J. OPERATOR THEORY
54:1(2005), 125–136

SIMILARITY OF PERTURBATIONS OF HESSENBERG MATRICES

LEONEL ROBERT

Communicated by Kenneth R. Davidson

ABSTRACT. To every infinite lower Hessenberg matrix D is associated a linear operator on l_2. In this paper we prove the similarity of the operator $D - \Delta$, where Δ belongs to a certain class of compact operators, to the operator $D - \Delta'$, where Δ' is of rank one. We first consider the case when Δ is lower triangular and has finite rank; then we extend this to Δ of infinite rank assuming that D is bounded. In Section 3 we examine the cases when $D = S^t$ and $D = (S + S^t)/2$, where S denotes the unilateral shift.

KEYWORDS: *Hessenberg matrix, similarity of operators, associated polynomials of k-th kind, perturbation determinants.*

MSC (2000): 47A55, 47B37.

1. INTRODUCTION

A lower infinite *Hessenberg matrix* is a matrix $D = \{d_{i,j}\}_{i,j=0}^{\infty}$ with complex entries and such that $d_{i,j} = 0$ for $j > i + 1$. In this paper, when we refer to a Hessenberg matrix, we will additionally assume that it satisfies $d_{i,i+1} \neq 0$ for all $i \geqslant 0$.

Let us denote by l_2 the space of square summable sequences and by $\{e_i\}_{i=0}^{\infty}$ its canonical basis.

Let A be an infinite matrix such that there is $k \in \mathbb{Z}$ such that $A_{i,j} = 0$ for $j > i + k$ (e.g., A is Hessenberg). Then for every $x \in l_2$, Ax is well defined. Thus we can associate a linear operator to A, which we also denote by A, with domain

$$\mathrm{domain}(A) = \{x \in l_2 : Ax \in l_2\}.$$

The operator A is a closed operator defined on its maximal domain of definition. Moreover, A is the adjoint of the densely defined operator $\overline{A}^t$ with domain C_0, the linear span of the canonical basis $\{e_i\}_{i=0}^{\infty}$ ([7], Theorem 1). All the operators that we consider here act on the space l_2.

In Section 2 we study the similarity of the operator $D - \Delta$ to an operator of the form $D - \Delta'$, where Δ' has rank one. The main theorems of this section are Theorems 2.2 and 2.5. Theorem 2.2 assumes that Δ is lower triangular and of finite rank. Theorem 2.5 extends Theorem 2.2 to Δ not necessarily lower triangular or of finite rank, but assumes that D is bounded. In Section 3 we apply these results to $D = S^t$ and $D = \frac{S+S^t}{2}$, where S is the unilateral shift.

In the sequel we will work mostly in the context of infinite matrices and vectors with complex entries. Bounded operators acting on l_2 are identified with matrices through their matrix representation with respect to the canonical basis of l_2. It is with this identification in mind, that sometimes we will refer to a matrix as being “bounded" or “compact".

2. PERTURBATIONS OF HESSENBERG MATRICES

2.1. THE FRIEDRICHS'S METHOD. Here we describe a method that has been frequently used to prove similarity between operators ([3], [2], [1]). Our exposition follows Freeman [2] and Chan [1].

Given D and Δ we want to find U such that $U(D-\Delta)U^{-1} = D$. This equation implies that

$$[U, D] = UD - DU = U\Delta. \tag{2.1}$$

We notice that $[\cdot, D]$ is a derivation. Thus equation (2.1) is analogous to the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = X(t)P(t)$$

in a Banach algebra.

Keeping in mind this analogy, we first try to integrate the equation $[X, D] = \Delta$; that is, find $\Sigma(\Delta)$ such that $[\Sigma(\Delta), D] = \Delta$. Then a solution of (2.1) is given by the Peano series

$$U(\Delta) = I + \Sigma(\Delta) + \Sigma(\Sigma(\Delta)\Delta) + \Sigma(\Sigma(\Sigma(\Delta)\Delta)\Delta)) + \cdots . \tag{2.2}$$

In this paper we will use the following modification of this method. Suppose that instead we have $\Sigma(\Delta)$, matrix valued, and $\psi(\Delta)$, vector valued, such that

$$[\Sigma(\Delta), D] = \Delta - \psi(\Delta)e_0^t. \tag{2.3}$$

Then $U = U(\Delta)$ defined by (2.2) satisfies

$$[U, D] = U\Delta - \psi(U\Delta)e_0^t. \tag{2.4}$$

From this equation, and assuming that U is invertible, we get that

$$U(D-\Delta)U^{-1} = D - \psi(U\Delta)(U^{-1}e_0)^t.$$

That is, $D - \Delta$ is similar to a rank one perturbation of D.

But how does one make sense of these formulas as formulas in the bounded operators, or even as formulas in the matrices?

In order to do this, we try to find a left ideal $\mathcal{F} \subset B(l_2)$ and a norm $|\cdot|_{\mathcal{F}}$, such that:

(i) $\|\Sigma(A)\| \leqslant |A|_{\mathcal{F}}$;

(ii) $|BA|_{\mathcal{F}} \leqslant \|B\| |A|_{\mathcal{F}}$.

