A REMARK ON PERTURBATIONS OF COMPACT OPERATORS

LARS HÖRMANDER and ANDERS MELIN

A generic linear transformation in $\mathbb{C}^N$ 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 $T - \lambda I$ is of dimension $\leq k + 1$ and the kernel of $(T - \lambda I)^2$ is of dimension $\leq 2k + 1$ 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 $\lambda \in \mathbb{C}$. Then

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

is a concave increasing sequence. The Legendre transform

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

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

Received October 27, 1993.
\[ \hat{n}(T, k, \lambda) = 0 \iff \dim \ker(T - \lambda I) = n(T, 1, \lambda) \leq k, \]

\[ \hat{n}(T, k, \lambda) \leq 1 \iff \dim \ker(T - \lambda I)^j = n(T, j, \lambda) \leq jk + 1, \quad j = 1, 2. \]

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

\[ n(T, r + 1, \lambda) - n(T, r, \lambda) \leq k \leq n(T, r, \lambda) - n(T, r - 1, \lambda). \]

One can recover \( n(T, r, \lambda) \) from \( \hat{n}(T, r, \lambda) \) by the inversion formula

\[ n(T, r, \lambda) = \min_{k \geq 0} (\hat{n}(T, k, \lambda) + kr), \quad r \geq 0. \]

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

\[ n(r + 1) - n(r) \leq n(r) - n(r - 1), \quad r \geq 1. \]

With the convention \( T^0 = I \) the map

\[ \ker T^{r+1}/\ker T^r \rightarrow \ker T^r/\ker T^{r-1} \]

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

\[ n(r - 1) - k(r - 1) \leq n(r) - kr, \quad n(r + 1) - k(r + 1) \leq n(r) - kr, \]

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

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

\[ n(r) - kr \leq \hat{n}(k), \quad \text{for all} \quad k \geq 0, \quad \text{hence} \quad n(r) \leq \min_{k \geq 0} (\hat{n}(k) + kr). \]

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

\[ n(s) \leq n(r) + k(s - r), \quad s \geq 0, \quad \text{hence} \quad \hat{n}(k) + kr \leq n(r). \]

The lemma is proved.

To explain the properties of \( \hat{n}(T, k, \lambda) \) we consider, still with \( \lambda = 0 \), a partial Jordan decomposition \( V = V_1 \oplus V_2 \) where \( TV_1 \subset V_j \), and the matrix of the restriction \( T_1 \) of \( T \) to \( V_1 = \mathbb{C}^m \) is a standard \( m \times m \) Jordan box.
while $V_2$ is the direct sum of the generalized eigenspaces belonging to eigenvalues $\pm 0$ and $T$ invariant spaces of dimension $\leq m$ where $T$ is nilpotent. Let $T_2$ be the restriction of $T$ to $V_2$ and define $n_2(r) = \dim \ker T_2^r$ when $r > 0$, $n_2(0) = 0$. Then $n(r) = r + n_2(r)$ when $r \leq m$, and we obtain if $k \geq 1$

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

for the maxima are attained when $r \leq m$ because $n(r)$ and $n_2(r)$ are constant when $r \geq m$. 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 $r \geq 1$.

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

**Theorem 2.** Let $V$ be a finite dimensional vector space over $C$, let $T \in L(V)$ and $S \in L_k(V)$ where $k \geq 1$. Then

$$\tilde{n}(T + S, v, \lambda) \geq \tilde{n}(T, v + k, \lambda), \quad v \geq 0.$$

For fixed $T$ there is an open dense subset $\Gamma_T$ of $L_k(V)$ such that for every $S \in \Gamma_T$

(i) there is equality in (5) if $\lambda$ is an eigenvalue of $T$,

(ii) all other eigenvalues of $T + S$ are simple.
PROOF. Set $T_\lambda = T - \lambda I$, and let $N = \text{Ker } S$. We have
\[ \text{Ker } (T_\lambda + S)' \ni N = \{ x \in \text{Ker } T_\lambda' \mid x \in N, \ T_\lambda x \in N, \ldots, T_\lambda^{n-1} x \in N \}. \]
In fact, if $x \in N$, then $(T_\lambda + S)x = T_\lambda x$, $(T_\lambda + S)^2 x = T_\lambda^2 x, \ldots, (T_\lambda + S)^r x = T_\lambda^r x = 0$. Hence
\[ \dim \text{Ker } (T_\lambda + S)' \geq \dim \text{Ker } T_\lambda' - r \text{ codim } N \geq \dim \text{Ker } T_\lambda' - rk, \]
which proves inequality (5).

If $k \geq d = \dim V$ then equality in (5) means that $\tilde{n}(T + S, v, \lambda) = 0$ for every $v$, that is, that $\lambda$ is not an eigenvalue of $T + S$. This condition is independent of $k$, and $L_k(V) = L_d(V)$ when $k \geq d$, so we may assume that $k \leq d$ 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 $V_1 = \mathbb{C}^m$ 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
\[ \tilde{n}(T_2 + S_2, v, 0) = \tilde{n}(T_2, v + k - 1, 0) = \tilde{n}(T_2, v + k, 0), \]
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 $\epsilon$ in the lower left corner and all others equal to 0. Then the characteristic equation of $T_1 + S_1$ is $\lambda^m - \epsilon = 0$ so the roots are distinct and different from the eigenvalues of $T_2 + S_2$ for suitable $\epsilon$. 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 $\lambda$ is an eigenvalue of $T + S$ and $\lambda \neq \lambda_j, j = 1, \ldots, J$, then $\lambda$ is a simple eigenvalue.

