# On the existence of solutions to the operator Riccati equation and the tan $\Theta$ theorem

Vadim Kostrykin

*Fraunhofer-Institut für Lasertechnik, Steinbachstraße 15, D-52074, Aachen, Germany*
*E-mail: kostrykin@t-online.de, kostrykin@ilt.fraunhofer.de*

Konstantin A. Makarov

*Department of Mathematics, University of Missouri, Columbia, MO 65211, USA*
*E-mail: makarov@math.missouri.edu*

Alexander K. Motovilov [1]

*Department of Mathematics, University of Missouri, Columbia, MO 65211, USA*
*E-mail: motovilv@thsun1.jinr.ru*

**Abstract**

Let $A$ and $C$ be self-adjoint operators such that the spectrum of $A$ lies in a gap of the spectrum of $C$ and let $d > 0$ be the distance between the spectra of $A$ and $C$. We prove that under these assumptions the sharp value of the constant $c$ in the condition $\|B\| < cd$ implying the solvability of the operator Riccati equation $XA - CX + XBX = B^*$ is equal to $\sqrt{2}$. We also prove an extension of the Davis-Kahan $\tan\Theta$ theorem.

*Key words:* invariant subspaces, operator Riccati equation, operator matrix, operator angle, factorization theorem
*2000 MSC:* 47A15, 39B42, 47A56

## 1 Introduction

In this paper we consider the operator Riccati equation

$$XA - CX + XBX = B^* \tag{1.1}$$

[1] On leave of absence from the Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia

associated with a $2 \times 2$ block operator matrix

$$\mathbf{H} = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \tag{1.2}$$

on the orthogonal sum $\mathcal{H} = \mathcal{H}_A \oplus \mathcal{H}_C$ of separable Hilbert spaces. Here $A$ is a bounded self-adjoint operator on the Hilbert space $\mathcal{H}_A$, $C$ possibly unbounded self-adjoint operator on the Hilbert space $\mathcal{H}_C$, and $B$ is a bounded operator from $\mathcal{H}_C$ to $\mathcal{H}_A$.

Solving the Riccati equation appears to be an adequate tool in the study of the invariant subspaces of the operator $\mathbf{H}$ which are the graphs of bounded operators from $\mathcal{H}_A$ to $\mathcal{H}_C$. More precisely, given a bounded solution $X$ to the Riccati equation (1.1) such that $\operatorname{Ran} X \subset \operatorname{Dom}(C)$, the graph

$$\mathcal{G}(X) = \{x \oplus Xx |\, x \in \mathcal{H}_A\}$$

of the operator $X$ reduces the operator $\mathbf{H}$. In the framework of this approach the following two problems naturally arise. The first problem is to study the spectrum of the part of $\mathbf{H}$ associated with the reducing subspace $\mathcal{G}(X)$ (respectively its orthogonal complement $\mathcal{G}(X)^\perp$), and the second one is to estimate the operator angle $\Theta$ (see, e.g., [11] for discussion of this notion) between the subspaces $\mathcal{G}(X)$ and $\mathcal{H}_A$ (respectively $\mathcal{G}(X)^\perp$ and $\mathcal{H}_C$). Both of these problems can efficiently be solved if a bounded solution $X$ to (1.1) is known: the operator $\mathbf{H}$ appears to be similar to the diagonal block operator matrix

$$\begin{pmatrix} A + BX & 0 \\ 0 & C - B^*X^* \end{pmatrix}$$

associated with the decomposition $\mathcal{H} = \mathcal{H}_A \oplus \mathcal{H}_C$, and the operator angle $\Theta$ between the subspaces $\mathcal{G}(X)$ and $\mathcal{H}_A$ has the representation

$$\Theta = \arctan \sqrt{X^*X}.$$

If the spectra of $A$ and $C$ overlap, then the Riccati equation (1.1) may have no solution at all (cf., e.g., [3, Example 3.2]). At the same time the spectra separation condition alone does not guarantee the existence of solutions either. Under the spectra separation hypothesis

$$\operatorname{dist}(\sigma(A), \sigma(C)) > 0,$$

a natural sufficient condition for the existence of solutions to the Riccati equation (1.1) requires a smallness assumption on the operator $B$ of the form

