operators in B. Since

$$\mathscr{L}_k(B) = \{S \in \mathscr{L}(B); \operatorname{rank} S \leqq k\}$$

is a closed subset, it is a complete metric space. If T is a compact operator in B, we shall denote by $\mathscr{L}_k^T(B)$ the closure in $\mathscr{L}_k(B)$ of the set of elements $S \in \mathscr{L}_k(B)$ such that there is a topological direct sum decomposition $B = B_1 \oplus B_2$ with B_1 finite dimensional, $TB_j \subset B_j$, $j = 1, 2$, and $SB_2 = 0$.

THEOREM 3. *Let T be a compact linear operator in the Banach space B. If $S \in \mathscr{L}_k(B)$ then* (5) *is valid when $\lambda \neq 0$. There is a first category subset Σ of $\mathscr{L}_k^T(B)$ such that for all $S \in \mathscr{L}_k^T(B) \setminus \Sigma$ there is equality in* (5) *when $\lambda \neq 0$ is an eigenvalue of T, and all other eigenvalues $\lambda \neq 0$ of $T + S$ are simple. In particular, the spectrum is then simple for all $\lambda \neq 0$ such that*

$$\dim \operatorname{Ker}(T - \lambda I) \leqq k + 1, \quad \dim \operatorname{Ker}(T - \lambda I)^2 \leqq 2k + 1. \tag{7}$$

PROOF. The proof of the inequality (5) in Lemma 2 works for Banach spaces with no real change. To discuss equality we introduce for $\varepsilon > 0$ the set Γ_ε of all $S \in \mathscr{L}_k^T(B)$ such that there is equality in (5) for all eigenvalues λ of T with $|\lambda| \geqq \varepsilon$ and all other eigenvalues λ of $T + S$ with $|\lambda| \geqq \varepsilon$ are simple. It is sufficient to prove that the complement Σ_ε of Γ_ε in $\mathscr{L}_k^T(B)$ is of the first category, for we can take $\Sigma = \cup_1^\infty \Sigma_{1/n}$. First we shall prove that Σ_ε is closed, that is, that Γ_ε is open.

Let $S_0 \in \Gamma_\varepsilon$ and let $|\lambda_0| \geqq \varepsilon$. If λ_0 is not an eigenvalue of $T + S_0$ then λ_0 has a compact neighborhood D such that $T + S$ has no eigenvalue in D when S is sufficiently close to S_0. If λ_0 is an eigenvalue of $T + S_0$ we can choose a compact neighborhood D of λ_0 containing no other eigenvalue. If λ_0 is a simple eigenvalue of $T + S_0$ it follows that $T + S$ has only a simple eigenvalue in D if S is sufficiently close to S_0. If λ_0 is an eigenvalue of $T + S_0$ which is not simple, then λ_0 is an eigenvalue of T since $S_0 \in \Gamma_\varepsilon$. For every $S \in \mathscr{L}_k^T(B)$ close to S_0 the equality in (5) remains valid at λ_0 when S_0 is replaced by S, for

$$\tilde{n}(T, v + k, \lambda_0) \leqq \tilde{n}(T + S, v, \lambda_0) \leqq \tilde{n}(T + S_0, v, \lambda_0) = \tilde{n}(T, v + k, \lambda_0),$$

where the first inequality follows from (5) and the second from the fact that

$$\dim \operatorname{Ker}(T + S - \lambda_0 I)^r \leqq \dim \operatorname{Ker}(T + S_0 - \lambda_0 I)^r, \tag{8}$$

if S is sufficiently close to S_0. To prove (8) we note that ϱ and $\varkappa$ can be chosen so that

$$\dim \operatorname{Ker}(T + S_0 - \lambda_0 I)^\varrho = \dim \operatorname{Ker}(T + S_0 - \lambda_0 I)^{\varrho+1} = \varkappa.$$

If S is sufficiently close to S_0, then (8) is valid for $r \leqq \varrho + \varkappa$, and both sides must be constant for $r \geqq \varrho + \varkappa$. In particular, the space of generalized eigenvectors of $T + S$ with eigenvalue λ_0 has the same dimension as when $S = S_0$ so there are no other eigenvalues in D. By the Borel-Lebesgue lemma we can cover $\{\lambda \in \mathbf{C}; \varepsilon \leqq |\lambda| \leqq \|T + S_0\| + 1\}$ by finitely many of the discs D discussed, which proves that Γ_ε is open. It remains to prove that the complement Σ_ε has no interior point.

If Σ_ε has an interior point it follows from the definition of $\mathscr{L}_k^T(B)$ (made for this purpose!) that there is an interior point S_0 such that for some topological direct sum decomposition $B = B_1 \oplus B_2$ with B_1 finite dimensional, $TB_j \subset B_j, j = 1, 2$, we have $S_0 B_2 = 0$. By standard Fredholm theory we can decompose B_2 further as $B_2 = B_3 \oplus B_4$ where $TB_3 \subset B_3$, $TB_4 \subset B_4$, B_3 is finite dimensional and the spectrum of the restriction of T to B_4 is contained in the disc $\{z; |z| \leqq \varepsilon/2\}$. Denote the restrictions of T to $E_1 = B_1 \oplus B_3$ and to $E_2 = B_4$ by T_1 and T_2. For operators S with $SE_2 = 0$, such as S_0, we denote by S_1 and S_2 the maps $E_1 \to E_1$ and E_1 to E_2 which it defines. Note that the equation $(T + S - z)^r x = 0$ for a generalized eigenvector $x = (x_1, x_2) \in E_1 \oplus E_2$ with eigenvalue z can be written

$$(T_1 + S_1 - z)^r x_1 = 0, \quad \sum_{j=0}^{r-1} (T_2 - z)^j S_2 (T_1 + S_1 - z)^{r-1-j} x_1 + (T_2 - z)^r x_2 = 0.$$

When $|z| > \varepsilon/2$ we conclude that the kernels of the powers of $T + S - z$ and $T_1 + S_1 - z$ have the same dimension, for the second equation can be solved for x_2.

By Theorem 2 we can choose S with $SE_2 = 0, S_2 = S_{02}$, and S_1 arbitrarily close to the component S_{01} of S_0, so that $T_1 + S_1$ has only simple eigenvalues apart from the eigenvalues λ_j of T_1, and (5) is valid with equality for T_1 and S_1 at each λ_j. Then

$$\tilde{n}(T, v + k, \lambda_j) = \tilde{n}(T_1, v + k, \lambda_j) = \tilde{n}(T_1 + S_1, v, \lambda_j) = \tilde{n}(T + S, v, \lambda_j),$$

if $|\lambda_j| \geqq \varepsilon$, which proves that $S \in \Gamma_\varepsilon$. Since S is arbitrarily close to S_0, this contradicts the assumption on S_0 and completes the proof, for the last assertion follows from Lemma 1.

We would like to thank Gerd Grubb for raising a question, motivated by control theory, which led to this note, and for constructive criticism of an earlier version.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF LUND
BOX 118
S-221 00 LUND
SWEDEN
e-mail: lvh@maths.lth.se, melin@maths.lth.se