DEPARTMENT OF MATHEMATICS

TECHNICAL REPORT

ON THE INVERTIBILITY OF SOME OPERATORS ON HILBERT SPACES

Ramesh V. Garimella

Volodymyr Hrynkiv

Yung-Way Liu

June 2000

No. 2000-3

TENNESSEE TECHNOLOGICAL UNIVERSITY

Cookeville, TN 38505

ON THE INVERTIBILITY OF SOME OPERATORS ON HILBERT SPACES

Ramesh V. Garimella *, Volodymyr Hrynkiv**, and Yung-Way Liu***

Box 5054, Department of Mathematics, Tennessee Technological University, Cookeville, TN 38505 USA

For any given bounded linear operator A on a complex Hilbert space H, we give sufficient conditions to ensure the existence of a bounded operator B on H such that $(i) AB + BA$ is of rank one, and $(ii) I + e^{xP(A)+tQ(A)}B$ is invertible for all $x, t \in \mathsf{R}$ where $P(A)$ and $Q(A)$ are polynomials in A. Our main results will provide a justification in general terms to a crucial step of the so-called operator method used by Aden, Carl, and Schiebold [1,3] to solve nonlinear partial differential equations like the Korteveg-deVries(KdV), modified KdV, Kadomtsev-Petviashvili equations.

1. INTRODUCTION

In [1] Aden and Carl showed that for a given bounded linear operator A on a Banach space E the family of operators $V(x,t) := (I+L)^{-1}e^{Ax+A^3t}(AB+BA)$ is a solution to the operator KdV equation $V_t = V_{xxx} + 3V_x^2$, provided the operator B satisfies (i) $AB + BA$ is of rank one, and (ii) $(I + L)^{-1}$ exists, where $L(x,t) := e^{Ax+A^3t}B$. Further, $v(x,t) := tr(V(x,t)))$, where tr is the continuous trace, gives a scalar solution to the scalar KdV equation $v_t = v_{xxx} + 3v_x^2$. A similar approach was used by Carl and Schiebold in [4] to solve some other partial differential equations of nonlinear physics like the modified KdV

* e-mail:rgarimella@tntech.edu
** e-mail:vh6884@tntech.edu
*** e-mail:ywliu@tntech.edu

equation, Kadomtsev-Petviashvili equation as well as the sine-Gordon equation. The approach mentioned above is known as the operator method. The main idea of the operator method can be described as follows. Given a nonlinear PDE of soliton physics as well as a specific scalar solution to the equation, the first step in the solution is to translate the given nonlinear equation to an operator equation. Using the specific scalar solution as an aid, one then searches for a family of operator solutions to the operator equation. Having obtained the operator solutions, the second step is to transfer the operator-valued solution into a scalar solution by using a suitable scalarization technique. The most important step in the above method is step one. In order to carry out this step, in most cases, given an operator A on a Hilbert space H, one needs to find an operator B such that

(1) $AB + BA$ is of rank one, and

(2) the operator $I + L$ is invertible, where $L := I + e^{xP(A)+tQ(A)}B$, $x, t \in \mathbb{R}$ and $P(A), Q(A)$ are some polynomials in A with real (or complex) coefficients.

In this paper we give sufficient conditions to ensure the existence of an operator B satisfying the above conditions (1) and (2) for a given operator A on a separable Hilbert space H. In section 3 we consider operators A acting on a finite dimensional space, whereas section 4 deals with operators A on the infinite dimensional sequence space ℓ_2.

2. PRELIMINARIES

Recall that an operator $T : E \to F$ (where E and F are Banach spaces) is said to be of rank one if the dimension of the range of T is equal to one. It is straightforward to verify that T is of rank one if and only if there exists $a \in E'$ (dual of E) and $y \in F$ such that $T = a \otimes y$, where $(a \otimes y)x := a(x)y, \ \ \forall x \in E$. It is obvious that for any $a, b \in E'$; $x, y \in E$, and complex number λ, we have (i) $\lambda(a\otimes x) = a\otimes \lambda x$, (ii) $(a\pm b)\otimes x = a\otimes x \pm b\otimes x$, and (iii) $(a\otimes x)\circ(b\otimes y) = b \otimes a(y)x$.

The following lemmas are quite useful in the proofs of the main results of the paper. Lemma 2.1 is also stated in [3] and a proof of Lemma 2.2 can be found in [1].

Lemma 2.1. *Suppose $R \in L(E, F)$ and $S \in L(F, E)$ where E and F are vector spaces. Then $I + RS$ is invertible if and only if $I + SR$ is invertible.*

Lemma 2.2. *Let E, F be Banach spaces, and let $A \in \mathcal{L}(E, E)$, $B \in \mathcal{L}(F, F)$ be bounded operators. If $\int_0^\infty \|e^{-Bs}Ce^{-As}\|ds < \infty$ then the operator X defined by the integral $X = \int_0^\infty e^{-Bs}Ce^{-As}ds$ satisfies the equation $XA + BX = C$.*

Throughout the paper $\mathbb{C}$ denotes the field of complex numbers, ℓ_2 stands for the Hilbert space of all absolute square summable complex sequences with the standard ℓ_2-norm, and $L_2(0,\infty)$ denotes the space of square integrable measurable functions on $(0,\infty)$. For any $x = (x_1, \dots, x_n) \in \mathbb{C}^n$ or $x = (x_n) \in \ell_2$,

x^{t} always denotes the corresponding column vector. For any square matrix A, $\mathcal{E}(A)$ denotes the set of eigenvalues of A. Finally, I always stands for the identity operator.