$$\|B\| < c \operatorname{dist}\Big(\sigma(A), \sigma(C)\Big) \tag{1.3}$$

with a constant $c > 0$ independent of the distance between the spectra $\sigma(A)$ and $\sigma(C)$ of the operators $A$ and $C$, respectively. The best possible constant $c_{\rm best}$ in (1.3) is still unknown. However, $c_{\rm best}$ is known to be within the interval $\left[\pi^{-1}, \sqrt{2}\right]$ (see [3]) and for bounded $A$ and $C$ to satisfy $c_{\rm best} \in \left[\frac{3\pi-\sqrt{\pi^2+32}}{\pi^2-4}, \sqrt{2}\right]$ (see [13]). In [13] the best possible constant $c_{\rm best}$ has been conjectured to be $\sqrt{3}/2$. Some earlier results in this direction can be found in [2], [17], [19], [20].

The best possible constant in inequality (1.3) may be different under additional assumptions upon mutual disposition of the spectra of $A$ and $C$. For instance, if the spectra of $A$ and $C$ are subordinated, e.g.,

$$\sup \sigma(A) < \inf \sigma(C), \tag{1.4}$$

the Riccati equation (1.1) is known to have a strictly contractive solution for any bounded $B$ (see, e.g., [1]). To some extent abusing the terminology one may say that in this case the best possible constant in inequality (1.3) is infinite: No smallness assumptions on $B$ are needed.

In the limiting case of (1.4),

$$\sup \sigma(A) = \inf \sigma(C),$$

the existence of contractive solutions has been established in [2] under some additional assumptions which has been dropped in [12]. See also [18] where the spectra separation condition has been somewhat relaxed and the existence of a bounded but not necessarily contractive solution has been established.

Our *first principal result* concerns the case where the operator $C$ has a finite spectral gap containing the spectrum of $A$. Recall that by a finite spectral gap of a self-adjoint operator $T$ one understands an *open* finite interval on the real axis lying in the resolvent set of $T$ such that both of its end points belong to the spectrum of $T$.

**Theorem 1.** Assume that the self-adjoint operator $C$ has a finite spectral gap $\Delta$ containing the spectrum of bounded self-adjoint operator $A$. Assume in addition that

$$\|B\| < \sqrt{d|\Delta|} \quad \text{where} \quad d = \operatorname{dist}(\sigma(A), \sigma(C)), \tag{1.5}$$

with $|\cdot|$ denoting Lebesgue measure on $\mathbb{R}$. Then the spectrum of the operator $\mathbf{H}$ in the gap $\Delta$ is a proper closed subset of $\Delta$. The spectral subspace of the operator $\mathbf{H}$ associated with the interval $\Delta$ is the graph of a bounded solution $X : \mathcal{H}_A \to \mathcal{H}_C$ to the Riccati equation (1.1). The operator $X$ is the unique solution to the Riccati equation in the class of bounded operators possessing the properties

$$
\begin{aligned}
&\sigma(A+BX) \subset \Delta, \\
&\sigma(C-B^*X^*) \subset \mathbb{R}\setminus\Delta, \quad \mathrm{Dom}(C-B^*X^*) = \mathrm{Dom}(C).
\end{aligned}
\tag{1.6}
$$

Moreover, under the assumption of the theorem that the operator $C$ has a finite spectral gap $\Delta$ containing the spectrum of $A$, we prove that $c=\sqrt{2}$ is best possible in (1.3) ensuring the existence of a bounded solution to the Riccati equation (1.1) (see Remark 3.5).

Our *second principal result* holds with no assumptions upon the mutual disposition of the spectra of $A$ and $C$ (in particular the spectra of $A$ and $C$ may overlap).