Then, for $|\Delta|_{\mathcal{F}} < 1$, equation (2.2) gives U bounded, and if one takes $|\Delta|_{\mathcal{F}} < \frac{1}{2}$, then U is invertible.

It is this general strategy what we will use to prove the Theorems 2.2 and 2.5 of this section.

2.2. FINITE RANK PERTURBATIONS. Let us denote by C^k the subspace of $\mathbb{C}^{\mathbb{N}}$ of infinite column vectors with the first $k-1$ components equal to zero and let $\mathcal{F}_k$ denote the space of matrices of finite rank such that $Xe_i = 0$ for $i > k$. Notice that if $X \in \mathcal{F}_k$ and A is an infinite matrix such that AX is well defined (e.g. A is Hessenberg), then $AX \in \mathcal{F}_k$.

LEMMA 2.1. *Let D be a Hessenberg matrix and ϕ an infinite column vector. There are $X_n^{\phi} \in \mathcal{F}_{n-1}$ and a vector ψ_n^{ϕ} such that*

$$[X_n^{\phi}, D] = \phi e_n^t - \psi_n^{\phi} e_0^t. \tag{2.5}$$

If $\phi \in C^n$, then X_n^{ϕ} is strictly lower triangular (i.e., X_n^{ϕ} is lower triangular with zeroes in the main diagonal).

The vector ψ_n^{ϕ} and the columns of X_n^{ϕ} belong to the subspace

$$\operatorname{span}\{D^i \phi : i \geqslant 0\}.$$

Proof. Let us define X_i^{ϕ} and ψ_i^{ϕ} by $X_0^{\phi} = 0$, $\psi_0^{\phi} = \phi$ and the recurrences (here $d_{n+1,n} \neq 0$ gets used)

$$\sum_{i=0}^{n+1} d_{n,i} X_i^{\phi} = D X_n^{\phi} + \phi e_n^t, \tag{2.6}$$

$$\sum_{i=0}^{n+1} d_{n,i} \psi_i^{\phi} = D \psi_n^{\phi}. \tag{2.7}$$

It is not difficult to prove by induction that X_n^{ϕ} and ψ_n^{ϕ} satisfy the conditions of the lemma. ∎

Let us define $\Sigma(\cdot)$ and $\psi(\cdot)$ by

$$\Sigma(\phi e_n^t) = X_n^{\phi}, \quad \psi(\phi e_n^t) = \psi_n^{\phi}, \tag{2.8}$$

and extend them by linearity to $\bigcup_{n=0}^{\infty} \mathcal{F}_n$. The following properties are readily verified:

(i) $\Sigma(\cdot)$ and $\psi(\cdot)$ satisfy (2.3);

(ii) $\Sigma(\mathcal{F}_n) \subset \mathcal{F}_{n-1}$;
(iii) if Δ is lower triangular, then $\Sigma(\Delta)$ is strictly lower triangular;
(iv) $\mathrm{Col}(\Sigma(\Delta)) \subset \sum_{n\geqslant 0} \mathrm{Col}(D^n\Delta)$ and $\psi(\Delta) \in \sum_{n\geqslant 0} \mathrm{Col}(D^n\Delta)$; here $\mathrm{Col}(A)$ denotes the space spanned by the columns of A.

THEOREM 2.2. *Let D be a Hessenberg matrix and Δ a lower triangular matrix of finite rank. Suppose that $\Delta \in \mathcal{F}_k$ for some k. Then there exist unique ψ and $U = I + X$ such that X is strictly lower triangular, $X \in \mathcal{F}_{k-1}$ and*

$$U(D-\Delta)U^{-1} = D - \psi e_0^t.$$

Suppose additionally that $\mathrm{Col}(D^n\Delta) \subset \mathrm{domain}(D)$ *for all* $n \geqslant 0$. *Then* $\psi \in l_2$ *and U is bounded with bounded inverse.*

Note: For every lower triangular matrix of finite rank there exists k such that $\Delta \in \mathcal{F}_k$.

Proof. Let us prove uniqueness first. Suppose that (U_1, ψ_1) and (U_2, ψ_2) satisfy the conditions of the theorem and $X_1 - X_2 \neq 0$. Let us write $U_1 = I + X_1$ and $U_2 = I + X_2$. Equation (2.4) implies that

$$[X_1 - X_2, D] = (X_1 - X_2)\Delta - (\psi_1 - \psi_2)e_0^t. \tag{2.9}$$

Let i be such that $0 \leqslant i \leqslant k-1$ and $X_1 - X_2 \in \mathcal{F}_i \backslash \mathcal{F}_{i-1}$.

We notice that:

1. If $A \in \mathcal{F}_i \backslash \mathcal{F}_{i-1}$ then $[A, D] \in \mathcal{F}_{i+1} \backslash \mathcal{F}_i$ (this is because $[A, D]e_{i+1} = d_{n-1,n}Ae_i \neq 0$).

2. If A and B are lower triangular and $A \in \mathcal{F}_i$, $B \in \mathcal{F}_n$ with $i < n$, then $AB \in \mathcal{F}_i$ (this is not difficult to verify).

