Zeitschrift für Analysis
und ihre Anwendungen
Bd. 6 (5) 1987, S. 409–419

On Certain Higher Order Riccati-Type Operator Equations with Possibly Unbounded Operator Coefficients

T. Mazumdar

Es seien $\mathcal{H}$ und $\mathcal{V}^1$ komplexe Hilbert-Räume, mit topologischer Einbettung von $\mathcal{V}^1$ in $\mathcal{H}$, und $\mathcal{V}^2$ ein komplexer Prähilbertraum. Es wird die Existenz einer Lösung $X = X_Q \in \mathcal{L}(\mathcal{V}^2, \mathcal{V}^1)$ der Operatorengleichung

$$A_1XA_2 - B_1XB_2 + XDX + XEXFX = Q$$

im Raum der (beschränkten oder unbeschränkten) linearen Operatoren in $\mathcal{H}$ in der Situation $A_1, B_1 \in \mathcal{L}(\mathcal{V}^1, \mathcal{H})$, $D, E, F \in \mathcal{L}(\mathcal{V}^1, \mathcal{V}^2)$, $A_2, B_2 : \mathcal{V}^2 \to \mathcal{V}^2$ linear und $Q \in \mathcal{L}(\mathcal{V}^2, \mathcal{H})$ eindimensional erörtert. Unter gewissen Voraussetzungen wird ein iteratives Näherungsverfahren für die Existenz solch einer Lösung angegeben und zwei Beispiele werden gebracht.

Пусть $\mathcal{H}$ и $\mathcal{V}^1$ комплексные гильбертовы пространства, $\mathcal{V}^1$ топологически вложено в $\mathcal{H}$, и $\mathcal{V}^2$ комплексное предгильбертово пространство. Обсуждается существование решения $X = X_Q \in \mathcal{L}(\mathcal{V}^2, \mathcal{V}^1)$ операторного уравнения

$$A_1XA_2 - B_1XB_2 + XDX + XEXFX = Q$$

в пространстве (ограниченных или неограниченных) операторов в $\mathcal{H}$ в ситуации $A_1, B_1 \in \mathcal{L}(\mathcal{V}^1, \mathcal{H})$, $D, E, F \in \mathcal{L}(\mathcal{V}^1, \mathcal{V}^2)$, $A_2, B_2 : \mathcal{V}^2 \to \mathcal{V}^2$ линейны и $Q \in \mathcal{L}(\mathcal{V}^2, \mathcal{H})$ одномерно. Под некоторыми предположениями даются итерационный метод для его решения и приводятся два примера.

Let $\mathcal{H}$ and $\mathcal{V}^1$ be complex Hilbert spaces, with $\mathcal{V}^1$ topologically included in $\mathcal{H}$, and $\mathcal{V}^2$ a complex pre-Hilbert space. There is considered the existence of a solution $X = X_Q \in \mathcal{L}(\mathcal{V}^2, \mathcal{V}^1)$ of the operator equation

$$A_1XA_2 - B_1XB_2 + XDX + XEXFX = Q$$

in the space of (bounded or not) linear operators in $\mathcal{H}$ under the data $A_1, B_1 \in \mathcal{L}(\mathcal{V}^1, \mathcal{H})$, $D, E, F \in \mathcal{L}(\mathcal{V}^1, \mathcal{V}^2)$; $A_2, B_2 : \mathcal{V}^2 \to \mathcal{V}^2$ linear and $Q \in \mathcal{L}(\mathcal{V}^2, \mathcal{H})$ one-dimensional. Under some hypotheses, an iterative analytic method to arrive at the existence of such a solution is given. Two examples are given.

§ 1 Introduction

The purpose of the present paper is to show how certain pertubation techniques coupled with results of [4] lead us to existence of operator solution X of Riccati-type equations of the form

$$A_1XA_2 - B_1XB_2 + XDX + XEXFX = Q \tag{1.1}$$

in which $A_1, A_2, B_1, B_2, D, E, F, Q$ are given linear operators that may be bounded or unbounded in an underlying Hilbert space, Q having one-dimensional range. The "magnitudes" of one or the other of the operators D, E, F, Q will be small enough so that the terms involving them may be looked upon as perturbations to the linear part $A_1XA_2 - B_1XB_2$. We mention [6], in passing, where problems with rank $Q = 1$ have been treated in the context of Lyapunov equations and its generalizations.

We find, in [2], that problems such as (1.1) above are dealt with from a purely algebraic point of view in a finite-dimentional matrix setting, whereas our approach will be largely analytic. In [7] also, we find such problems treated in relation to the theory of backscattering of a travelling beam of particles. If, in equation (1.1) above, we take E or F to be zero, we obtain the familiar equation $A_1XA_2 - B_1XB_2 + XDX = Q$, which has been dealt with in widely differing contexts in literature (cf. [1—3, 5]). We particularly mention [3] in the context of feedback optimal control theory of distributed parameter systems, in which the setting is infinite-dimensional and D is non-zero, with E or F zero. In this context, the restriction "rank $Q = 1$" would signify that the observation operator or the detection mechanism has one-dimensional range. In the transport theory, such a restriction on Q might describe some kind of special relationship among the various probabilities that a moving particle has in changing over from one state to another.

Our approach to the problem at hand is different from those found in any of the references above. We permit the possibility that one or the other of D, E, F may be zero, Q remaining non-zero. We have tried to give a unified approach to classes of problems somewhat similar to the problems appearing in the references above. Whether our methods are directly applicable to the actual problems arising in practice remains to be investigated.