**Theorem 2.** Assume that the self-adjoint operator $C$ has a spectral gap $\Delta$ (finite or infinite) and the self-adjoint operator $A$ is bounded. Assume that the Riccati equation (1.1) has a bounded solution $X$ and hence the graph subspace $\mathcal{G}(X)$ reduces the block operator matrix $\mathbf{H}$. Suppose that the spectrum of the part $\mathbf{H}|_{\mathcal{G}(X)}$ of the operator $\mathbf{H}$ associated with the reducing subspace $\mathcal{G}(X)$ is a closed subset of $\Delta$. Then the operator $X$ satisfies the norm estimate

$$
\|X\| \le \frac{\|B\|}{\delta} \quad \text{with } \delta = \mathrm{dist}\Big(\sigma(\mathbf{H}|_{\mathcal{G}(X)}), \sigma(C)\Big). \tag{1.7}
$$

Equivalently,

$$
\|\tan\Theta\| \le \frac{\|B\|}{\delta}, \tag{1.8}
$$

where $\Theta$ is the operator angle between the subspaces $\mathcal{H}_A$ and $\mathcal{G}(X)$.

Estimate (1.8) generalizes the Davis-Kahan $\tan\Theta$ theorem [8] previously known only in the case where the operator $C$ is semibounded and the spectrum of the part $\mathbf{H}|_{\mathcal{G}(X)}$ lies in the *infinite* spectral gap of $C$, i.e., the spectra of $C$ and $\mathbf{H}|_{\mathcal{G}(X)}$ are subordinated. This generalization extends the list of the celebrated $\sin\Theta$ and $\sin 2\Theta$ theorems, proven in the case where the operator $C$ has a gap of finite length [8].

Our main techniques are based on applications of the Virozub-Macaev factorization theorems for analytic operator-valued functions [22] in the spirit of the work [18] and the Daletsky-Krein factorization formula [5]. Under the hypothesis of Theorem 1 we prove that

- for $\lambda \notin \sigma(C)$ the operator-valued Herglotz function $M(\lambda) = \lambda I - A + B(C-\lambda I)^{-1}B^*$ admits a factorization

$$
M(\lambda) = W(\lambda)(Z-\lambda I) \tag{1.9}
$$

with $W$ being an operator-valued function holomorphic on the resolvent set of the operator $C$ and $Z$ a bounded operator with the spectrum in the spectral gap $\Delta$ of the operator $C$,

- the Riccati equation (1.1) has a bounded solution given by

$$X = -\frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (C - \zeta I)^{-1} B^* (Z - \zeta I)^{-1},$$

  where $\Gamma$ is an appropriate Jordan contour encircling the spectrum of the operator $Z$,
- the spectral subspace of the $2 \times 2$ operator matrix $\mathbf{H}$ (1.2) associated with the interval $\Delta$ is a graph of the operator $X$,
- the spectrum of the operator $\mathbf{H}$ in the interval $\Delta$ coincides with that of the operator $Z$, i.e., $\sigma(H) \cap \Delta = \sigma(Z)$.

In Section 2 we prove factorization formula (1.9) and give bounds on the location of the spectrum of the operator $Z$ (see Theorem 2.4). In Section 3 we prove Theorem 1. Finally, in Section 4 we prove Theorem 2 and, combining the results of Theorems 1 and 2, under condition (1.5) we estimate the operator angle between the spectral subspaces of the operators $\mathbf{H}_0 = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}$ and $\mathbf{H} = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}$ corresponding to the interval $\Delta$ (see Corollary 4.1).

## 2 Factorization theorems for operator-valued Herglotz functions

The principal purpose of this section is to obtain factorization results for a class of operator-valued Herglotz functions associated with the Riccati equation (1.1).

Given a Hilbert space $\mathcal{K}$ by $I_\mathcal{K}$ we denote the identity operator on $\mathcal{K}$. If it does not lead to any confusion we will simply write $I$ instead of more pedantic notation $I_\mathcal{K}$. The set of all linear bounded operators on $\mathcal{K}$ will be denoted by $\mathcal{B}(\mathcal{K})$. We also adopt the following notation. Let $K$ and $L$ be self-adjoint operators on a Hilbert space $\mathcal{K}$. We say $K < L$ (or, equivalently, $L > K$) if there is a number $\gamma > 0$ such that $L - K > \gamma I$. If $H$ is a closed operator on a Hilbert space $\mathcal{K}$, by $\rho(H)$ we denote the resolvent set of $H$.