3. FINITE DIMENSIONAL CASE

Recall that for any $h, g \in \mathsf{C}^{\mathsf{n}}$, $h \otimes g$ gives a linear operator on C^{n} which is defined as follows $(h \otimes g)x = \left(\sum_{\mathsf{i}=1}^{\mathsf{n}} \bar{h}_{\mathsf{i}} x_{\mathsf{i}}\right) g$ for each $x \in \mathsf{C}^{\mathsf{n}}$. Even though for any $h, g \in \mathsf{C}^{\mathsf{n}}$ there always exists a matrix B such that $AB + BA = h \otimes g$ provided $0 \notin \mathcal{E}(A) + \mathcal{E}(A)$ for a given A (see [1], [4], [6]), we show that a careful choice of h and g will also ensure the invertibility of the matrix $I + e^{\mathsf{x}\mathsf{P}(\mathsf{A}) + \mathsf{t}\mathsf{Q}(\mathsf{A})} B$ for all $x, t \in \mathsf{R}$, where $P(A)$ and $Q(A)$ are some polynomials in A. First we prove the result for any diagonalizable matrix of size n and then extend it to a general square matrix. We call a matrix diagonalizable if it is similar to a diagonal matrix.

Theorem 3.1. *Let A be a diagonalizable square matrix of size n such that $0 \notin \mathcal{E}(A) + \mathcal{E}(A)$. Then there always exist non-zero vectors $h, g \in \mathsf{C}^{\mathsf{n}}$ and an $n \times n$ matrix B such that*

(i) $AB + BA = h \otimes g$, and

(ii) $I + e^{\mathsf{x}\mathsf{P}(\mathsf{A}) + \mathsf{t}\mathsf{Q}(\mathsf{A})} B$ is invertible for all $x, t \in \mathsf{R}$ where $P(A)$ and $Q(A)$ are polynomials in A.

Proof. Let S be a similarity transformation which diagonalizes A, i.e., there exists a diagonal matrix $\tilde{A}$ such that $\tilde{A} = SAS^{-1}$. Let $\{e_j\}$ be the standard basis, and let $\tilde{A}e_j = \mu_j e_j$, $1 \le j \le n$. Define the following matrix $\tilde{B} := (b_{ij})_{i,j=1}^n$ where $b_{ij} = \dfrac{\beta_i \bar{\alpha}_j}{\mu_i + \mu_j}$, $\alpha = (\alpha_1, \dots, \alpha_n)^t$, and $\beta = (\beta_1, \dots, \beta_n)^t$. Note that the entries (b_{ij}) of the matrix $\tilde{B}$ are well-defined since $0 \notin \mathcal{E}(A) + \mathcal{E}(A)$.

Claims: (1) $\tilde{A}\tilde{B} + \tilde{B}\tilde{A} = \alpha \otimes \beta$.

(2) The following choice of α and β guarantees the invertibility for the matrix $(I + e^{xP(\tilde{A})+tQ(\tilde{A})}\tilde{B})$ for all $x, t \in \mathsf{R}$: for $i < n$, $\alpha_i = 0$ if i is odd, $\alpha_i \ne 0$ if i is even; $\beta_i \ne 0$ if i is odd, $\beta_i = 0$ if i is even, and $\dfrac{\beta_n \bar{\alpha}_n}{2\mu_n} \ge 0$.

Proof of claim (1). It is enough to show that for each j, $1 \le j \le n$ $(\tilde{A}\tilde{B} + \tilde{B}\tilde{A})e_j = (\alpha \otimes \beta)e_j$. We have $(\tilde{A}\tilde{B} + \tilde{B}\tilde{A})e_j = \tilde{A}\tilde{B}e_j + \tilde{B}(\mu_j e_j) = \tilde{A}(b_{1j}, \dots, b_{nj})^t + \mu_j(b_{1j}, \dots, b_{nj})^t = (\mu_1 b_{1j}, \dots, \mu_n b_{nj})^t + \mu_j(b_{1j}, \dots, b_{nj})^t = ((\mu_1 + \mu_j)b_{1j}, \dots, (\mu_n + \mu_j)b_{nj})^t = (\beta_1\bar{\alpha}_j, \dots, \beta_n\bar{\alpha}_j)^t = (\alpha \otimes \beta)e_j$.

This proves claim (1).

To prove claim (2), note that because of the choice of α and β the matrix $\tilde{B}$ will have the following structure

$$\tilde{B} = \begin{pmatrix} 0 & b_{12} & 0 & b_{14} & 0 & \dots & b_{1n} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & b_{32} & 0 & b_{34} & 0 & \dots & b_{3n} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & b_{52} & 0 & b_{54} & 0 & \dots & b_{5n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \\ 0 & b_{n2} & 0 & b_{n4} & 0 & \dots & b_{nn} \end{pmatrix}.$$

Hence $\det(I + e^{xP(\tilde{A})+tQ(\tilde{A})}\tilde{B}) = 1 + b_{nn}e^{xP(\mu_n)+tQ(\mu_n)}$. Since $b_{nn} = \dfrac{\beta_n\bar{\alpha}_n}{2\mu_n} \ge 0$, it is easy to see that $1 + b_{nn}e^{xP(\mu_n)+tQ(\mu_n)} \ne 0$. This proves claim (2).