§ 2 Notations

Let $\mathbf{C}$ denote the set of complex numbers and $\mathbf{N}$ the set of natural ones. $\mathcal{H}$ will denote a complex Hilbert space with norm $|\cdot|_{\mathcal{H}}$ and inner product $(\cdot, \cdot)_{\mathcal{H}}$, $\mathcal{V}^1$ a complex Hilbert space with norm $\|\cdot\|_1$ and inner product $((\cdot, \cdot))_1$ such that $\mathcal{V}^1$ is a dense subspace of $\mathcal{H}$ with continuous inclusion injection from $\mathcal{V}^1$ into $\mathcal{H}$. Consequently, there exists a constant $\gamma > 0$ such that

$$|v|_{\mathcal{H}} \leqq \gamma \|v\|_1 \quad \text{for all} \quad v \in \mathcal{V}^1. \tag{2.1}$$

$\mathcal{V}^2$ will be a complex pre-Hilbert space with norm $\|\cdot\|_2$ and inner product $((\cdot, \cdot))_2$. If X is a normed linear space and Y a Banach space, then $\mathcal{L}(X, Y)$ will denote the Banach space of bounded linear operators from X into Y, with the usual norm topology on it. The suffixes to the norm notations $\|\cdot\|$, indicative of the space in which the norm is taken, will usually be omitted, because the space is often clear from the context. Let $\mathcal{X} = \mathcal{L}(\mathcal{V}^2, \mathcal{H})$, $\mathcal{W} = \mathcal{L}(\mathcal{V}^2, \mathcal{V}^1)$. Considered given in our problem are

$$A_1, B_1 \in \mathcal{L}(\mathcal{V}^1, \mathcal{H}), \qquad D, E, F \in \mathcal{L}(\mathcal{V}^1, \mathcal{V}^2), \qquad Q \in \mathcal{X}$$

and

$$\text{linear operators} \quad A_2, B_2 : \mathcal{V}^2 \to \mathcal{V}^2.$$

(We may look upon A_2, B_2 as elements of $\mathcal{L}(\mathcal{V}^2, \mathcal{K})$ for an appropriate Banach space $\mathcal{K}$). Because of (2.1), the possibility is kept open that A_1, B_1 are unbounded linear operators in $\mathcal{H}$. If $\mathcal{V}^2$ is also a subspace of $\mathcal{H}$, as we will have in examples in § 5, $\mathcal{H}$ will be our underlying Hilbert space, and then the rest of the operators above may also turn out to be unbounded in $\mathcal{H}$.

With domains and ranges laid out as above, equation (1.1) is now well-defined. Our concern of proving existence of solution $X = X_Q \in \mathcal{W}$ of equation (1.1) under a suitable set of sufficient conditions will now be precisely stated as the main result in the next section. The proof given in § 4 will consist of an iterative approximation prodecure converging to the solution X_Q.

§ 3 The main result

We start by listing the **hypotheses** we will work under.

(**H 1**) *Q has one-dimensional range; say $Q(\mathcal{V}^2) = \{\alpha h : \alpha \in \mathbf{C}\}$, denoted by $[h]$, for some $h \in \mathcal{H}$ with $|h|_{\mathcal{H}} = 1$.*

(**H 2**) $[h] \subset \mathcal{V}^1$; $A_1([h]) = [h] = B_1([h])$.
We will use the notations $A_{1,h}$, $B_{1,h}$, respectively, for the restrictions of A_1, B_1 to $[h]$.

(**H 3**) *There exists an orthonormal basis $\mathcal{B} = \{b_i : i \in \mathbf{N}\}$ of $\mathcal{V}^2$ such that each b_i is an eigenvector of both A_2 and B_2 belonging to eigenvalues λ_i and μ_i, respectively.*

We will denote by $\mathcal{V}_n^2$ the subspace of $\mathcal{V}^2$ generated by $b_1, b_2, \ldots, b_n$. $A_{2,n}$, $B_{2,n}$, Q_n are restrictions of A_2, B_2, Q to $\mathcal{V}_n^2$. $\mathcal{W}(n)$ is the set of restrictions to $\mathcal{V}_n^2$ of all those $X \in \mathcal{W}$ such that $Xb_i = 0$ for all $i > n$ and $Xb_i \in [h]$ for all $i \leqq n$. $\mathcal{W}(n)$ will be considered isomorphic to $\mathcal{L}(\mathcal{V}_n^2, [h])$, as can be easily seen to be the case. Also, we will use the notation $\mathcal{W}_h = \mathcal{L}(\mathcal{V}^2, [h])$. Both $\mathcal{L}(\mathcal{V}_n^2, [h])$ and $\mathcal{W}_h$ will be considered to have the subspace topology of $\mathcal{W}$.

(**H 4**) *There exists a constant $\beta > 0$ such that for all nonzero $Y \in \mathcal{W}_h$, there exists a $\Phi_Y \in \mathcal{V}^2$ satisfying the dominance relation*

$$|(A_{1,h} Y A_2 - B_{1,h} Y B_2)\, \Phi_Y|_{\mathcal{H}} > \beta\, \|Y\|_{\mathcal{W}}\, \|\Phi_Y\|_2 .$$

This condition is obviously an extension of the well-known concept of ellipticity or coercivity (cf. [3]), and may be called a one-sided coercivity condition. In § 5 we will indicate a class of examples for which (H 4) may be verified. A direct consequence of (H 4) is the following condition, utilized in [4]:
There exists a constant $\beta > 0$ such that for all $n \in \mathbf{N}$ and for all nonzero $Y \in \mathcal{W}(n)$, there exists a $\Phi_Y \in \mathcal{V}_n^2$ satisfying the dominance relation

$$|(A_{1,h} Y A_{2,n} - B_{1,h} Y B_{2,n})\, \Phi_Y|_{\mathcal{H}} > \beta\, \|Y\|_{\mathcal{W}}\, \|\Phi_Y\|_2 .$$

We note in passing that if $Y \in \mathcal{W}(n)$, then $\|Y\|_{\mathcal{W}(n)} = \|Y\|_{\mathcal{W}_h} = \|Y\|_{\mathcal{W}}$, and if $X \in \mathcal{W}_h$, then $\|X\|_{\mathcal{W}_h} = \|X\|_{\mathcal{W}}$, with the norms defined via the usual suprema.

In the sequel, we will frequently use the notations (see (2.1))

$$\alpha_1 = \frac{\gamma}{\beta^2}\, \|Q\|\, \|D\|, \qquad \alpha_2 = \frac{\gamma}{\beta^3}\, \|Q\|^2\, \|E\|\, \|F\|, \tag{3.1}$$

$\|Q\|$ always meaning $\|Q\|_{\mathcal{T}}$, and norms of D, E, F being always taken in $\mathcal{L}(\mathcal{V}^1, \mathcal{V}^2)$.