For notational setup we assume the following

**Hypothesis 2.1.** *Assume that $A$ is a bounded self-adjoint operator on a Hilbert space $\mathcal{H}_A$ and $C$ possibly unbounded self-adjoint operator with domain $\mathrm{Dom}(C)$ on a Hilbert space $\mathcal{H}_C$. Suppose that $C$ has a spectral gap $\Delta = (\alpha, \beta)$, $\alpha < \beta$ and* *the spectrum of $A$ lies in $\Delta$, i.e., $\sigma(A) \subset \Delta$, and let*

$$d = \operatorname{dist}\Big(\sigma(A), \sigma(C)\Big) > 0.$$

*Assume in addition that $B$ is a bounded operator from $\mathcal{H}_C$ and $\mathcal{H}_A$.*

Under Hypothesis 2.1 introduce the operator-valued Herglotz function

$$M(\lambda) = \lambda I - A + B(C - \lambda I)^{-1}B^*, \quad \lambda \in \rho(C). \tag{2.1}$$

By definition the spectrum $\sigma(M)$ of the function $M$ coincides with the set of all $\lambda \in \mathbb{C}$ such that either the operator $M(\lambda)$ is not invertible or the inverse $[M(\lambda)]^{-1}$ is an unbounded operator.

**Lemma 2.2.** *Under Hypothesis 2.1 the function $M(\lambda)$ given by* (2.1) *is holomorphic in $\rho(C)$. In particular, the holomorphy domain of $M$ contains the interval $\Delta$ and*

$$\frac{d}{d\lambda} M(\lambda) > 0 \quad \textit{for} \quad \lambda \in \Delta. \tag{2.2}$$

*Assume in addition that*

$$\|B\| < \sqrt{d|\Delta|}. \tag{2.3}$$

*Then*

$$M(\lambda) < 0 \quad \textit{for} \quad \lambda \in \Big(\alpha, \inf \sigma(A) - \delta_B^-\Big) \tag{2.4}$$

*and*

$$M(\lambda) > 0 \quad \textit{for} \quad \lambda \in \Big(\sup \sigma(A) + \delta_B^+, \beta\Big), \tag{2.5}$$

*where*

$$\delta_B^- = \|B\| \tan\left(\frac{1}{2} \arctan \frac{2\|B\|}{\beta - \inf \sigma(A)}\right) < \inf \sigma(A) - \alpha, \tag{2.6}$$

$$\delta_B^+ = \|B\| \tan\left(\frac{1}{2} \arctan \frac{2\|B\|}{\sup \sigma(A) - \alpha}\right) < \beta - \sup \sigma(A). \tag{2.7}$$

*Proof.* By the spectral theorem

$$B(C - \lambda I)^{-1}B^* = \int\limits_{\mathbb{R}\setminus\Delta} B\mathsf{E}_C(d\mu)B^* \frac{1}{(\mu - \lambda)^2}, \quad \lambda \in \rho(C),$$

where $\mathsf{E}_C(\mu)$ stands for the spectral family of the self-adjoint operator $C$. Hence

$$\frac{d}{d\lambda}M(\lambda) = I + \int\limits_{\mathbb{R}\setminus\Delta} B\mathsf{E}_C(d\mu)B^*\frac{1}{(\mu-\lambda)^2}, \qquad \lambda \in \rho(C). \tag{2.8}$$

For $\lambda \in \Delta$ the integral in (2.8) is a non-negative operator. Therefore,

$$\frac{d}{d\lambda}M(\lambda) \geq I, \quad \lambda \in \Delta,$$

proving (2.2).