Applying 1 and 2 in equation (2.9) we get that the right side is in $\mathcal{F}_i \backslash \mathcal{F}_{i-1}$, while the left side is in $\mathcal{F}_{i+1} \backslash \mathcal{F}_i$. This is a contradiction. Therefore $U_1 = U_2$, and this implies that $\psi_1 = \psi_2$.

Since $\Delta \in \mathcal{F}_k$, $\Sigma(\Delta) \in \mathcal{F}_{k-1}$ and Δ is lower triangular, property 2 above implies that $\Sigma(\Delta)\Delta \in \mathcal{F}_{k-1}$. In the same way, we see that $\Sigma(\Sigma(\Delta)\Delta)\Delta \in \mathcal{F}_{k-2}$ and so on. Thus, the right side of (2.2) is a finite sum. Let us take $U = U(\Delta)$ and $\psi = \psi(U\Delta)$. Since U is of the form $I + X$ with X strictly lower triangular, U is an invertible matrix. This proves the first part of the theorem.

Suppose that $\mathrm{Col}(D^n\Delta) \subset \mathrm{domain}(D)$ for all $n \geqslant 0$. By property (iv) of $\Sigma(\cdot)$ listed before, we have

$$\begin{aligned}
\mathrm{Col}(\Sigma(\Delta)) &\subset \sum_{n\geqslant 0} \mathrm{Col}(D^n\Delta) \subset \mathrm{domain}(D), \\
\mathrm{Col}(\Sigma(\Sigma(\Delta)\Delta)) &\subset \sum_{n\geqslant 0} \mathrm{Col}(D^n\Sigma(\Delta)) \subset \sum_{n\geqslant 0} \mathrm{Col}(D^n\Delta) \subset \mathrm{domain}(D), \\
\cdots \qquad & \qquad \cdots \qquad\qquad \cdots \qquad\qquad \cdots
\end{aligned}$$

This implies that $\mathrm{Col}(X) \subset \mathrm{domain}(D)$. Hence the columns of X are in l_2. Since only a finite number of the columns of X are nonzero, X and U represent bounded operators. Also, X is strictly lower triangular of finite rank. So U is injective and a compact perturbation of the identity. Hence, U has bounded inverse. ∎

Let $D_1 = (d^1_{i,j})$ and $D_2 = (d^2_{i,j})$ be two Hessenberg matrices. Denote by Υ the diagonal matrix such that $\Upsilon_{i,j} = 0$ if $i \neq j$, $\Upsilon_{0,0} = 1$ and $\Upsilon_{i,i} = \prod_{k=0}^{i-1} \frac{d^2_{k,k+1}}{d^1_{k,k+1}}$ for $i > 0$. Then $\Upsilon D_2 \Upsilon^{-1} - D_1$ is lower triangular, and if we assume that

$$\prod_{k=0}^{\infty} \left| \frac{d^2_{k,k+1}}{d^1_{k,k+1}} \right| < \infty, \tag{2.10}$$

then Υ is bounded.

Taking into account this remark, the hypothesis in last theorem that Δ is lower triangular can be relaxed to Δ such that $D - \Delta$ is still a Hessenberg matrix. In order to see this we take $D_1 = D$, $D_2 = D - \Delta$ and Υ as before. Since $\Delta \in \mathcal{F}_k$ for some k, the condition (2.10) clearly holds, so D_2 is similar to a perturbation of D_1 by a lower triangular matrix Δ' of finite rank. Now we can apply Theorem 2.2.

In the same way, some of the results that follow, even though they are stated for perturbations by a lower triangular matrix, can be extended to Δ such that D and $D - \Delta$ are both Hessenberg matrices.

Next we will extend Theorem 2.2 to include perturbations by a Δ of infinite rank.

2.3. THE ASSOCIATED POLYNOMIALS OF k-TH KIND. A Hessenberg matrix D can always be written as $D = S^t V$, where V is a lower triangular matrix such that $V_{0,0} = 1$ and $S = (\delta_{i+1,j})_{i,j}$ is the unilateral shift. The matrix V has nonzero elements on the main diagonal; thus it is an invertible matrix. Let us define $\widetilde{D}(0) = V^{-1}$ and

$$\widetilde{D}(z) = (I + (\widetilde{D}(0)S)z + (\widetilde{D}(0)S)^2 z^2 + \cdots)\widetilde{D}(0). \tag{2.11}$$

Notice that, since the first n rows of $(\widetilde{D}(0)S)^n$ are zero, the entries of $\widetilde{D}(z)$ are well defined. The matrix $\widetilde{D}(z)$ plays an important role in several questions related to D; we refer to [7] for a more thorough treatment of the properties of this matrix. We have

$$(D - zI)\widetilde{D}(z) = S^t, \quad \widetilde{D}(z)S(D - zI) = I - \mathbf{p}(z)e_0^t,$$

where $\mathbf{p}(z)$ is the first column of $\widetilde{D}(z)$.

We denote the entries of $\widetilde{D}(z)$ by $(\widetilde{D}(z))_{n,k} = p^k_{n-k}(z)$. These are the so called *associated polynomials of k-th kind* in the theory of orthogonal polynomials.