(**H 5**) *There exists a $\Delta > 0$ such that $1 + \alpha_1 + \alpha_2 < \Delta$ and $1 + \alpha_1 \Delta^2 + \alpha_2 \Delta^3 < \Delta$.*

This condition delineates in what sense the part $XDX + XEXFX$ of (1.1) may be considered as perturbation to the part $A_1XA_2 - B_1XB_2$. $\|D\|$, $\|E\|$, $\|F\|$, though, need not be small if $\|Q\|$ is small enough. A small $\|Q\|$ might correspond to a weak observation or detection process in optimal control theory. A small $\|D\|$ might indicate that the feedback process is weak. In nuclear transport theory, a small $\|Q\|$ might indicate a low probability of moving particles changing over to particles in different states moving in the opposite direction. Examples of validity of (H 5) are not difficult to come by. For example, if $\Delta = 2$, $\alpha_1 < 1/8$, $\alpha_2 < 1/16$, then (H 5) is satisfied. Or else, if $\Delta > \left(1 + \sqrt{5}\right)/2$, $\alpha_1 < \delta_1/\Delta^2$, $\alpha_2 < \delta_2/\Delta^3$ where δ_1 and δ_2 are such that $\delta_1 + \delta_2 < \Delta - 1$, then also (H 5) is satisfied. The estimates (H 5) lead to a successful existence theorem. It is not claimed that they are the best possible estimates.

(**H6**) *There exists a fixed number $k_0 \geqq 1$ such that*

(i) $k_0(\alpha_1 + \alpha_2) < 1$,

(ii) $$\frac{2\alpha_1}{1 - k_0(\alpha_1 + \alpha_2)} + \frac{3\alpha_2 k_0(\alpha_1 + \alpha_2)}{1 - k_0(\alpha_1 + \alpha_2)} + \frac{3\alpha_2}{[1 - k_0(\alpha_1 + \alpha_2)]^2}$$
$$+ \alpha_1 k_0(\alpha_1 + \alpha_2) + \alpha_2[k_0(\alpha_1 + \alpha_2)]^2 \leqq k_0(\alpha_1 + \alpha_2),$$

(iii) *if $\Delta > 2$, then* $\dfrac{\alpha_1}{[1 - k_0(\alpha_1 + \alpha_2)]^2} + \dfrac{\alpha_2}{[1 - k_0(\alpha_1 + \alpha_2)]^3} \leqq 1$.

An example when these conditions are valid is obtained by taking $\alpha_1 < 1/18$, $\alpha_2 < 1/61$, $k_0 = 1/2(\alpha_1 + \alpha_2)$. Indeed, we have the following proposition[1])

Proposition 3.1: a) *For an arbitrary $\Delta > 1$, there exist $a_1 > 0$, $a_2 > 0$ such that* (H 5) *is satisfied whenever $0 \leqq \alpha_1 < a_1$, $0 \leqq \alpha_2 < a_2$.*

b) *For arbitrary $k_0 > 2$ and $\tau > 0$, there exists an $m_\tau > 0$ such that if $0 < \alpha_1 < m_\tau$ and $\alpha_2 = \alpha_1^{1+\tau}$, then* (H 6) *is satisfied.*

c) *For arbitrary $\Delta > 1$ and $k_0 > 2$, there exist $a_1 > 0$ and $\tau > 0$ such that if $0 \leqq \alpha_1 < a_1$ and $\alpha_2 = \alpha_1^{1+\tau}$, then both* (H 5) *and* (H 6) *hold.*

Proof: It suffices to prove statement b). First we show the validity of the inequality (H 6)/(ii) which may be alternatively written as (on division by $k_0(\alpha_1 + \alpha_2)$, since $\alpha_2 = \alpha_1^{1+\tau}$)

$$\frac{2}{k_0(1 + \alpha_1^\tau)\,[1 - k_0\alpha_1(1 + \alpha_1^\tau)]} + \frac{3\alpha_1}{1 - k_0\alpha_1(1 + \alpha_1^\tau)}$$
$$+ \frac{3\alpha_1^\tau}{k_0(1 + \alpha_1^\tau)\,[1 - k_0\alpha_1(1 + \alpha_1^\tau)]^2} + \alpha_1 + \alpha_1^{2+\tau}k_0(1 + \alpha_1^\tau) \leqq 1. \tag{3.2}$$

The first term on the left side of this inequality approaches $2/k_0$ as $\alpha_1 \to 0$, whereas the other terms tend to zero. Since, by hypothesis, $(2/k_0) < 1$, we have the inequality (3.2), and hence (H 6)/(ii), true for α_1 small enough, say $\alpha_1 < m_\tau$. To complete the proof, one has now to take care of parts (i) and (iii) of (H 6). This can be easily done ∎

Let us note here that the hypothesis (H 6) above takes a simplified form if $\alpha_2 = 0$, i.e., if the term $XEXFX$ does not appear in the original equation (1.1). We also become unpleasantly aware of the fact that if we want to apply our methods to a still higher dimensional Riccati-type equation, namely one that includes a fourth degree term in X, then the complexities of our estimates will grow rapidly.

Now we are ready to state our main result.

Theorem 3.2: *Under the hypotheses* (H 1)—(H 6) *above, there exists a solution $X_Q \in \mathscr{W}$ of the equation* (1.1) ∎

We can put a regularity feature on this solution X_Q. For all $i \in \mathrm{N}$, define $f_i \in \mathscr{W}$ by $f_i(\alpha b_j) = \alpha\delta_{ij}h$ for all $j \in \mathrm{N}$, where $\alpha \in \mathbf{C}$, b_j's are given by (H 3), and $\delta_{ij} = 0$ if $i \neq j$, $\delta_{ii} = 1$. Let $\mathscr{W}_1$ be the set of all finite linear combinations of the f_i's over $\mathbf{C}$. The topology on $\mathscr{W}_1$ is the one inherited from $\mathscr{W}$. Let $\overline{\mathscr{W}}_1$ be the closure of $\mathscr{W}_1$ in $\mathscr{W}$. Clearly $\overline{\mathscr{W}}_1 \subset \mathscr{W}_h$. Let $\tilde{\mathcal{V}}^2$ be the completion of $\mathcal{V}^2$.