Next we estimate the quadratic form of $M(\lambda)$. Let $f \in \mathcal{H}_A$, $\|f\| = 1$. Then

$$\begin{aligned}
\langle M(\lambda)f, f\rangle &= \lambda - \langle Af, f\rangle + \langle (C-\lambda)^{-1}B^*f, B^*f\rangle \\
&= \lambda - \langle Af, f\rangle + \int\limits_{-\infty}^{\alpha} \frac{1}{\mu-\lambda}\langle \mathsf{E}_C(d\mu)B^*f, B^*f\rangle \\
&\quad + \int\limits_{\beta}^{+\infty} \frac{1}{\mu-\lambda}\langle \mathsf{E}_C(d\mu)B^*f, B^*f\rangle, \quad \lambda \in \rho(C).
\end{aligned} \tag{2.9}$$

Since for $\lambda \in \Delta$ the integral in the second line of (2.9) is non-positive and that in the third line is non-negative, one obtains the two-sided estimate

$$\left(\lambda - \sup\sigma(A) - \frac{\|B\|^2}{\lambda-\alpha}\right)I \leq M(\lambda) \leq \left(\lambda - \inf\sigma(A) - \frac{\|B\|^2}{\lambda-\beta}\right)I$$

for any $\lambda \in \Delta$. Now, a simple calculation shows that (2.4) and (2.5) hold, completing the proof. $\square$

For convenience of the reader we reproduce here the Virozub-Matsaev factorization theorem [22] (also see [16]). We present the statement following Propositions 1.1 and 1.2 of [18].

**Theorem 2.3.** *Let $\mathcal{K}$ be a Hilbert space and $F(\lambda)$ a holomorphic $\mathcal{B}(\mathcal{K})$-valued function on a simply connected domain $\Omega \subset \mathbb{C}$. Assume that $\Omega$ includes an interval $[a,b] \subset \mathbb{R}$ such that*

$$F(a) < 0, \quad F(b) > 0, \quad \textit{and} \quad \frac{d}{d\lambda}F(\lambda) > 0 \quad \textit{for all} \quad \lambda \in [a,b].$$

*Then there are a domain $\widetilde{\Omega} \subset \Omega$ containing $[a,b]$ and a unique bounded operator $Z$ on $\mathcal{K}$ with $\sigma(Z) \subset (a,b)$ such that $F(\lambda)$ admits the factorization*

$$F(\lambda) = G(\lambda)(Z - \lambda I), \quad \lambda \in \widetilde{\Omega},$$

*where $G(\lambda)$ is a holomorphic operator-valued function on $\widetilde{\Omega}$ whose values are bounded and boundedly invertible operators in $\mathcal{K}$, that is,*

$$G(\lambda) \in \mathcal{B}(\mathcal{K}) \quad \textit{and} \quad [G(\lambda)]^{-1} \in \mathcal{B}(\mathcal{K}) \quad \textit{for} \quad \lambda \in \widetilde{\Omega}.$$

As a direct corollary of Theorem 2.3, under Hypothesis 2.1 we get the following factorization result for Herglotz functions of the form (2.1).

**Theorem 2.4.** *Assume hypothesis of Lemma 2.2 and let $\delta_B^{\pm}$ be given by* (2.6) *and* (2.7). *Then there is a unique operator $Z \in \mathcal{B}(\mathcal{H}_A)$ with*

$$\sigma(Z) \subset [\inf \sigma(A) - \delta_B^-, \sup \sigma(A) + \delta_B^+] \tag{2.10}$$

*such that the function* (2.1) *admits the factorization*

$$M(\lambda) = W(\lambda)(Z - \lambda I), \quad \lambda \in \rho(C). \tag{2.11}$$

*Here $W$ is a holomorphic $\mathcal{B}(\mathcal{H}_A)$-valued function on $\rho(C)$ such that for any*

$$\lambda \in \Big(\mathbb{C} \setminus \sigma(M)\Big) \cup \Delta$$

*the operator $W(\lambda)$ has a bounded inverse.*

*Proof.* By Lemma 2.2, the function $F(\lambda) = M(\lambda)$ satisfies assumptions of Theorem 2.3 for any $a \in (\alpha, \inf \sigma(A) - \delta_B^-)$ and any $b \in (\sup \sigma(A) + \delta_B^+, \beta)$, proving the existence of the unique bounded operator $Z \in \mathcal{B}(\mathcal{H}_A)$ satisfying (2.11) and (2.10). Theorem 2.3 also states that the factor $W(\lambda)$ in (2.11) has a bounded inverse in a complex neighborhood $U \subset \mathbb{C}$ of the interval $[\inf \sigma(A) - \delta_B^-, \sup \sigma(A) + \delta_B^+]$. The last assertion of the theorem then follows from the representation

$$[W(\lambda)]^{-1} = (Z - \lambda I)[M(\lambda)]^{-1}, \quad \lambda \in U \setminus \sigma(M)$$

by analytic continuation to the domain $\Big(\mathbb{C} \setminus \sigma(M)\Big) \cup \Delta$. □

We will call the operator $Z$ referred to in Theorem 2.4 the operator root of the Herglotz function $M$ (cf. [22, Remark 4]).