Now we are ready to exhibit $h, g \in \mathbb{C}^n$ and an $n \times n$ matrix B as stated in the theorem. Let $B := S^{-1}\tilde{B}S$,

$h := \alpha_2(\bar{s}_{21}, \ldots, \bar{s}_{2n})^t + \alpha_4(\bar{s}_{41}, \ldots, \bar{s}_{4n})^t + \ldots + \alpha_n(\bar{s}_{n1}, \ldots, \bar{s}_{nn})^t$ and

$g := \beta_1(q_{11}, \ldots, q_{n1})^t + \beta_3(q_{13}, \ldots, q_{n3})^t + \ldots + \beta_n(q_{1n}, \ldots, q_{nn})^t$

where $S = (s_{ij})$ and $S^{-1} = (q_{ij})$. Notice also that $S^{-1}(\alpha \otimes \beta)S = h \otimes g$. Then

$$AB + BA = S^{-1}\tilde{A}SS^{-1}\tilde{B}S + S^{-1}\tilde{B}SS^{-1}\tilde{A}S = S^{-1}(\tilde{A}\tilde{B} + \tilde{B}\tilde{A})S = h \otimes g.$$

This proves part (i) of the theorem. Since

$$I + e^{xP(A)+tQ(A)}B = S(I + e^{xP(\tilde{A})+tQ(\tilde{A})}\tilde{B})S^{-1}$$

and $I + e^{xP(\tilde{A})+tQ(\tilde{A})}\tilde{B}$ is invertible, statement (ii) of the theorem follows. $\square$

Now we are ready to extend the above theorem to a general situation.

Theorem 3.2. *Let A be any square matrix of size n such that $0 \notin \mathcal{E}(A)+\mathcal{E}(A)$. Then there always exist non-zero vectors $h, g \in \mathbb{C}^n$ and an $n \times n$ matrix B such that*

(i) $AB + BA = h \otimes g$, and

(ii) $I + e^{xP(A)+tQ(A)}B$ is invertible for all $x, t \in \mathbb{R}$ where $P(A)$ and $Q(A)$ are polynomials in A.

Proof. Let S be such a similarity transformation which reduces the matrix A to a Jordan canonical form in which all the 1×1 Jordan blocks appear at the bottom of the matrix $\tilde{A}$. Obviously, this can always be achieved since the Jordan canonical form is unique up to permutations of the Jordan blocks. The matrix $\tilde{A}$ will be as follows

$$\tilde{A} = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$

where

$$J_i = \begin{pmatrix} \mu_i & 1 & & & \\ & \mu_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \mu_i & 1 \\ & & & & \mu_i \end{pmatrix}.$$

All the entries in matrices $\tilde{A}$ and J_i which are not shown are zeros. Suppose that there are r 1×1 Jordan blocks (the transformation S is such that all of them are at the bottom of the matrix) and l Jordan blocks which are not of size 1×1. Let the first Jordan block J_1 be of the size $k_1 \times k_1$, second of the size $k_2 \times k_2$ and so on, and J_l block of size $k_l \times k_l$.

Notice that the matrix $\tilde{A}$ can always be represented as a sum of a diagonal matrix D and a nilpotent matrix N (the last r rows of N will be zeroes) such that $DN = ND$ [7]. Thus $\tilde{A} = D + N, \quad DN = ND$. Let $\{e_j\}$ be the standard basis, and let $De_j = \mu_j e_j$, $1 \leq j \leq n$. Now, choose two vectors $\alpha = (\alpha_1, \ldots, \alpha_n)^t, \quad \beta = (\beta_1, \ldots, \beta_n)^t$ in $\mathbb{C}^n$ in the following way: all $\alpha_i{}'s$ are zero except when $i = k_1, k_1+k_2, \ldots, k_1+\ldots+k_l, k_1+\ldots+k_l+2, k_1+\ldots+k_l+4, \ldots, n$; all $\beta_i{}'s$ are zero except when $i = 1, k_1+1, k_1+k_2+1, \ldots, k_1+\ldots+k_l+1$, $k_1+\ldots+k_l+3, k_1+\ldots+k_l+5, \ldots, n$; assume also that $\dfrac{\beta_n \bar{\alpha}_n}{2\mu_n} \geq 0$. In other words, α and β should look like this

$\alpha = (0, \ldots, 0, \alpha_{m_1}, 0, \ldots, 0, \alpha_{m_2}, 0, \ldots, 0, \alpha_{m_l}, 0, \alpha_{m_l+2}, 0, \alpha_{m_l+4}, \ldots, \alpha_n)^t$ and

$\beta = (\beta_1, 0, \ldots, 0, \beta_{m_1+1}, 0, \ldots, 0, \beta_{m_2+1}, 0, \ldots, 0, \beta_{m_l+1}, 0, \beta_{m_l+3}, 0, \ldots, \beta_n)^t$

where $m_i := \sum_{s=1}^{i} k_s$. Define the matrix $\tilde{B} := (b_{ij})_{i,j=1}^{n}$ as follows $b_{ij} := \dfrac{\beta_i \bar{\alpha}_j}{\mu_i + \mu_j}$ with α and β as specified above. The entries (b_{ij}) of the matrix $\tilde{B}$ are well-defined since $0 \notin \mathcal{E}(A) + \mathcal{E}(A)$.

Claims: (1) $\tilde{A}\tilde{B} + \tilde{B}\tilde{A} = \alpha \otimes \beta$.

(2) The above choice of α and β guarantees the invertibility of the matrix $(I + e^{xP(\tilde{A}) + tQ(\tilde{A})}\tilde{B})$ for all $x, t \in \mathsf{R}$.