Lemma 3.3: *$\overline{\mathscr{W}}_1$ is isomorphic to the Hilbert space $\tilde{\mathcal{V}}^2$, and so $\overline{\mathscr{W}}_1$ is separable and reflexive.*

The proof consists of standard methods of functional analysis, and is omitted ∎

---

[1]) The author thanks the referees whose suggestions led him to this proposition.

Theorem 3.4: *Under the hypotheses* (H 1)–(H 6) *above, equation* (1.1) *has a solution* $X_Q \in \overline{\mathscr{W}}_1$.

It is evident that Theorem 3.4 includes the result of Theorem 3.2.

§ 4 Proof of Theorem 3.4

We repeatedly apply the following existence — uniqueness theorem taken from [4].

Theorem 4.1: *Under the hypotheses* (H1)–(H4) *above, there exists a unique solution* $X = X^{(Q)} \in \overline{\mathscr{W}}_1$ *of the equation*

$$(A_1XA_2 - B_1XB_2)\,\Phi = Q\Phi \quad \text{for all} \quad \Phi \in \mathcal{V}^2. \tag{4.1}$$

Moreover, $X^{(Q)}$ *has the same range as* Q *has, and* $\|X^{(Q)}\|_{\mathscr{W}} \leqq \|Q\|/\beta$.

We now proceed with the iterative proof of Theorem 3.4. We first construct a sequence of "approximate solutions" of equation (1.1) with the following lemma.

Lemma 4.2: *Under the hypotheses* (H 1)–(H 6) *above, there exists a sequence* $\{X^{(i)}\}_{i\in\mathbf{N}} \subset \overline{\mathscr{W}}_1$ *such that*

(i) $A_1X^{(1)}A_2 - B_1X^{(1)}B_2 = Q$,

(ii) $A_1X^{(n)}A_2 - B_1X^{(n)}B_2$
$= Q - X^{(n-1)}DX^{(n-1)} - X^{(n-1)}EX^{(n-1)}FX^{(n-1)}$ *for all* $n > 1$,

(iii) $X^{(n-1)}$ *has the same range as* Q *has, for all* $n > 1$,

(iv) $\|X^{(1)}\|_{\mathscr{W}} \leqq \dfrac{1}{\beta}\|Q\|$,

(v) $\|X^{(2)}\|_{\mathscr{W}} < \dfrac{1}{\beta}\|Q\|\,T_2$, *where* $T_2 = 1 + \alpha_1 + \alpha_2 < \Delta$;

(vi) *for all* $n > 3$,
$\|X^{(n-1)}\|_{\mathscr{W}} \leqq \dfrac{1}{\beta}\|Q\|\,T_{n-1}$, *where* $T_{n-1} = 1 + \alpha_1T_{n-2}^2 + \alpha_2T_{n-2}^3 < \Delta$,

(vii) *for all* $n > 1$, *the right side of* (ii) *is non-zero and, consequently (keeping* (iii) *in view), has the same range as* Q *has,*

(viii) $\|X^{(2)} - X^{(1)}\|_{\mathscr{W}} < \dfrac{k_0(\alpha_1 + \alpha_2)}{\beta}\|Q\|$,

(ix) $\|X^{(i)} - X^{(i-1)}\|_{\mathscr{W}} < \dfrac{[k_0(\alpha_1 + \alpha_2)]^{i-1}}{\beta}\|Q\|$ *for all* $i \geqq 2$,

(x) $\|X^{(n-1)}\|_{\mathscr{W}} < \dfrac{1}{\beta}\|Q\|\min\left(\Delta, \dfrac{1}{1 - k_0(\alpha_1 + \alpha_2)}\right)$ *for all* $n \geqq 2$.

Proof: Let $X^{(1)} \in \overline{\mathscr{W}}_1$ be the solution of equation (4.1) obtained by applying Theorem 4.1. It is non-zero because Q is. Also, from Theorem 4.1 we have that

$$\|X^{(1)}\|_{\mathscr{W}} \leqq \frac{1}{\beta}\|Q\| \tag{4.2}$$

and $X^{(1)}$ has the same range as Q has. We can now see that $Q - X^{(1)}DX^{(1)} - X^{(1)}EX^{(1)}FX^{(1)}$ is non-zero, so that it has the same range as Q has. Indeed, if it were zero, then, using (4.2), (3.1) and (H 6)/(i),

$$\begin{aligned}\|Q\| &= \|X^{(1)}DX^{(1)} + X^{(1)}EX^{(1)}FX^{(1)}\|_{\mathfrak{X}}\\ &\leqq \gamma \|X^{(1)}\|_{\mathscr{W}}^2 \|D\| + \gamma \|X^{(1)}\|_{\mathscr{W}}^3 \|E\| \|F\| \leqq \|Q\| (\alpha_1 + \alpha_2) < \|Q\|,\end{aligned}$$

an inpossibility. This allows us to consider the second approximation $X^{(2)} \in \overline{\mathscr{W}}_1$ which is the solution given by Theorem 4.1 of the case $n = 2$ of equation (ii). It has the same range as Q has, and, as seen above and by (H 5),

$$\begin{aligned}\|X^{(2)}\|_{\mathscr{W}} &\leqq \frac{1}{\beta} \|Q - X^{(1)}DX^{(1)} - X^{(1)}EX^{(1)}FX^{(1)}\|_{\mathfrak{X}}\\ &\leqq \frac{1}{\beta} [\|Q\| + \|Q\| (\alpha_1 + \alpha_2)] = \frac{1}{\beta} \|Q\| T_2,\end{aligned}$$

$$T_2 = 1 + \alpha_1 + \alpha_2 < \Delta.$$

Let us note in passing that if $X^{(2)} = X^{(1)}$, then $X^{(2)}$ is a solution of equation (1.1), and we do not need to proceed any further. So we assume that $X^{(2)} \neq X^{(1)}$. To prove (viii), we subtract equation (i) from the case $n = 2$ of equation (ii). We get