## 3 Existence and uniqueness results. Proof of Theorem 1

We start this section by an existence result for the operator Riccati equation (1.1).

Under Hypothesis 2.1, a bounded operator $X$ from $\mathcal{H}_A$ to $\mathcal{H}_C$ is said to be an operator solution to (1.1) if $\mathrm{Ran}(X) \subset \mathrm{Dom}(C)$ and (1.1) holds as an operator equality.

**Theorem 3.1.** *Assume Hypothesis 2.1 and suppose that*

$$\|B\| < \sqrt{d|\Delta|}.$$

*Then the operator Riccati equation* (1.1) *has a bounded operator solution.*

*Proof.* Let $Z$ be the operator root of the operator-valued function $M(\lambda)$ (2.1) referred to in Theorem 2.4 and $\Gamma$ an arbitrary bounded Jordan contour encircling the spectrum of $Z$ in the clockwise direction with winding number $0$ with respect to the spectrum of the operator $C$. Put

$$X = -\frac{1}{2\pi \mathrm{i}} \int\limits_{\Gamma} d\zeta (C - \zeta I)^{-1} B^* (Z - \zeta I)^{-1}, \tag{3.1}$$

Clearly, for $\zeta \in \Gamma$ we have $\mathrm{Ran}(C - \zeta I)^{-1} \subset \mathrm{Dom}(C)$ and hence

$$\mathrm{Ran}(X) \subset \mathrm{Dom}(C),$$

which immediately follows from (3.1). Multiplying both sides of (3.1) by $B$ from the left yields

$$BX = -\frac{1}{2\pi \mathrm{i}} \int_{\Gamma} d\zeta B (C - \zeta I)^{-1} B^* (Z - \zeta I)^{-1}. \tag{3.2}$$

Meanwhile,

$$B(C - \zeta I)^{-1} B^* = A - \zeta I + M(\zeta) = A - \zeta I + W(\zeta)(Z - \zeta I), \quad \zeta \in \Gamma,$$

and, hence, using (3.2)

$$BX = -\frac{1}{2\pi \mathrm{i}} \int_{\Gamma} d\zeta \left[ W(\zeta) + A(Z - \zeta I)^{-1} - \zeta (Z - \zeta I)^{-1} \right]. \tag{3.3}$$

The function $W(\zeta)$ is holomorphic in the domain bounded by the contour $\Gamma$ and, thus, the first term in the integrand on the r.h.s. of (3.3) gives no contribution. Since $\Gamma$ encircles the spectrum of $Z$, the integration of the remaining two terms in (3.3) can be performed explicitly using the operator version of the residue theorem, which yields $BX = -A + Z$ and hence

$$Z = A + BX. \tag{3.4}$$

Since the spectra of the operators $C$ and $Z$ are disjoint and $Z$ is a bounded operator, it is straightforward to show (see, e.g., [4] or [21]) that the operator $X$ given by (3.1) is the unique operator solution to the Sylvester equation

$$XZ - CX = B^*,$$

which by (3.4) proves that $X$ solves the Riccati equation (1.1). $\square$

The existence of solutions to the Riccati equation (1.1) is directly related to the possibility of block diagonalization of the $2 \times 2$ self-adjoint block operator matrix

$$\mathbf{H} = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}, \quad \mathrm{Dom}(\mathbf{H}) = \mathcal{H}_A \oplus \mathrm{Dom}(C), \tag{3.5}$$

in the Hilbert space

$$\mathcal{H} = \mathcal{H}_A \oplus \mathcal{H}_C. \tag{3.6}$$

The precise statement describing this connection is as follows (see Lemma 5.3 and Theorem 5.5 in [3]; cf. [2], [6], [11], [20]).