To prove claim (1) it is enough to show that for each j, $1 \leq j \leq n$

$$(\tilde{A}\tilde{B} + \tilde{B}\tilde{A})e_j = (\alpha \otimes \beta)e_j.$$

Note that because of the choice of α and β the matrix $\tilde{B}$ has the following structure

$$\tilde{B} = \begin{pmatrix} 0 & \ldots & b_{1m_1} & \ldots & b_{1m_2} & \ldots & b_{1m_l} & \ldots & b_{1n} \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 & \ldots & 0 \\ 0 & \ldots & b_{m_1+1,m_1} & \ldots & b_{m_1+1,m_2} & \ldots & b_{m_1+1,m_l} & \ldots & b_{m_1+1,n} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \ldots & b_{m_l+1,m_1} & \ldots & b_{m_l+1,m_2} & \ldots & b_{m_l+1,m_l} & \ldots & b_{m_l+1,n} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \ldots & b_{n,m_1} & \ldots & b_{n,m_2} & \ldots & b_{n,m_l} & \ldots & b_{n,n} \end{pmatrix}.$$

By direct calculation it can be shown that $N\tilde{B} = \tilde{B}N = 0$. Thus, we have

$$(\tilde{A}\tilde{B} + \tilde{B}\tilde{A})e_j = (D + N)\tilde{B}e_j + \tilde{B}(D + N)e_j = D\tilde{B}e_j + \tilde{B}(\mu_j e_j)$$

$= D(b_{1j}, \ldots, b_{nj})^t + \mu_j(b_{1j}, \ldots, b_{nj})^t = (\mu_1 b_{1j}, \ldots, \mu_n b_{nj})^t + \mu_j(b_{1j}, \ldots, b_{nj})^t =$

$((\mu_1 + \mu_j)b_{1j}, \ldots, (\mu_n + \mu_j)b_{nj})^t = (\beta_1 \bar{\alpha}_j \ldots \beta_n \bar{\alpha}_j)^t = (\alpha \otimes \beta)e_j.$

This proves claim (1).

To prove claim (2) notice that

$$(I + e^{xP(\tilde{A})+tQ(\tilde{A})}\tilde{B}) = (I + e^{xP(D+N)+tQ(D+N)}\tilde{B}) = (I + e^{xP(D)+tQ(D)}\tilde{B})$$

since $N\tilde{B} = 0$. Thus $\det(I + e^{xP(\tilde{A})+tQ(\tilde{A})}\tilde{B}) = 1 + b_{nn}e^{xP(\mu_n)+tQ(\mu_n)}$. Since $b_{nn} = \dfrac{\beta_n\bar{\alpha}_n}{2\mu_n} \geq 0$, it follows that $1 + b_{nn}e^{xP(\mu_n)+tQ(\mu_n)} \neq 0$.

This proves claim (2).

Now we are ready to exhibit $h, g \in \mathbb{C}^n$ and an $n \times n$ matrix B as stated in the theorem. Let $B := S^{-1}\tilde{B}S$,

$h := \alpha_{m_1}(\bar{s}_{m_1 1}, \ldots, \bar{s}_{m_1 n})^t + \alpha_{m_2}(\bar{s}_{m_2 1}, \ldots, \bar{s}_{m_2 n})^t + \cdots + \alpha_n(\bar{s}_{n1}, \ldots, \bar{s}_{nn})^t$

$g := \beta_1(q_{11}, \ldots, q_{n1})^t + \beta_{m_1+1}(q_{1,m_1+1}, \ldots, q_{n,m_1+1})^t + \cdots + \beta_n(q_{1n}, \ldots, q_{nn})^t$

where $S = (s_{ij})$ and $S^{-1} = (q_{ij})$. Note also that $S^{-1}(\alpha \otimes \beta)S = h \otimes g$. Then $AB + BA = S^{-1}\tilde{A}SS^{-1}\tilde{B}S + S^{-1}\tilde{B}SS^{-1}\tilde{A}S = S^{-1}(\tilde{A}\tilde{B} + \tilde{B}\tilde{A})S$

$= S^{-1}(\alpha \otimes \beta)S = h \otimes g$. This proves part (i) of the theorem. Since

$$I + e^{xP(A)+tQ(A)}B = S(I + e^{xP(\tilde{A})+tQ(\tilde{A})}\tilde{B})S^{-1}$$

and $I + e^{xP(\tilde{A})+tQ(\tilde{A})}\tilde{B}$ is invertible, statement (ii) of the theorem follows. □

4. INFINITE DIMENSIONAL CASE

It was shown in [1] that if D is a diagonal operator on ℓ_2 defined by $D(x_i) = (\lambda_i x_i)$, where $\lambda_i > 0$, and $\inf \lambda_i > 0$, then there exist non-trivial vectors $h, g \in \ell_2$ and an operator B on ℓ_2 such that $BD + DB = h \otimes g$. One can easily prove that for such D and B the operator $I + e^{xD+tD^3}B$ is invertible. In this section we extend the above result to a more general situation by using the results of the previous section.

Suppose

$$A = \begin{pmatrix} A_1 & & & & & \\ & A_2 & & & & \\ & & A_3 & & & \\ & & & \ddots & & \\ & & & & A_k & \\ & & & & & \ddots \end{pmatrix}$$

is the matrix representation of an operator on ℓ_2 in the standard basis $\{e_j\}$, where A_k ($k \in \mathbb{N}$) is a normal square matrix of size n_k placed on the main diagonal of the infinite matrix A and there exists a positive integer n_0 such that the size of each A_k is less than or equal to n_0. All the entries of A which are not shown are equal to zero. Assume that there exists a real number $M > 0$ such that $\|A_k\| \leq M$ for all k. Here $\|A_k\|$ is the Euclidian norm of the matrix A_k. Also assume that all eigenvalues $\{\mu_i\}$ of A are positive and constitute a bounded sequence. Since each A_k is normal, for each k there exists a unitary matrix U_k such that

$$U_k A_k U_k^{-1} = D_k \tag{4.1}$$

where D_k is diagonal and $U_k^{-1} = U_k^*$. Let U, U^{-1} and D be infinite matrices constructed as the matrix A but by replacing each block A_k in A by U_k, U_k^{-1} and D_k, respectively.