$$A_1(X^{(2)} - X^{(1)}) A_2 - B_1(X^{(2)} - X^{(1)}) B_2 = -X^{(1)}DX^{(1)} - X^{(1)}EX^{(1)}FX^{(1)}. \tag{4.3}$$

Since $X^{(2)} - X^{(1)} \neq 0$, there exists $\Phi_{2,1} \in \mathcal{V}^2$ such that (see (H 4))

$$\begin{aligned}\beta \|X^{(2)} - X^{(1)}\|_{\mathscr{W}} \|\Phi_{2,1}\|_2 &< \|[A_1(X^{(2)} - X^{(1)}) A_2 - B_1(X^{(2)} - X^{(1)})B_2] \Phi_{2,1}|_{\mathscr{X}}\\ &\geqq \|A_1(X^{(2)} - X^{(1)}) A_1 - B_1(X^{(2)} - X^{(1)}) B_2\|_{\mathfrak{X}} \|\Phi_{2,1}\|_2,\end{aligned}$$

and so, using (4.3), (4.2), (3.1) and $k_0 \geqq 1$,

$$\begin{aligned}\beta \|X^{(2)} - X^{(1)}\|_{\mathscr{W}} &< \|X^{(1)}\|_{\mathfrak{X}} \|D\| \|X^{(1)}\|_{\mathscr{W}} + \|X^{(1)}\|_{\mathfrak{X}} \|E\| \|X^{(1)}\|_{\mathscr{W}} \|F\| \|X^{(1)}\|_{\mathscr{W}}\\ &\leqq \gamma \|X^{(1)}\|_{\mathscr{W}}^2 \|D\| + \gamma \|X^{(1)}\|_{\mathscr{W}}^3 \|E\| \|F\| \leqq \|Q\| (\alpha_1 + \alpha_2)\\ &\leqq k_0(\alpha_1 + \alpha_2) \|Q\|.\end{aligned}$$

Now we proceed with inductive reasoning. Suppose, the elements $X^{(1)}, X^{(2)}, \ldots, X^{(n-1)}$ have been constructed in the prescribed manner. By Theorem 4.1, equation (ii) has a unique non-zero solution $X^{(n)} \in \overline{\mathscr{W}}_1$ having the same range as Q has. If $X^{(n)} = X^{(n-1)}$, then this is a solution of equation (1.1), and we need not proceed any further. So, we assume that $X^{(n)} - X^{(n-1)} \neq 0$. By Theorem 4.1 again, using (2.1),

$$\begin{aligned}\|X^{(n)}\|_{\mathscr{W}} &\leqq \frac{1}{\beta} \|Q - X^{(n-1)}DX^{(n-1)} - X^{(n-1)}EX^{(n-1)}FX^{(n-1)}\|_{\mathfrak{X}}\\ &\leqq \frac{1}{\beta} [\|Q\| + \gamma \|X^{(n-1)}\|_{\mathscr{W}}^2 \|D\| + \gamma \|X^{(n-1)}\|_{\mathscr{W}}^3 \|E\| \|F\|]\\ &\leqq \frac{1}{\beta} \left[\|Q\| + \frac{\gamma}{\beta^2} \|Q\|^2 \|D\| T_{n-1}^2 + \frac{\gamma}{\beta^3} \|Q\|^3 \|E\| \|F\| T_{n-1}^3\right]\\ &= \frac{1}{\beta} \|Q\| T_n,\end{aligned}$$

where, by inductive hypotheses and (H 5),

$$T_n = 1 + \alpha_1 T_{n-1}^2 + \alpha_2 T_{n-1}^3 < 1 + \alpha_1 \Delta^2 + \alpha_2 \Delta^3 < \Delta.$$

We have thus proved that

$$\|X^{(n)}\|_{\mathscr{W}} < \frac{1}{\beta} \|Q\| \Delta \quad \text{for all } n \in \mathbf{N}. \tag{4.4}$$

Next, we assume that the inequality (ix) is true whenever $2 \leq i \leq n$. Before we prove that this inequality is true for $i = n + 1$ also, let us note that as a consequence of this inductive assumption, we have, for all $i = 1, 2, \ldots, n$,

$$\|X^{(i)}\|_{\mathscr{W}} \leq \sum_{j=2}^{i} \|X^{(j)} - X^{(j-1)}\|_{\mathscr{W}} + \|X^{(1)}\|_{\mathscr{W}}$$

$$< \frac{1}{\beta} \sum_{j=2}^{i} [k_0(\alpha_1 + \alpha_2)]^{j-1} \|Q\| + \frac{1}{\beta} \|Q\| < \frac{1}{\beta} \frac{1}{1 - k_0(\alpha_1 + \alpha_2)} \|Q\| \tag{4.5}$$

because $0 < k_0(\alpha_1 + \alpha_2) < 1$. Inequalities (4.4) and (4.5) together yield (X).

We next let $X_n = X^{(n)} - X^{(n-1)}$ and write, using Theorem 4.1,

$$A_1 X^{(n+1)} A_2 - B_1 X^{(n+1)} B_2 = Q - X^{(n)} D X^{(n)} - X^{(n)} E X^{(n)} F X^{(n)}$$

$$= Q - (X^{(n-1)} + X_n) D (X^{(n-1)} + X_n)$$

$$- (X^{(n-1)} + X_n) E (X^{(n-1)} + X^n) F (X^{(n-1)} + X_n).$$

Expanding the right hand side, and subtracting equation (ii), we get

$$A_1(X^{(n+1)} - X^{(n)}) A_2 - B_1(X^{(n+1)} - X^{(n)}) B_2$$

$$= -X^{(n-1)} D X_n - X_n D X^{(n-1)} - X_n D X_n - X^{(n-1)} E X^{(n-1)} F X_n$$

$$- X^{(n-1)} E X_n F X^{(n-1)} - X^{(n-1)} E X_n F X_n$$

$$- X_n E X^{(n-1)} F X^{(n-1)} - X_n E X^{(n-1)} F X_n$$

$$- X_n E X_n F X^{(n-1)} - X_n E X_n F X_n. \tag{4.6}$$

We assume that $X^{(n+1)} \neq X^{(n)}$, otherwise we would get a solution of equation (1.1) right away. By Theorem 4.1, $X^{(n+1)} \in \overline{\mathscr{W}}_1$. By (H 4), there exists a $\Phi_{n+1,n} \in \mathcal{V}^2$ such that

$$\beta \|X^{(n-1)} - X^{(n)}\|_{\mathscr{W}} \|\Phi_{n+1,n}\|_2$$

$$< |A_1(X^{(n+1)} - X^{(n)}) A_2 - B_1(X^{(n+1)} - X^{(n)}) B_2|_{\mathscr{X}} \|\Phi_{n+1,n}\|_2. \tag{4.7}$$