**Theorem 3.2.** *Assume that $A$ is a bounded self-adjoint operator on a Hilbert space $\mathcal{H}_A$, $C$ possibly unbounded self-adjoint operator with domain* $\mathrm{Dom}(C)$ *on a Hilbert space $\mathcal{H}_C$, and $B$ a bounded operator from $\mathcal{H}_C$ to $\mathcal{H}_A$. Then a bounded operator $X$ from $\mathcal{H}_A$ to $\mathcal{H}_C$ is a solution to the Riccati equation* (1.1) *iff the graph $\mathcal{G}(X)$ of $X$ is a reducing subspace for the $2 \times 2$ block operator matrix* (3.5). *Moreover, if $X \in \mathcal{B}(\mathcal{H}_A, \mathcal{H}_C)$ is a solution to* (1.1) *then:*

(i) *The operator $\mathbf{V}^{-1}\mathbf{H}\mathbf{V}$ with*

$$\mathbf{V} = \begin{pmatrix} I & -X^* \\ X & I \end{pmatrix}.$$

*is block diagonal with respect to decomposition* (3.6). *Furthermore,*

$$\mathbf{V}^{-1}\mathbf{H}\mathbf{V} = \begin{pmatrix} Z & 0 \\ 0 & \widehat{Z} \end{pmatrix},$$

*where $Z = A + BX$ and $\widehat{Z} = C - B^*X^*$ with* $\mathrm{Dom}(\widehat{Z}) = \mathrm{Dom}(C)$.

(ii) *The operators*

$$\Lambda = (I + X^*X)^{1/2} Z (I + X^*X)^{-1/2} \tag{3.7}$$

*and*

$$\widehat{\Lambda} = (I + XX^*)^{1/2} \widehat{Z} (I + XX^*)^{-1/2}$$

*with* $\mathrm{Dom}(\widehat{\Lambda}) = (I + XX^*)^{1/2}(\mathrm{Dom}(C))$ *are self-adjoint operators in $\mathcal{H}_A$ and $\mathcal{H}_C$, respectively.*

Theorem 3.2 yields the following uniqueness result as a corollary.

**Corollary 3.3.** *Assume Hypothesis of Theorem 3.2. Suppose that $\Sigma$ and $\widehat{\Sigma}$ are disjoint Borel subsets of $\mathbb{R}$ such that $\operatorname{dist}(\Sigma, \widehat{\Sigma}) > 0$. Let $\mathcal{X} = \mathcal{X}(A, B, \Sigma, \widehat{\Sigma})$ be the set of all bounded operators $X$ from $\mathcal{H}_A$ to $\mathcal{H}_C$ with the properties*

$$\sigma(A + BX) \subset \Sigma, \tag{3.8}$$

$$\sigma(C - B^*X^*) \subset \widehat{\Sigma}, \quad \operatorname{Dom}(C - B^*X^*) = \operatorname{Dom}(C). \tag{3.9}$$

*Then if $X, Y \in \mathcal{X}$ satisfy the Riccati equation* (1.1)*, then $X = Y$.*

*Proof.* Suppose $X$ and $Y$ are two bounded solutions to (1.1) both satisfying (3.8) and (3.9). Then by Theorem 3.2 the graphs of $X$ and $Y$ both coincide with the spectral subspace of the $2 \times 2$ operator matrix (3.5) associated with the set $\Sigma$, and hence, $X = Y$. $\square$

Now we are ready to prove the key result of the section.