They satisfy the recurrence relations, $p_0^k(z) = \frac{1}{d_{k-1,k}}$,

$$zp_{n-k}^k(z) = \sum_{i=0}^{n+1} d_{n,i} p_{i-k}^k(z). \tag{2.12}$$

The next lemma shows the relationship between the polynomials $p_{n-k}^k(z)$ and the functions $\Sigma(\cdot)$ and $\psi(\cdot)$ that were constructed in Subsection 2.2.

LEMMA 2.3. *We have*

$$\Sigma(\phi e_n^t) = (p_{n-1}^1(D)\phi, \dots, p_0^n(D)\phi, 0, 0, \dots), \tag{2.13}$$

$$\psi(\phi e_n^t) = p_n(D)\phi. \tag{2.14}$$

Notice that $(p_n(z), p_{n-1}^1(z), \dots, p_0^n(z), 0, 0, \dots)$ is the n-th row of $\widetilde{D}(z)$.

Proof. Recall from (2.8) that $\Sigma(\phi e_n^t) = X_n^\phi$ and $\psi(\phi e_n^t) = \psi_n^\phi$. It is enough to check that the right sides of (2.13) and (2.14) satisfy the recurrence relations used in Lemma 2.1 to define X_n^ϕ and ψ_n^ϕ. This follows from the recurrence relation (2.12). ∎

2.4. EXTENSION OF THEOREM 2.2. From this point on we assume that D represents a bounded operator on l_2.

Let $\{a_i\}_{i=0}^\infty$ be a sequence of positive numbers. For $X \in B(l_2)$ define

$$|X|_{\mathcal{F}} = \Big(\sum_{i=0}^{\infty} a_i \|Xe_i\|^2\Big)^{1/2}, \quad \mathcal{F} = \{X \in B(l_2) : |X|_{\mathcal{F}} < \infty\}.$$

LEMMA 2.4. *$\mathcal{F}$ is a left ideal of $B(l_2)$ and $|BX|_{\mathcal{F}} \leqslant \|B\||X|_{\mathcal{F}}$.*

Proof. $|BX|_{\mathcal{F}}^2 = \sum a_i \|BXe_i\|^2 \leqslant \|B\|^2 (\sum a_i \|Xe_i\|^2)$. ∎

Let us take $a_n = \sum\limits_{i=0}^{n} \|p_{n-i}^i(D)\|^2$ and define $|\cdot|_{\mathcal{F}}$ and $\mathcal{F}$ accordingly. We have the estimate

$$\|\psi(\phi e_n^t)\|^2 + \|\Sigma(\phi e_n^t)\|_2^2 \leqslant \|\phi\|^2 \sum_{i=0}^{n} \|p_{n-i}^i(D)\|^2 \leqslant a_n \|\phi\|^2. \tag{2.15}$$

Here $\|\cdot\|_2$ is the Hilbert-Schmidt norm (the sum of the squares of the absolute values of the entries). This estimate shows that we can extend $\Sigma(\cdot)$ and $\psi(\cdot)$ by continuity to $\mathcal{F}$ (recall that they were originally defined in $\bigcup\limits_{n=0}^{\infty} \mathcal{F}_n$). We have

$$\|\psi(A)\|^2 + \|\Sigma(A)\|_2^2 \leqslant |A|_{\mathcal{F}}^2$$

for all $A \in \mathcal{F}$.

For the functions $\Sigma(\cdot)$ and $\psi(\cdot)$ defined on $\mathcal{F}$ we have:

(i) $\Sigma(\cdot)$ and $\psi(\cdot)$ satisfy (2.3);
(ii) $\Sigma(\mathcal{F}) \subset B_2(l_2)$ ($B_2(l_2)$ denotes the space of Hilbert-Schmidt operators);
(iii) if A is lower triangular, then $\Sigma(A)$ is strictly lower triangular.

Let us write $|f|_{\infty,M} = \sup_{|x|\leqslant M} |f(x)|$. Taking into account that one has the inequality $\|p(D)\| \leqslant |p|_{\infty,\|D\|}$, the coefficients a_n in the definition of $|\cdot|_{\mathcal{F}}$ can be replaced by the more computable numbers $b_n = \sum_i |p^i_{n-i}|^2_{\infty,\|D\|}$. However, in many cases b_n has geometric growth, while a_n has polynomial growth (e.g., the Chebyshev matrix).

Now we are ready to state the extension of Theorem 2.2.

THEOREM 2.5. *Let D be a bounded Hessenberg matrix and Δ be a member of $\mathcal{F}$ that satisfies one of the following two conditions:*

(i) $|\Delta|_{\mathcal{F}} < \frac{1}{2}$;

(ii) *Δ is lower triangular.*

Then there are $\psi, v \in l_2$ and $U = I + X$ with X of Hilbert-Schmidt class such that $U(D-\Delta)U^{-1} = D - \psi v^t$. In case (ii) *we can take X to be strictly lower triangular and $v = e_0^t$.*

Proof. (i) Take $U = U(\Delta)$ as in (2.2). Notice that if Δ is lower triangular then $U(\Delta)$ is lower triangular too.