Since inequality (ix) is assumed to be true for $2 \leq i \leq n$, we have

$$\|X_n\|_{\mathscr{W}} < \frac{[k_0(\alpha_1 + \alpha_2)]^{n-1}}{\beta} \|Q\|.$$

Combining this with (4.5)—(4.7), we get

$$\beta \|X^{(n+1)} - X^{(n)}\|_{\mathscr{W}} < \gamma \left[\frac{2}{\beta^2} \frac{[k_0(\alpha_1 + \alpha_2)]^{n-1}}{1 - k_0(\alpha_1 + \alpha_2)} \|Q\|^2 \|D\| \right.$$

$$+ \frac{[k_0(\alpha_1 + \alpha_2)]^{2n-2}}{\beta^2} \|Q\|^2 \|D\| + \frac{3}{\beta^3} \frac{[k_0(\alpha_1 + \alpha_2)]^{n-1}}{[1 - k_0(\alpha_1 + \alpha_2)]^2} \|Q\|^3 \|E\| \|F\|$$

$$+\frac{3}{\beta^3}\frac{[k_0(\alpha_1+\alpha_2)]^{2n-2}}{1-k_0(\alpha_1+\alpha_2)}\|Q\|^3\,\|E\|\,\|F\|+\frac{[k_0(\alpha_1+\alpha_2)]^{3n-3}}{\beta^3}\|Q\|^3\,\|E\|\,\|F\|\Bigg]$$

$$=[k_0(\alpha_1+\alpha_2)]^{n-1}\,\|Q\|\left[\frac{2\alpha_1}{1-k_0(\alpha_1+\alpha_2)}+\alpha_1[k_0(\alpha_1+\alpha_2)]^{n-1}\right.$$

$$\left.+\frac{3\alpha_2}{[1-k_0(\alpha_1+\alpha_2)]^2}+\frac{3\alpha_2[k_0(\alpha_1+\alpha_2)]^{n-1}}{1-k_0(\alpha_1+\alpha_2)}+\alpha_2[k_0(\alpha_1+\alpha_2)]^{2n-2}\right]$$

$$\leqq[k_0(\alpha_1+\alpha_2)]^{n-1}\,\|Q\|\,k_0(\alpha_1+\alpha_2)$$

because of the hypotheses (H 6)/(i) and (ii). This completes the inductive proof of (ix).

To complete the proof of the Lemma, it remains to show that

$$Q-X^{(n)}DX^{(n)}-X^{(n)}EX^{(n)}FX^{(n)}\neq 0. \tag{4.8}$$

If that were not the case, we would have

$$\|Q\|=\|X^{(n)}DX^{(n)}+X^{(n)}EX^{(n)}FX^{(n)}\|_{\mathfrak{X}}$$
$$\leqq\gamma\,\|X^{(n)}\|_{\mathscr{W}}^2\,\|D\|+\gamma\,\|X^{(n)}\|_{\mathscr{W}}^3\,\|E\|\,\|F\|,$$

whence we conclude both (I) and (II) below:

(I) By (4.4), (3.1) and (H 5), $\|Q\|<\|Q\|\,(\alpha_1\Delta^2+\alpha_2\Delta^3)<\|Q\|\,(\Delta-1)$, which yields $\Delta>2$.

(II) By (4.5) and (3.1), $\|Q\|<\dfrac{\alpha_1\,\|Q\|}{[1-k_0(\alpha_1+\alpha_2)]^2}+\dfrac{\alpha_2\,\|Q\|}{[1-k_0(\alpha_1+\alpha_2)]^3}\leqq\|Q\|$.

This is impossible. Hence (4.8) is true ∎

We continue with the proof of Theorem 3.4. Indeed, the rest of the proof is very easy as pointed out by the reviewers of this paper, to whom the author's thanks are due:

Since $\mathcal{V}^1$ is complete, so is $\mathscr{W}$. Therefore, $\overline{\mathscr{W}}_1$ is complete. Next we observe that (ix) of Lemma 4.2 together with (i) if (H 6) imply that the sequence $\{X^{(i)}\}$ is Cauchy in $\overline{\mathscr{W}}_1$, and so converges to some $X_Q\in\overline{\mathscr{W}}_1$ in the norm operator topology. Now, in (ii) of Lemma 4.2 we pass to the limit in the norm operator topology as $n\to\infty$, and end up with

$$A_1X_QA_2-B_1X_QB_2=Q-X_QDX_Q-X_QEX_QF_Q,$$

solving equation (1.1). Furthermore, it is clear from (X) of Lemma 4.2 that

$$\|X_Q\|_{\mathscr{W}}\leqq\frac{1}{\beta}\|Q\|\min\left[\Delta,\frac{1}{1-k_0(\alpha_1+\alpha_2)}\right] \quad ∎$$

§ 5 Examples

Our first example is taken from [2] where a solution is provided for the finite-dimensional equation (1.1) with

$$A_1=B_2=\begin{bmatrix}2&0\\0&-1\end{bmatrix},\quad A_2=-B_1=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

$$D=-E=-\begin{bmatrix}1&1\\1&1\end{bmatrix},\quad F=\begin{bmatrix}1&0\\1&1\end{bmatrix},\quad Q=\begin{bmatrix}0&0\\0&-2\end{bmatrix}.$$