Proposition 4.1. *The operators A, D, U and U^{-1} constructed as described above are bounded linear operators on ℓ_2.*

Proof. Obviously, the operators A, D, U and U^{-1} are linear. Let $x \in \ell_2$. For convenience we write x in cycles $x = (x_1^{(1)}, \ldots, x_{n_1}^{(1)}; \ldots, x_1^{(k)}, \ldots, x_{n_k}^{(k)}; \ldots)$, where n_k is the size of the matrix A_k, $k = 1, 2, \ldots$. To show that A is bounded note that $Ax = (A_1 r_1, \ldots, A_k r_k, \ldots)$ where $r_k = (x_1^{(k)}, \ldots, x_{n_k}^{(k)})$.

Then $\|Ax\|^2 = \sum\limits_{k=1}^{\infty} \|A_k r_k\|^2$. Further, by using Hölder's inequality

$\|A_k r_k\|^2 = |a_{11}^{(k)} x_1^{(k)} + \ldots + a_{1n_k}^{(k)} x_{n_k}^{(k)}|^2 + \ldots + |a_{n_k 1}^{(k)} x_1^{(k)} + \ldots + a_{n_k n_k}^{(k)} x_{n_k}^{(k)}|^2$

$\leq \sum_{j=1}^{n_k} |a_{1j}^{(k)}|^2 \sum_{j=1}^{n_k} |x_j^{(k)}|^2 + \ldots + \sum_{j=1}^{n_k} |a_{n_k j}^{(k)}|^2 \sum_{j=1}^{n_k} |x_j^{(k)}|^2$

$= \|A_k\|^2 \sum_{j=1}^{n_k} |x_j^{(k)}|^2 \leq M^2 \sum_{j=1}^{n_k} |x_j^{(k)}|^2$. Thus we have

$\|Ax\|^2 = \sum\limits_{k=1}^{\infty} \|A_k r_k\|^2 \leq M^2 \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} |x_j^{(k)}|^2 = M^2 \|x\|^2$. This implies that $Ax \in \ell_2$, and thus A is a well-defined and bounded operator on ℓ_2. Obviously, the diagonal operator $D : \ell_2 \to \ell_2$ will be bounded since the sequence of its eigenvalues $\{\mu_i\}$ is bounded by assumption. To show that U is a well-defined and bounded operator on ℓ_2 notice that since each finite matrix U_k in U is a unitary matrix it follows that $\|U_k\| \leq n_0$. Hence, exactly the same argument that was used to show the boundedness of A can be applied to prove that U is bounded. The same is true about the boundedness of U^{-1} since $\|U_k^{-1}\| = \|U^*\| \leq n_0$. $\square$

Remark 4.2. Note that if U^{-1} is bounded, then U is a bijection from ℓ_2 onto ℓ_2. To show this, first notice that the range of U is dense in ℓ_2. This can be seen from the fact that any sequence which is eventually zero, is in the range of U. Since by the above proposition U is bounded, and ℓ_2 is complete, it follows that U is a closed operator. Hence, U^{-1} is also closed. If U^{-1} is bounded, then the domain of U^{-1} (which is the range of U) is closed [8]. Hence, range(U)=ℓ_2.

Before proceeding further, we need two lemmas. The operators R and S defined in the following lemmas are used in [3]. Even though some of the properties of R and S are implicitly present in [3], we provide some explicit proofs for the sake of completion.

Lemma 4.3. *Let $\alpha = \{\alpha_i\} \in \ell_2, \beta = \{\beta_i\} \in \ell_2$, and let $\{\mu_i\}$ be an infinite bounded sequence of positive real numbers such that $\epsilon = \inf_{i \in N} \mu_i > 0$. Then*

(i) $R : \ell_2 \to L_2(0,\infty)$ defined by $R(x_1, x_2, \ldots)(s) = \sum_{j=1}^{\infty} \bar{\alpha}_j e^{-s\mu_j} x_j$ is a bounded linear operator, and

(ii) $S : L_2(0,\infty) \to \ell_2$ defined by $S(f) = (\ldots, \int_0^\infty f(s) e^{-s\mu_i} \beta_i ds, \ldots)$ is a bounded linear operator.