(ii) Let us write $\Delta = \Delta_1 + \Delta_2$, with Δ_2 lower triangular of finite rank and $|\Delta_1|_{\mathcal{F}} < \frac{1}{2}$. By (i) there is U_1 lower triangular such that

$$U_1(D - \Delta_1 - \Delta_2)U_1^{-1} = D - \psi e_0^t - U_1 \Delta_2 U_1^{-1}.$$

The right side is a finite rank perturbation of D. Thus, an application of Theorem 2.2 completes the proof. ∎

We can improve (ii) of the previous theorem. Let $\mathcal{TF}$ denote the subspace of the lower triangular elements of $\mathcal{F}$. For every n we write $\mathcal{TF} = \mathcal{TF}_n \oplus \mathcal{TF}^n$, where $\mathcal{TF}_n = \{X \in \mathcal{TF} : Xe_i = 0, i > n\}$ and $\mathcal{TF}^n = \{X \in \mathcal{TF} : Xe_i = 0, i \leqslant n\}$.

PROPOSITION 2.6. *Formula* (2.2) *is well defined for every $\Delta \in \mathcal{TF}$ and*

$$U(\cdot) : \mathcal{TF}, |\cdot|_{\mathcal{F}} \to B(l_2), \|\cdot\|$$

is a continuous map.

Proof. For $A \in \mathcal{TF}$ we define the linear map $T_A : B(l_2) \to B(l_2)$ by $T_A(X) = \Sigma(XA)$. Notice that formula (2.2) can be written as $U(\Delta) = \sum_{i=0}^{\infty} T^i_\Delta(1)$.

First we prove that $U(\cdot)$ is continuous on the unit ball of $\mathcal{TF}$. Take $A, B \in \mathcal{TF}$ such that $|A|_{\mathcal{F}} = a < 1$, $|B|_{\mathcal{F}} = b < 1$ and $|A - B|_{\mathcal{F}} < \varepsilon$. We have,

$$T^k_A(1) - T^k_B(1) = \Sigma((T^{k-1}_A(1) - T^{k-1}_B(1))A) + \Sigma(T^{k-1}_B(1)(A - B)).$$

From this we get by induction that $\|T^k_A(1) - T^k_B(1)\| < \varepsilon \frac{a^k - b^k}{a-b}$. This implies that $\|U(A) - U(B)\| < \varepsilon \frac{1}{(1-a)(1-b)}$. Hence $U(\cdot)$ is continuous on the unit ball of $\mathcal{TF}$.

Next we prove that the sum on the right side of (2.2) is well defined for all $\Delta \in \mathcal{TF}$. Let $\Delta = \Delta_1 + \Delta_2$ with $\Delta_1 \in \mathcal{TF}^n$, $|\Delta_1|_{\mathcal{F}} < 1$ and $\Delta_2 \in \mathcal{TF}_n$. We notice that if A and X are lower triangular, then by property 2 in the proof of Theorem 2.2 we get $T_A(X\Delta_2) \in \mathcal{TF}_{n-1}$. Hence we have $T_A^n(X\Delta_2) = 0$.

Let us prove by induction that $T_\Delta^k(1) = T_\Delta^n T_{\Delta_1}^{k-n}(1)$, for $k \geqslant n$. This is clearly true for $k = n$. Assuming it is true for k, we get

$$\begin{aligned} T_\Delta^{k+1}(1) &= T_\Delta^{n+1} T_{\Delta_1}^{k-n}(1) = T_\Delta^n T_\Delta(T_{\Delta_1}^{k-n}(1)) = T_\Delta^n(\Sigma(T_{\Delta_1}^{k-n}(1)(\Delta_1 + \Delta_2))) \\ &= T_\Delta^n T_{\Delta_1}^{k+1-n}(1) + T_\Delta^n(\Sigma(T_{\Delta_1}^{k-n}(1)\Delta_2)) = T_\Delta^n T_{\Delta_1}^{k+1-n}(1). \end{aligned}$$

This completes the induction.

Adding $T_\Delta^k(1) = T_\Delta^n T_{\Delta_1}^{k-n}(1)$ for all k we get $U(\Delta) = \sum_{i=0}^{n-1} T_\Delta^i(1) + T_\Delta^n(U(\Delta_1))$. Hence $U(\Delta)$ is well defined.