In this example Q has one-dimensional range. We can now pose the question: For which Q's of the from $\begin{bmatrix} 0 & 0 \\ 0 & q \end{bmatrix}$ does the aforementioned example have a solution? The results of the previous section provide us with at least a partial answer, namely that a solution exists for all sufficiently small $|q|$ consistent with the hypotheses (H 5) and (H 6). We are going to see below how this conclusion is arrived at. As a matter of convenience we take $\mathcal{V}^1 = \mathcal{V}^2 = \mathcal{H} =$ the space of all two-dimensional vectors. (We could have taken, though, $\mathcal{H} =$ the sequence space l_2 with basis $\{b_i\}_{i\in\mathrm{N}}$ given by $b_1 = (1, 0, \ldots)$, $b_2 = (0, 1, 0, \ldots)$ etc., and then redefine $A_1, A_2, B_1, B_2, D, E, F, Q$ by extending them by zero, e.g. define A_1 by $A_1b_1 = 2b_1$, $A_1b_2 = -b_2$, $A_1b_n = 0$ for all $n \geqq 3$, etc.)

Clearly, conditions (H 1) and (H 2) are satisfied with $h = [0, 1]^{\mathrm{T}}$. Condition (H 3) is also satisfied with $b_1 = [1, 0]^{\mathrm{T}}$, $b_2 = [0, 1]^{\mathrm{T}}$. To see that (H 4) is satisfied, we note that a $Y \in \mathcal{W}_h$ is of the form $Y = \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}$. If Y is non-zero, then choosing $\Phi_Y = [c_1 - d]^{\mathrm{T}}$ we have $|(A_1YA_2 - B_1YB_2)\,\Phi_Y|_{\mathcal{H}} = c^2 + 2d^2$, and $\|Y\|\,\|\Phi_Y\|_2 = \sqrt{c^2 + d^2}\sqrt{c^2 + d^2} = c^2 + d^2$, so that (H_4) is satisfied with any $\beta < 1$. In this example, $\gamma = 1$. Any choise of D, E, F, q consistent with the hypotheses (H 5) and (H 6) will give us an equation possessing solutions.

Our next example is the infinite-dimensional example given in [4], which we reproduce here. Let $\mathcal{H} = L^2(\Omega; \mathrm{C})$, where $\Omega = (0,2\pi) \times (0,2\pi)$. $\mathcal{H}$ is a Hilbert space under the inner product,

$$(u, v)_{\mathcal{H}} = \int_0^{2\pi}\int_0^{2\pi} u(x, y)\,\overline{v(x, y)}\,dx\,dy,$$

where $\overline{v(x, y)}$ is the complex conjugate of $v(x, y)$. Let $e_{i,j}(x, y) = e_i(x)\,e_j(y)$ where $e_0(\xi) = (2\pi)^{-1/2}$, $e_{2n-1}(\xi) = \pi^{-1/2}\sin n\xi$, $e_{2n}(\xi) = \pi^{-1/2}\cos n\xi$ for all $n \in \mathrm{N}$. Let $b_{p,q} = \gamma_{p,q}e_{p,q}$ where $\gamma_{p,q}$ are constants so chosen as to make $\mathcal{B} = \{b_{p,q} : p \in \mathrm{N}, q \in \mathrm{N}\}$ an orthonormal set in the Sobolev space $\mathcal{H}^3$. Let us recall that if ν is a positive integer, then $\mathcal{H}^\nu$ is the set of all those elements of $\mathcal{H}$ whose distributional derivatives of all order not exceeding ν are again elements of $\mathcal{H}$. $\mathcal{H}^\nu$ is known to be a Hilbert space under the inner product

$$\big((u, v)\big)_{\mathcal{H}^\nu} = \sum_{\substack{m,n=0\\ m+n\leqq\nu}}^{\nu} (\partial_x^m\,\partial_y^n u, \partial_x^m\,\partial_y^n v)_{\mathcal{H}},$$

where ∂_x, ∂_y are the distributional derivatives. Let $\mathcal{V}^2$ be the subspace of $\mathcal{H}^3$ consisting of all finite linear combinations of the $b_{p,q}$'s. Let $\mathcal{V}^1 = \mathcal{H}_0^2 =$ the set of all those elements of $\mathcal{H}^2$ which, together with their first distributional derivatives, vanish at the boundary of Ω. $\mathcal{H}_0^2$ is a Hilbert space under the same inner product under which $\mathcal{H}^2$ is a Hilbert space. The γ of inequality (2.1) may obviously be taken to be 1. We now construct the example by setting

$$A_1u = -a\,\partial_x^2u + k_1u, \quad \text{where} \quad a > 0, k_1 > 0,$$

$$B_1u = -b\,\partial_y^2, \quad \text{where} \quad b > 0,$$

$$A_2v = \partial_x^2v + \partial_y^2v - \frac{1}{2}v,$$

$$B_2v = \partial_x^4v + \partial_y^4v + k_2v, \quad \text{where} \quad k_2 > 0,$$

for all $u \in \mathcal{V}^1$, $v \in \mathcal{V}^2$. This is typical of a class of similarly constructed examples to which our methods may be applied.

Condition (H 3) is clearly satisfied.

As an example of a $Q \in \mathcal{X}$ with one-dimensional range, we may take the one defined by $Q\Phi = d_{5,3}h$ where $h \in \mathcal{H}$ is given by $h(x, y) = e_5(x)\, e_3(y)$, and

$$(\partial_y{}^2\Phi)(x, y) = \sum_{p,q=0}^{\infty} d_{p,q} b_{p,q}(x, y) \quad \text{with} \quad d_{p,q} \in \mathbf{C},\ \Phi \in \mathcal{V}^2.$$

Clearly, (H 1) and (H 2) are satisfied, and $\|Q\| \leqq 1$.