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

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

\[
(iii) \quad v_j = \tilde{n}(T, k, \lambda_j),
\]

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

\[
Q_S(\lambda) = \det(\lambda I - (T + S))_{j \neq j} \prod_{v_j \geq 1} (\lambda - \lambda_j)^{\mu_j - 1}, \quad \mu_j = \tilde{n}(T, k, \lambda_j),
\]

has only simple zeros, that is, \( D(S) = 0 \) 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 \( \mathcal{L}_k(V) \). The full condition (ii) means in addition that \( S \) shall avoid a proper algebraic subvariety of \( \mathcal{L}_k(V) \). In fact, inequality in (5) at \( \lambda_j \) means that

\[
(6) \quad \dim \ker (T + S - \lambda_j I)^r - vr > \tilde{n}(T, v + k, \lambda_j)
\]

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

It remains to prove that there is some \( S \in \mathcal{L}_k(V) \) such that \( D(S) = 0 \), that is, such that \( T + S \) has only simple eigenvalues apart from the points \( \lambda_j \) which have algebraic multiplicity \( \mu_j \). Choose closed disjoint discs \( D_1, \ldots, D_J \) with centers \( \lambda_1, \ldots, \lambda_J \). 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 \( S \). 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 \( \Phi \) defined for small \( S \) by

\[
\Phi(S) : V = \bigoplus_i V_j \ni (x_1, \ldots, x_j) \mapsto \sum_i P_j(S)x_j \in V,
\]

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 $\lambda_j$ which has algebraic multiplicity $\mu_j$. 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 $S \in \mathcal{L}_k(V)$. Hence we can choose $S \in \mathcal{L}_k(V)$ 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 $\lambda_j$, which have algebraic multiplicity $\mu_j$, which completes the proof.

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

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

$$\mathcal{L}_k(B) = \{S \in \mathcal{L}(B); \text{rank } S \leq k\}$$

is a closed subset, it is a complete metric space. If $T$ is a compact operator in $B$, we shall denote by $\mathcal{L}_k^{\mathcal{T}}(B)$ the closure in $\mathcal{L}_k(B)$ of the set of elements $S \in \mathcal{L}_k(B)$ such that there is a topological direct sum decomposition $B = B_1 \oplus B_2$ with $B_1$ finite dimensional, $TB_1 \subset B_2$, $j = 1, 2$, and $SB_2 = 0$.

**Theorem 3.** Let $T$ be a compact linear operator in the Banach space $B$. If $S \in \mathcal{L}_k(B)$ then (5) is valid when $\lambda \neq 0$. There is a first category subset $\Sigma$ of $\mathcal{L}_k^{\mathcal{T}}(B)$ such that for all $S \in \mathcal{L}_k^{\mathcal{T}}(B) \setminus \Sigma$ there is equality in (5) when $\lambda = 0$ 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

$$\dim \text{Ker}(T - \lambda I) \leq k + 1, \quad \dim \text{Ker}(T - \lambda I)^2 \leq 2k + 1.$$  

**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 $\Gamma_{\varepsilon}$ of all $S \in \mathcal{L}_k^{\mathcal{T}}(B)$ such that there is equality in (5) for all eigenvalues $\lambda$ of $T$ with $|\lambda| \geq \varepsilon$ and all other eigenvalues $\lambda$ of $T + S$ with $|\lambda| \geq \varepsilon$ are simple. It is sufficient to prove that the complement $\Sigma_{\varepsilon}$ of $\Gamma_{\varepsilon}$ in $\mathcal{L}_k^{\mathcal{T}}(B)$ is of the first category, for we can take $\Sigma = \cup_{\varepsilon} \Sigma_{\varepsilon}$. First we shall prove that $\Sigma_{\varepsilon}$ is closed, that is, that $\Gamma_{\varepsilon}$ is open.

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

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

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

(8) \hspace{1cm} \dim \text{Ker}(T + S - \lambda_0)^r \leq \dim \text{Ker}(T + S_0 - \lambda_0)^r,$$

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

$$\dim \text{Ker}(T + S_0 - \lambda_0)^n = \dim \text{Ker}(T + S_0 - \lambda_0)^{n+1} = \kappa.$$

If $S$ is sufficiently close to $S_0$, then (8) is valid for $r \leq \varrho + \kappa$, and both sides must be constant for $r \geq \varrho + \kappa$. In particular, the space of generalized eigenvectors of $T + S$ with eigenvalue $\lambda_0$ 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 $\{\lambda \in C; \varepsilon \leq |\lambda| \leq \|T + S_0\| + 1\}$ by finitely many of the discs $D$ discussed, which proves that $\Gamma_\varepsilon$ is open. It remains to prove that the complement $\Sigma_\varepsilon$ has no interior point.

If $\Sigma_\varepsilon$ has an interior point it follows from the definition of $\mathcal{L}_k^T(B)$ (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 $\{z; |z| \leq \varepsilon/2\}$. 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

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