Proof. Proof of (i). Clearly, R is a linear operator. Let $x = (x_i) \in \ell_2$. Then

$$\left(\int_0^\infty |\bar{\alpha}_j e^{-s\mu_j} x_j|^2 ds \right)^{1/2} = \left(|\bar{\alpha}_j x_j|^2 \int_0^\infty e^{-2s\mu_j} ds \right)^{1/2} = \frac{|\bar{\alpha}_j x_j|}{\sqrt{2\mu_j}} \leq \frac{|\bar{\alpha}_j x_j|}{\sqrt{\epsilon}}.$$

Thus

$$\sum_{j=1}^{\infty} \|\bar{\alpha}_j e^{-s\mu_j} x_j\|_{L_2} \leq \sum_{j=1}^{\infty} \frac{|\bar{\alpha}_j x_j|}{\sqrt{\epsilon}} < \infty.$$

Since $L_2(0,\infty)$ is a Banach space and absolute convergence in a Banach space implies convergence, it follows that $\sum_{j=1}^{\infty} \bar{\alpha}_j e^{-s\mu_j} x_j$ is convergent in $L_2(0,\infty)$, and hence R is well-defined. To show that R is bounded, notice that by using the Hölder's inequality

$\|R(x_1,x_2,\dots)\|_{L_2} = \left\| \sum_{j=1}^{\infty} \bar{\alpha}_j e^{-s\mu_j} x_j \right\|_{L_2} \leq \sum_{j=1}^{\infty} |\bar{\alpha}_j x_j| \cdot \|e^{-s\mu_j}\|_{L_2}$

$\leq \sum_{j=1}^{\infty} \frac{|\bar{\alpha}_j x_j|}{\sqrt{\epsilon}} \leq \frac{1}{\sqrt{\epsilon}} \left(\sum_{j=1}^{\infty} |\alpha_j|^2 \right)^{1/2} \left(\sum_{j=1}^{\infty} |x_j|^2 \right)^{1/2}$. Thus

$$\|R(x_1,x_2,\dots)\|_{L_2} \leq \frac{1}{\sqrt{\epsilon}} \|\alpha\|_{\ell_2} \|x\|_{\ell_2}.$$

This completes the proof of part (i).

Proof of part (ii). It is clear that S is linear. Next, note that

$\sum_{i=1}^{\infty} \left| \int_0^{\infty} f(s) e^{s\mu_i} \beta_i ds \right|^2 \leq \sum_{i=1}^{\infty} |\beta_i|^2 \left(\int_0^{\infty} |f(s)|^2 ds \right) \left(\int_0^{\infty} |e^{-s\mu_i}|^2 ds \right)$

$= \|f\|_{L_2}^2 \sum_{i=1}^{\infty} \frac{|\beta_i|^2}{2\mu_i} \leq \frac{\|f\|_{L_2}^2}{2\epsilon} \sum_{i=1}^{\infty} |\beta_i|^2 < \infty.$

This completes the proof of the lemma. $\square$

Lemma 4.4. *If K is a compact operator on $L_2(0,\infty)$ and -1 is not an eigenvalue of K, then $(I+K)^{-1}$ exists and is bounded on $L_2(0,\infty)$.*

Proof. Since any non-zero spectral value of K is an eigenvalue, it follows that -1 is in the resolvent set of K. Hence, $(I+K)^{-1}$ is defined on L_2 and is bounded. $\square$

Before proceeding further we need the following definition. A sequence of square matrices $\{T_k\}$, $k \in \mathsf{N}$ is said to be a sequence of matrices of bounded size if there exists a positive integer n_0 such that the size of a matrix T_k is less than or equal to n_0 for all k. Now we are ready to prove one of the main results of the section. Some arguments in the proof of the following theorem are similar to those given in [3].

Theorem 4.5. *Let $\{A_k\}$ be the sequence of normal matrices of bounded size such that all eigenvalues $\{\mu_i\}$ of A_k's are positive and constitute a bounded sequence such that $\epsilon = \inf_{i \in N} \mu_i > 0$. Let $A : \ell_2 \to \ell_2$ be an operator constructed as in the discussion preceding the proposition 4.1. If there is a positive real number M such that $\|A_k\| \le M$ for all k, then there exist non-trivial vectors $h, g \in \ell_2$ and a bounded linear operator $B : \ell_2 \to \ell_2$ such that*

(i) $AB + BA = h \otimes g$, *and*

(ii) $I + e^{xP(A)+tQ(A)}B$ *is invertible for all $x, t \in R$, where $P(A)$ and $Q(A)$ are polynomials in A with real coefficients.*

Proof. For any $x, t \in R$ define an operator $K_{x,t} : L_2(0,\infty) \to L_2(0,\infty)$ by $(K_{x,t}f)(w) = \int_0^\infty k_{x,t}(s,w)f(s)ds$ where $k_{x,t}(s,w) := \sum_{k=1}^\infty c_k e^{-\mu_k(w+s)+xP(\mu_k)+tQ(\mu_k)}$, $c_k > 0$. It is straightforward to see that $K_{x,t}$ is well defined. Claim 1: -1 is not an eigenvalue of the operator $K_{x,t}$. This can be seen from the following consideration.

$$\begin{aligned}
\langle K_{x,t}f, f\rangle &= \int_0^\infty (K_{x,t}f)(w)\bar{f}(w)dw \\
&= \int_0^\infty \left(\int_0^\infty \sum_{k=1}^\infty c_k e^{-\mu_k(w+s)+xP(\mu_k)+tQ(\mu_k)} f(s)ds\right)\bar{f}(w)dw \\
&= \sum_{k=1}^\infty c_k e^{xP(\mu_k)+tQ(\mu_k)}\left(\int_0^\infty f(s)e^{-s\mu_k}ds\right)\int_0^\infty \bar{f}(w)e^{-w\mu_k}dw \\
&= \sum_{k=1}^\infty c_k e^{xP(\mu_k)+tQ(\mu_k)}\left(\int_0^\infty f(s)e^{-s\mu_k}ds\right)\overline{\int_0^\infty f(w)e^{-w\mu_k}dw} \\
&= \sum_{k=1}^\infty c_k e^{xP(\mu_k)+tQ(\mu_k)}\left|\int_0^\infty f(s)e^{-s\mu_k}ds\right|^2 \ge 0.
\end{aligned}$$

Thus

$$\langle (I+K_{x,t})f,f\rangle = \langle f,f\rangle + \langle K_{x,t}f,f\rangle = \|f\|^2 + \langle K_{x,t}f,f\rangle \geq 0$$

Hence, -1 is not an eigenvalue of $K_{x,t}$. This completes the proof of claim 1.