Finally we prove the continuity of $U(\cdot)$ on all $\Delta \in \mathcal{TF}$. Let $\Delta \in \mathcal{TF}$ and take $\delta = \delta(\Delta, n) > 0$ with the property that for every Δ' such that $|\Delta - \Delta'|_{\mathcal{F}} < \delta$, it is true that if $\Delta' = \Delta'_1 + \Delta'_2$ with $\Delta'_1 \in \mathcal{TF}^n$ and $\Delta'_2 \in \mathcal{TF}_n$, then $|\Delta'_1|_{\mathcal{F}} < 1$. That this is always possible follows from the continuity of the projection of $\mathcal{TF}$ onto $\mathcal{TF}^n$. Then

$$U(\Delta') = \sum_{i=0}^{n-1} T_{\Delta'}^i(1) + T_\Delta^n(U(\Delta'_1))$$

for all Δ' in the neighborhood of Δ of radius δ. The right side is continuous on a neighborhood of Δ, hence $U(\Delta)$ is continuous. ∎

3. PERTURBATIONS OF THE SHIFT AND THE CHEBYSHEV MATRIX

In this section we study perturbations of S^t and the Chebyshev matrix $J = \frac{1}{2}(S + S^t)$, where $S = (\delta_{i,j+1})_{i,j}$ denotes the unilateral shift. In some cases the perturbing matrix Δ is assumed to be lower triangular, but taking into account the remark at the end of Subsection 2.2, these results can be extended to perturbations by a matrix Δ of Hessenberg form.

3.1. THEOREM 2.5 APPLIED TO S^t AND J. *The shift.* The similarity of perturbations of the shift to the shift itself has been considered in papers [2], [5], [8] and [1]. In [2] J.S. Freeman proved that if Δ is a strictly lower triangular matrix satisfying $\sum_{i,j} |\Delta_{i,j}| < \infty$ and $\Delta_{i+1,i} \neq -1$, then the $S + \Delta$ is similar to S. O.I. Soibelman and B. Chan proved independently that $S + \Delta$ is similar to S when Δ is an infinite matrix that satisfies $\sum_{i,j} |\Delta_{i,j}| < 1$.

The matrix S^t is Hessenberg, so the results of the previous section apply to it. Using (2.12) we compute the associated polynomials of k-th kind. We obtain

$p^k_{n-k}(z) = z^{n-k}$, and from this we get $a_n = n$ for the coefficients in the definition of $|\cdot|_{\mathcal{F}}$. Hence, by Theorem 2.5, we get that if Δ satisfies one of the two conditions:

(I) $\sum\limits_j j\|\Delta e_j\|^2 < \frac{1}{4}$;

(II) Δ is lower triangular and $\sum\limits_j j\|\Delta e_j\|^2 < \infty$;

then $S^t - \Delta$ is similar to $S^t - \phi v^t$ for some $\phi, v \in l_2$. In case (II) we can take $v = e_0$.

The following theorem is proven in [1]:

Let Δ be such that $\sum\limits_{i,j} |\Delta_{i,j}| < 1$. Then there is a bounded operator U bounded such that $U(S^t - \Delta)U^{-1} = S^t$. Moreover, if Δ is lower triangular then U can be taken to be lower triangular too.

Combining Theorem 2.2 with this assertion we get that if:

(III) Δ is lower triangular and $\sum\limits_{j=N}^{\infty}\sum\limits_{i=0}^{\infty} |\Delta_{i,j}| < \infty$ for some $N > 0$,

then $S^t - \Delta$ is similar to $S^t - \phi e_0^t$ for some $\phi \in l_2$.

REMARKS. (i) Taking adjoints, these results can be restated for perturbations of S.

(ii) There are rank one perturbations of S^t of arbitrarily small norm that are not similar to S^t.

The Chebyshev matrix. Now we consider perturbations of $J = \frac{S+S^t}{2}$. A computation shows that the associated polynomials of k-th kind are $p^k_{n-k}(z) = T_{n-k}(z)$, where $\{T_n\}_{n=0}^{\infty}$ are the Chebyshev polynomials. The Chebyshev polynomials are defined by $T_0 = 1$ and the recurrence

$$2zT_n(z) = T_{n-1}(z) + T_{n+1}(z).$$

We have ([6])

$$\|T_n(J)\| = \sup_{x\in[-1,1]} |T_n(x)| = 1.$$

Thus, we can take $a_n = n$ in the definition of $|\cdot|_{\mathcal{F}}$. Hence, by Theorem 2.5, we see that if Δ satisfies conditions (I) or (II) stated above, then $J - \Delta$ is similar to $J - \phi v^t$ for some $\phi, v \in l_2$. In case (II) we can take $v = e_0$.

3.2. PERTURBATION DETERMINANTS. Let us denote by $R(z, A) = (zI - A)^{-1}$ the resolvent function of the bounded operator A. If A' is a bounded operator such that $A - A'$ is of trace class, then the analytic function

$$\det(I + (A - A')R(z, A))$$

is called a *perturbation determinant*. A reference for several properties of perturbation determinants is [4].

In this subsection we discuss the link between perturbation determinants and the similarity of perturbations of S^t and J.

We start with the shift. Let Δ be lower triangular and of trace class (hence bounded). Let us write $D = S^t - \Delta$. Let us define the function

$$\psi_{-1}(z) = \det(I + \Delta R(z^{-1}, S^t)).$$

This function is analytic in the interior of the unit circle. It is known as Szego's function in the theory of orthogonal polynomials. (One has to allow that Δ be Hessenberg, and not just lower triangular, to make $S^t - \Delta$ the Hessenberg matrix of orthogonal polynomials in the unit circle. See [7].)