Let us now verify condition (H 4). Take an arbitrary non-zero $Y \in \mathcal{W}_h$. Suppose $Yb_{p,q} = \xi_{p,q}h$ with $\xi_{p,q} \in \mathbf{C}$. Then, $\|Y\| \geqq \|Yb_{p,q}\|_1 = |\xi_{p,q}|\, \|h\|_1$. Let $\bar{\xi}_{p,q}$ denote the complex conjugate of $\xi_{p,q}$. Let r, s be arbitrary positive integers. In what follows we write $\sum\limits^{r,s}$ to mean $\sum\limits_{p=0}^{r} \sum\limits_{q=0}^{s}$ and $\sum\limits^{\infty,\infty}$ to mean $\sum\limits_{p=0}^{\infty} \sum\limits_{q=0}^{\infty}$. If $v = \sum\limits^{r,s} \bar{\xi}_{p,q} b_{q,q} \in \mathcal{V}^2$, then $Yv = \sum\limits^{r,s} |\xi_{p,q}|^2\, h$, and so

$$\|Y\|_{\mathcal{W}} \geqq \frac{\|Yv\|_1}{\|v\|_2} \left[\sum^{r,s} |\xi_{p,q}|^2\right]^{1/2} \|h\|_1. \tag{5.1}$$

This is true for all r and all s. Thus, $\sum\limits^{r,s} |\xi_{p,q}|^2$ converges as $r \to \infty$, $s \to \infty$, and $\|Y\|_{\mathcal{W}} \geqq \left(\sum\limits^{\infty,\infty} |\xi_{p,q}|^2\right)^{1/2} \|h\|_1$. Noting that

$$\|Y\|_{\mathcal{W}} = \sup_{\substack{w \in \mathcal{V}^2 \\ w \neq 0}} \frac{\left|\sum\limits^{r,s} \xi_{p,q}\eta_{p,q}\right| \|h\|_1}{\left(\sum\limits^{r,s} |\eta_{p,q}|^2\right)^{1/2}} \leqq \left[\sum^{\infty,\infty} |\xi_{p,q}|^2\right]^{1/2} \|h\|_1$$

where $w = \sum\limits^{r,s} \eta_{p,q} b_{p,q}$ is an arbitrary element of $\mathcal{V}^2$, we see from (5.1) that $\|Y\|_{\mathcal{W}} = \left[\sum\limits^{\infty,\infty} |\xi_{p,q}|^2\right]^{1/2} \|h\|_1$. If $0 < \varepsilon < 1$, then finite positive integers r, s exist such that

$$\varepsilon\, \|Y\|_{\mathcal{W}} = \varepsilon \left(\sum^{\infty,\infty} |\xi_{p,q}|^2\right)^{1/2} \|h\|_1 < \left(\sum^{r,s} |\xi_{p,q}|^2\right)^{1/2} \|h\|_1.$$

With this r, s, let $v_0 = \sum\limits^{r,s} \bar{\xi}_{p,q} b_{p,q}$. Then, exactly as in [4: Section 3], we get

$$\begin{aligned} &|(A_1YA_2 - B_1YB_2)\, v_0|_{\mathcal{H}} \\ &\geqq \left[\frac{1}{2}(9a + k_1) + 4bk_2\right] \left(\sum^{r,s} |\xi_{p,q}|^2\right)^{1/2} \left(\sum^{r,s} |\xi_{p,q}|^2\right)^{1/2} \\ &> \left[\frac{1}{2}(9a + k_1) + 4bk_2\right] \frac{\varepsilon\, \|Y\|_{\mathcal{W}}}{\|h\|_1} \|v_0\|_2. \end{aligned}$$

So, we may take $\beta = \varepsilon[2^{-1}(9a + k_1) + 4bk_2]/\|h\|_1$ for a convenient ε, and then (H 4) is satisfied.

We also note that

$$\frac{\gamma}{\beta^2} \|Q\| \leqq \frac{\|h\|_1{}^2}{\varepsilon^2 \left[\frac{1}{2}(9a + k_1) + 4bk_2\right]^2} \quad \text{and} \quad \frac{\gamma}{\beta^3} \|Q\|^2 \leqq \frac{\|h\|_1{}^3}{\varepsilon^3 \left[\frac{1}{2}(9a + k_1) + 4bk_2\right]^3}.$$

Thus, if a, b, k_1 or k_2 are chosen large enough, or D, E, F are chosen with their magnitudes small enough, then hypotheses (H 5) and (H 6) will be satisfied, and the operator Riccati-type equation

$$(-a\,\partial_x^2 + k_1 I)\,X\left(\partial_x^2 + \partial_y^2 - \frac{1}{2}I\right) + (b\,\partial_y^2)\,X(\partial_x^4 + \partial_y^4 + k_2 I)$$
$$+ XDX + XEXFX = Q,$$

with I representing the identity operator, has a solution X whose norm does not exceed the right member of the inequality (X) of Lemma 4.2. The actual values of Δ, α_1, α_2 will depend on what exactly D, E, F are. The constant k_0 plays no essential role — it is retained solely for possible computational advantage in a numerical situation.

REFERENCES

[1] Bellman, R., Ueno, S., und R. Vasudevan: On the matrix Riccati equations of transport processes. J. Math. Anal. Appl. **44** (1973), 472–481.

[2] Jones, J. Jr., Martin, C. R., and H. A. Undem: Solutions of higher-order Riccati-type matrix equations. Int. J. Systems Sci. **15** (1984), 1001–1010.

[3] Lions, J. L.: Optimal Control of Systems Governed by Partial Differential Equations. Berlin—Heidelberg—New York: Springer-Verlag 1971.

[4] Mazumdar, T.: On the operator equation $A_1XA_2 - B_1XB_2 = Q$ when Q has one-dimensional range. ZAMM **66** (1986), 443–444.

[5] Nelson, P., Ray, A. K., and G. M. Wing: On the effectiveness of the inverse Riccati transformation in the matrix case. J. Math. Anal. Appl. **65** (1978), 201–210.

[6] Rosza, P.: Lineare Matrizengleichungen und Kroneckersche Produkte. ZAMM **58** (1978), T395–T397.

[7] Vasudevan, R.: On a generalization of the matrix and scalar Riccati equations. J. Math. Anal. Appl. **45** (1974), 629–638.

Manuskripteingang: 12. 11. 1985; in revidierter Fassung 28. 05. 1986

VERFASSER:

Prof. Dr. Tapas Mazumdar
Department of Mathematics and Statistics
Wright State University
Dayton, Ohio 45435, USA