Claim 2: $K_{x,t}$ is a compact operator. To prove this it suffices to show that $k_{x,t}(s,w) \in L_2\Big((0,\infty)\times(0,\infty)\Big)$ [5]. Let $c=(c_k)\in \ell_2$. For fixed $x,t \in \mathsf{R}$

$$\begin{aligned}
&\int_0^\infty\int_0^\infty |c_j e^{-\mu_j(w+s)+xP(\mu_j)+tQ(\mu_j)}|^2 ds dw \\
&= |c_j|^2 e^{2xP(\mu_j)+2tQ(\mu_j)} \int_0^\infty\int_0^\infty e^{-2\mu_j(w+s)} ds dw \\
&= \frac{|c_j|^2}{4\mu_j^2} e^{2xP(\mu_j)+2tQ(\mu_j)} \leq \frac{|c_j|^2}{\epsilon^2} e^{2xP(\|D\|)+2tQ(\|D\|)}
\end{aligned} \tag{4.2}$$

where D is a diagonal operator constructed from the matrices (4.1) as described at the beginning of the section. Then (4.2) implies that

$$\left(\int_0^\infty\int_0^\infty |c_j e^{-\mu_j(w+s)+xP(\mu_j)+tQ(\mu_j)}|^2 ds dw\right)^{1/2} \leq \frac{|c_j|}{\epsilon} e^{xP(\|D\|)+tQ(\|D\|)}.$$

This, in turn, implies that the series $\sum_{k=1}^{\infty} c_k e^{-\mu_k(w+s)+xP(\mu_k)+tQ(\mu_k)}$ converges absolutely, and hence, since L_2 is a Banach space $k_{x,t}(s,w) \in L_2\Big((0,\infty)\times(0,\infty)\Big)$. This completes the proof of claim 2.

Let U be a unitary transformation which diagonalizes A, i.e., there exists a diagonal infinite matrix D such that $D = UAU^{-1}$. The construction of such operators U and U^{-1} was described at the beginning of this section. Let $\{e_j\}$ be the standard basis, and let $De_j = \mu_j e_j$, $j \in \mathsf{N}$. Let $\alpha, \beta \in \ell_2$

$$\alpha = (\alpha_1, \ldots, \alpha_n, \ldots)^t, \quad \beta = (\beta_1, \ldots, \beta_n, \ldots)^t$$

be such that $c_k := \bar{\alpha_k}\beta_k > 0$ for any $k \in \mathsf{N}$. According to Lemma 2.2 the operator equation $D\tilde{B}+\tilde{B}D = \alpha\otimes\beta$ has the solution $\tilde{B} = \int_0^\infty e^{-sD}(\alpha\otimes\beta)e^{-sD}ds$. Thus, we have

$$\begin{aligned}
\tilde{B}(x_1,\ldots,x_n,\ldots) &= \int_0^\infty e^{-sD}(\alpha\otimes\beta)(e^{-s\mu_1}x_1,\ldots,e^{-s\mu_n}x_n,\ldots)ds \\
&= \int_0^\infty \sum_{j=1}^\infty \bar{\alpha_j}\, e^{-s\mu_j}x_j\,(e^{-s\mu_1}\beta_1,\ldots,e^{-s\mu_n}\beta_n,\ldots)ds \\
&= (\ldots,\underbrace{\int_0^\infty \sum_{j=1}^\infty \bar{\alpha_j}\, e^{-s\mu_j}x_j\, e^{-s\mu_i}\beta_i ds}_{\text{ith component}},\ldots) \\
&= SR(x_1,\ldots,x_n,\ldots)
\end{aligned}$$

where operators $R : \ell_2 \to L_2$ and $S : L_2 \to \ell_2$ are defined in Lemma 4.3. Then according to Lemma 2.1,

$I + e^{xP(D)+tQ(D)}\tilde{B} = I + e^{xP(D)+tQ(D)}SR$ is invertible iff $I + Re^{xP(D)+tQ(D)}S$ is invertible. Now

$$\begin{aligned}
Re^{xP(D)+tQ(D)}S(f)(w) &= Re^{xP(D)+tQ(D)}(\ldots,\int_0^\infty f(s)e^{-s\mu_i}\beta_i ds,\ldots)(w) \\
&= R(\ldots,\int_0^\infty f(s)e^{-s\mu_i}\beta_i e^{xP(\mu_i)+tQ(\mu_i)}ds,\ldots)(w) \\
&= \sum_{j=1}^\infty \bar{\alpha_j}\int_0^\infty f(s)e^{-s\mu_i}\beta_i e^{xP(\mu_i)+tQ(\mu_i)}ds\cdot e^{-w\mu_j} \\
&= \int_0^\infty \sum_{k=1}^\infty \bar{\alpha_k}\beta_k e^{-\mu_k(w+s)+xP(\mu_k)+tQ(\mu_k)}f(s)ds \\
&= (K_{x,t}f)(w).
\end{aligned}$$

Since $I + K_{x,t}$ is invertible it follows that $I + e^{xP(D)+tQ(D)}\tilde{B}$ is invertible. Let

$B := U^{-1}\tilde{B}U$,

$h := \alpha_1(\bar{u}_{11},\ldots,\bar{u}_{1n},\ldots)^t + \ldots + \alpha_n(\bar{u}_{n1},\ldots,\bar{u}_{nn},\ldots)^t + \ldots$