Next we define the operator-valued function

$$\Psi(z) = \psi_{-1}(z) R(z^{-1}, D).$$

An application of Carleman's inequality shows that this function is analytic in the interior of the unit circle (see [7]). Let us denote by $\psi_n(z) = (\Psi(z))_{0,n}$, for $n = 0, 1 \ldots$, the entries of the first row of $\Psi(z)$. The following formula is proven in Theorem 8 of [7]:

$$\psi_n(z) = z^{n+1} + \sum_{j=0}^{\infty} \Big(\sum_{i=j}^{\infty} \Delta_{i,j} z^{i+1} \Big) p^{n+1}_{j-n-1}(z), \tag{3.1}$$

for $n = -1, 0, 1, \ldots$. The right side of (3.1) is a series converging uniformly in compact sets of the open unit disk. From this formula we see that $\psi_n(z) = z^{n+1} + O(z^{n+2})$ for $n \geqslant -1$. Moreover, $\psi_n(z)$ should be close to z^{n+1} if we make the perturbation small in a suitable sense.

Let us write $\psi_j(z) = \sum_{i=j}^{\infty} \psi_{i,j} z^{i+1}$ for $j = -1, 0, 1, \ldots$. Let H_2 denote the Hardy space of the unit circle. The following theorem establishes a simple relationship between the functions $U(\cdot)$ and $\psi(\cdot)$ constructed in Section 2 and the coefficients $\psi_{i,j}$ of the functions $\psi_j(z)$.

THEOREM 3.1. *If Δ is lower triangular of trace class satisfying condition* (II), *then* $U(\Delta) = (\psi_{i,j})_{i,j=0}^{\infty}$ *and* $\psi(U(\Delta)\Delta) = (\psi_{i,-1})_{i=0}^{\infty}$.

Notice that by Theorem 2.5 and Proposition 2.6 one has $U = U(\Delta)$ and $\psi = \psi(U(\Delta)\Delta)$ are precisely the matrix and the vector implementing the similarity between $S^t - \Delta$ and $S^t - \psi e_0^t$.

Proof. Let us write $\widetilde{U} = (\psi_{i,j})_{i,j=0}^{\infty}$ and $\widetilde{\psi} = (\psi_{i,-1})_{i=-1}^{\infty}$ ($\widetilde{\psi}$ is a column vector).

First we assume that $\Delta \in \mathcal{TF}_k$ for some k. Let $\widetilde{U} = I + \widetilde{X}$. Using (3.1) we can prove that:

1. $\widetilde{X}$ is strictly lower triangular (because $\psi_n(z) = z^{n+1} + O(z^{n+2})$).
2. $\widetilde{X} \in \mathcal{TF}_{k-1}$ (because $\psi_n(z) = z^{n+1}$ for $n \geqslant k$).
3. The columns of $\widetilde{X}$ are in l_2 (because $\sum_i \Delta_{i,j} z^{i+1} \in H_2$ and $p^{n+1}_{j-n-1}(z)$ is a polynomial. Hence $\psi_n(z) \in H_2$ for $n \geqslant -1$).

The definition of $\Psi(z)$ implies that $(\psi_0(z), \psi_1(z), \dots)(z^{-1}I - D) = \psi_{-1}(z)e_0^t$, which can be written as $(z, z^2, \dots)\widetilde{U}(z^{-1}I - D) = (1, z, \dots)\widetilde{\psi}e_0^t$, and then we get $(1, z, \dots)(\widetilde{U} - S\widetilde{U}D) = (1, z, \dots)\widetilde{\psi}e_0^t$. Thus $\widetilde{U} - S\widetilde{U}D = \widetilde{\psi}e_0^t$. Multiplying by S^t on the left and by $\widetilde{U}^{-1}$ on the right we get

$$\widetilde{U}D\widetilde{U}^{-1} = S^t - (S^t\widetilde{\psi})e_0^t. \tag{3.2}$$

Now recall that, by Theorem 2.2, there are unique U and ψ satisfying (3.2) and properties 1 and 2 listed above. This proves the theorem when $\Delta \in \mathcal{TF}_k$.

Next we assume that $\Delta \in \mathcal{TF}$. Take $\Delta^{(n)} \in \mathcal{TF}_n$ such that $|\Delta - \Delta^{(n)}|_{\mathcal{F}} \to 0$. We know by Proposition 2.6 that $U(\Delta^{(n)})$ and $\psi(U(\Delta^{(n)})\Delta^{(n)})$ converge in the norm topology to $U(\Delta)$ and $\psi(U(\Delta)\Delta)$ respectively. On the other hand, by the definition of $\psi_{-1}(z)$ and $\Psi(z)$, and the continuity of the determinant and the resolvent function, we have $\psi_{-1}^{(n)}(z) \to \psi_{-1}(z)$ and $\Psi^{(n)}(z) \to \Psi(z)$ uniformly on compact sets of the unit circle. So $\psi_{i,j}^n \to \psi_{i,j}$ for $i = 0, 1, \dots, j = -1, 0, 1, \dots$. Therefore $\widetilde{U} = U(\Delta)$ and $S^t\widetilde{\psi} = \psi(U(\Delta)\Delta)$ for every $\Delta \in \mathcal{TF}$ of trace class. ∎