$g := \beta_1(q_{11}, \ldots, q_{n1}, \ldots)^t + \ldots + \beta_n(q_{1n}, \ldots, q_{nn}, \ldots)^t + \ldots$

where $U = (u_{ij})$ and $U^{-1} = (q_{ij})$. Notice also that $U^{-1}(\alpha \otimes \beta)U = h \otimes g$. Then one checks

$$\begin{aligned} AB + BA &= U^{-1}DUU^{-1}\tilde{B}U + U^{-1}\tilde{B}UU^{-1}DU \\ &= U^{-1}(D\tilde{B} + \tilde{B}D)U \\ &= U^{-1}(\alpha \otimes \beta)U = h \otimes g. \end{aligned}$$

This proves part (i) of the theorem. Since

$$I + e^{xP(A)+tQ(A)}B = U(I + e^{xP(D)+tQ(D)}\tilde{B})U^{-1}$$

and $I + e^{xP(D)+tQ(D)}\tilde{B}$ is invertible, statement (ii) of the theorem follows. $\square$

The previous theorem was a special case of the operator A made up of infinitely many normal matrices. What happens if all the conditions of Theorem 4.5 are satisfied except for the normality of matrices A_k? Thus, suppose an operator $A : \ell_2 \to \ell_2$ is given which has the following matrix representation in the standard basis $\{e_i\}$

$$A = \begin{pmatrix} A_1 & & & & & \\ & A_2 & & & & \\ & & A_3 & & & \\ & & & \ddots & & \\ & & & & A_k & \\ & & & & & \ddots \end{pmatrix}$$

where A_k $(k \in \mathsf{N})$ is a square matrix of size n_k placed on the main diagonal of the infinite matrix A, and there exists a positive integer n_0 such that the size of each A_k is less than or equal to n_0. All the entries of A which are not shown are equal to zero. Assume that there exists a real number $M > 0$ such that $\|A_k\| \leq M$ for all k. Also assume that all eigenvalues $\{\mu_i\}$ of A are positive and constitute a bounded sequence. Then, there exists a similarity transformation S_k such that $S_k A_k S_k^{-1} = \tilde{A}_k \quad \forall k \in \mathsf{N}$ where $\tilde{A}_k$ is in Jordan canonical form. Thus, from the previous section

$$\tilde{A}_k = D_k + N_k, \quad D_k N_k = N_k D_k$$

where D_k is a diagonal matrix and N_k is a nilpotent matrix. Moreover, S_k is chosen in such a way that all 1×1 Jordan blocks in $\tilde{A}_k$ appear at the bottom of $\tilde{A}_k$. Next, construct operators $\tilde{A}, S, S^{-1}, D$ and N in the same way as described before. Notice that $SAS^{-1} = \tilde{A}$ and $\tilde{A} = D + N$ so that $DN = ND$, where D is a diagonal operator and N is a nilpotent operator. Now

$$S_k A_k S_k^{-1} = \left(\frac{S_k}{\|S_k\|}\right) A_k \|S_k\| S_k = \tilde{A}_k.$$

Let $T_k = \dfrac{S_k}{\|S_k\|}$. It can be seen that $\|T_k\| = 1$, implying that T is bounded. It is easy to see that, under the assumption of boundedness of T^{-1}, Proposition 4.1 and Remark 4.2 carry over to operators A, D and T. Using arguments similar to those of theorem 4.5 and theorem 3.2 we can prove the following

Theorem 4.6. *Let $\{A_k\}$ be a sequence of matrices of bounded size such that all eigenvalues $\{\mu_i\}$ of the A_k's are positive and constitute a bounded sequence such that $\inf_{i \in \mathsf{N}} \mu_i > 0$. Let $A : \ell_2 \to \ell_2$ be an operator constructed as in the* *discussion preceding the statement of this theorem. If there is a positive real number M such that $\|A_k\| \leq M$ for all k and the operator T^{-1} is bounded from ℓ_2 onto ℓ_2, then there always exist non-trivial vectors $h, g \in \ell_2$ and a bounded linear operator $B : \ell_2 \to \ell_2$ such that*

(i) $AB + BA = h \otimes g$, *and*

(ii) $I + e^{xP(A)+tQ(A)}B$ *is invertible for all* $x, t \in \mathbb{R}$, *where* $P(A)$ *and* $Q(A)$ *are polynomials in* A *with real coefficients.*

REFERENCES

1. H. Aden and B. Carl, On realizations of solutions of the KdV equation by determinants on operator ideals, Journal of Mathematical Physics, 37 (1996), 1833-1857.
2. R. Bhatia and P. Rosenthal, How and why to solve the operator equation $AX - XB = Y$, Bull. London Math. Soc., 29 (1997), 1-21.
3. H. Blohm, Solution of inverse scattering problems and non-linear equations by trace methods, submitted to Nonlinearity.
4. B. Carl and C. Schiebold, Nonlinear equations in soliton physics and operator ideals, Nonlinearity, 12 (1999), 333-364.
5. J. B. Conway, A Course in Functional Analysis, Springer-Verlag, New York, 1998.
6. A. Eschmeier, Tensor products and elementary operators, J. reine und angew. Math, 390 (1988), 47-66.
7. P. R. Halmos, Finite-Dimensional Vector Spaces, Van Nostrand, New York, 1958.
8. E. Kreyszig, Introductory Functional Analysis with Applications, John Wiley and Sons, New York, 1978.