The arguments in the case of the Chebyshev matrix are analogous to the case of the shift, so we will just give a succinct presentation. Let $J = \frac{S+S^t}{2}$, Δ be a lower triangular matrix of trace class and $D = J - \Delta$. Define the function (sometimes called Jost function)

$$\psi_{-1}(z) = \det\Big(I + \Delta R\Big(\frac{z+z^{-1}}{2}, J\Big)\Big),$$

and the operator valued function

$$\Psi(z) = \psi_{-1}(z) R\Big(\frac{z+z^{-1}}{2}, D\Big).$$

The functions $\psi_{-1}(z)$ and $\Psi(z)$ are analytic in the interior of the unit circle. If $\psi_n(z)$, $n = 0, 1, \dots$ denotes the elements in the first row of $\Psi(z)$, then (Theorem 8 of [7])

$$\psi_n(z) = z^{n+1} + \sum_{j=0}^{\infty}\Big(\sum_{i=j}^{\infty}\Delta_{i,j}z^{i+1}\Big)p_{j-n-1}^{n+1}\Big(\frac{z+z^{-1}}{2}\Big), \tag{3.3}$$

for $n = -1, 0, 1, \dots$. Again let us $\psi_j(z) = \sum\limits_{i=j}^{\infty}\psi_{i,j}z^{i+1}$ for $j = -1, 0, 1, \dots$.

THEOREM 3.2. *If Δ is lower triangular of trace class satisfying condition* (II)*, then $U(\Delta) = (\psi_{i,j})_{i,j=0}^{\infty}$ and $\psi(U(\Delta)\Delta) = (\psi_{i,-1})_{i=0}^{\infty}$.*

Proof. Suppose that $\Delta \in \mathcal{TF}_k$ for some k. Define $\widetilde{U}$ and $\widetilde{\psi}$ as in the proof of Theorem 3.1. Using (3.3), again we see that $\widetilde{U}$ satisfies properties 1, 2 and 3 listed in the proof of Theorem 3.1 (the computation must be more careful now, since $p_{j-n-1}^{n+1}(\frac{z+z^{-1}}{2})$ is not a polynomial. Nevertheless, $z^{j-n-1}p_{j-n-1}^{n+1}(\frac{z+z^{-1}}{2})$ is a polynomial). The definition of $\Psi(z)$ implies that

$$
\begin{aligned}
(\psi_0(z), \psi_1(z), \dots)\Big(\frac{z+z^{-1}}{2}I - D\Big) &= \psi_{-1}(z)e_0^t \\
(z, z^2, \dots)\widetilde{U}\Big(\frac{z+z^{-1}}{2}I - D\Big) &= (1, z, \dots)\widetilde{\psi}e_0^t \\
(1, z, \dots)\Big(\frac{I+S^2}{2}\widetilde{U} - S\widetilde{U}D\Big) &= (1, z, \dots)\widetilde{\psi}e_0^t.
\end{aligned}
$$

Thus $\frac{I+S^2}{2}\widetilde{U} - S\widetilde{U}D = \widetilde{\psi}e_0^t$. Multiplying by S^t on the left and by $\widetilde{U}^{-1}$ on the right we get

$$
\widetilde{U}D\widetilde{U}^{-1} = J - (S^t\widetilde{\psi})e_0^t. \tag{3.4}
$$

Now we invoke the uniqueness of U and ψ satisfying (3.4) and the properties 1 and 2 listed in the proof of Theorem 3.1. This proves the theorem when $\Delta \in \mathcal{TF}_k$.

In order to finish the proof, we use the same continuity argument that was used in Theorem 3.1. ∎

Acknowledgements. I am grateful to Peter Rosenthal for encouraging me to write this paper and for suggesting many corrections to one of its early drafts.

REFERENCES

[1] B. CHAN, Perturbations of shifts, *J. Math. Anal. Appl.* **88**(1982), 572–590.
[2] J.M. FREEMAN, Perturbations of the shift operator, *Trans. Amer. Math. Soc.* **114**(1965), 251–260.
[3] K.O. FRIEDRICHS, *Perturbation of Spectra in Hilbert Space*, Lectures in Appl. Math., vol. 3, Amer. Math. Soc., Providence, RI 1965.
[4] I.C. GOHBERG, M.G. KREIN, *Introduction to the Theory of Linear Non-Selfadjoint Operators*, Amer. Math. Soc., Providence RI 1969.
[5] S.-C. LIN, Perturbations of the unilateral shift, *Bull. Amer. Math. Soc.* **77**(1971), 621–624.
[6] T.J. RIVLIN, *The Chebyshev Polynomials*, New York, Wiley 1974.
[7] L. ROBERT, L. SANTIAGO, Finite sections method for Hessenberg matrices, *J. Approx. Theory* **123**(2003), 69–88.
[8] O.I. SOIBELMAN, Perturbations of the shift operator [Russian], *Mat. Issled.* **8**(1973), 76–91, 181.

LEONEL ROBERT, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, 100 ST. GEORGE STREET, ROOM 4072, TORONTO, ON, CANADA M5S 3G3
E-mail address: lrobert@math.utoronto.ca

Received September 5, 2003.