**Theorem 3.4.** *Assume Hypothesis 2.1 and suppose that*

$$\|B\| < \sqrt{d|\Delta|}.$$

*Then the spectral subspace of the operator* (3.5) *associated with the interval $\Delta$ is the graph of a bounded operator $X$ from $\mathcal{H}_A$ to $\mathcal{H}_C$.*

*Proof.* It is well known (see, e.g., [17]) that the resolvent of the operator $\mathbf{H}$ can be represented as the following $2 \times 2$ operator matrix

$$\begin{aligned}(\mathbf{H} - \lambda I)^{-1} = &\begin{pmatrix} 0 & 0 \\ 0 & (C - \lambda I)^{-1} \end{pmatrix} \\ &- \begin{pmatrix} I \\ -(C - \lambda I)^{-1}B^* \end{pmatrix} M^{-1}(\lambda) \begin{pmatrix} I & -B(C - \lambda I)^{-1} \end{pmatrix}, \\ &\lambda \in \rho(\mathbf{H}),\end{aligned} \tag{3.10}$$

where $M$ is the Herglotz function given by (2.1). By Theorem 2.4 the Herglotz function $M(\lambda)$ admits the factorization

$$M(\zeta) = W(\zeta)(Z - \zeta I), \quad \zeta \in \rho(C), \tag{3.11}$$

where $Z$ is a bounded operator with

$$\sigma(Z) \subset \Delta. \tag{3.12}$$

Representation (3.10) shows that for $\lambda \in \Delta$ the operator $\mathbf{H} - \lambda I$ has a bounded inverse iff $M(\lambda)$ does. By Theorem 2.4 $W(\zeta)$ is holomorphic on $\rho(C)$ and for any

$$\lambda \in \Big(\mathbb{C} \setminus \sigma(M)\Big) \cup \Delta$$

the operator $W(\lambda)$ has a bounded inverse. Therefore, using (3.11) and (3.12) proves that the spectrum of $\mathbf{H}$ in the interval $\Delta$ coincides with that of $Z$,

$$\sigma(\mathbf{H}) \cap \Delta = \sigma(Z),$$

and, hence, it is isolated from the remaining part of the spectrum of $\mathbf{H}$.

Using representation (3.10), for the spectral projection $\mathsf{E}_{\mathbf{H}}\big(\Delta\big)$ of the operator $\mathbf{H}$ associated with the interval $\Delta$ the Riesz integration yields

$$\begin{aligned}\mathsf{E}_{\mathbf{H}}\big(\Delta\big) =& \mathsf{E}_{\mathbf{H}}\big(\sigma(Z)\big) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (\mathbf{H} - \zeta I)^{-1} \\ =& -\frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta \begin{pmatrix} I \\ -(C - \zeta I)^{-1} B^* \end{pmatrix} M^{-1}(\zeta) \begin{pmatrix} I & -B(C - \zeta I)^{-1} \end{pmatrix},\end{aligned}$$

where $\Gamma$ is an arbitrary bounded Jordan contour in $\mathbb{C}$ encircling $\sigma(Z)$ in the clockwise direction and having winding number $0$ with respect to the spectrum of $C$. Hence,

$$\mathsf{E}_{\mathbf{H}}\big(\sigma(Z)\big) = \begin{pmatrix} E & G^* \\ G & F \end{pmatrix}, \tag{3.13}$$

where

$$\begin{aligned} E &= -\frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta M^{-1}(\zeta), \\ F &= -\frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (C - \zeta I)^{-1} B^* M^{-1}(\zeta) B (C - \zeta I)^{-1}, \end{aligned} \tag{3.14}$$

and

$$\begin{aligned} G &= \frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (C - \zeta I)^{-1} B^* M^{-1}(\zeta) \\ &= \frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (C - \zeta I)^{-1} B^* (Z - \zeta I)^{-1} W^{-1}(\zeta), \end{aligned}$$

using factorization formula (3.11). By Theorem 2.4 the function $W^{-1}(\zeta)$ is holomorphic in the domain bounded by $\Gamma$. Then applying the Daletsky-Krein lemma ([5, Lemma I.2.1]) yields

$$G = \left[\frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (C - \zeta I)^{-1} B^* (Z - \zeta I)^{-1}\right] \left[\frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (Z - \zeta I)^{-1} W^{-1}(\zeta)\right].$$

Hence, combining (3.11) and (3.14) proves the representation

$$G = XE, \tag{3.15}$$

where

$$X = -\frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (C - \zeta I)^{-1} B^* (Z - \zeta I)^{-1}. \tag{3.16}$$

In an analogous way one also proves that

$$F = XEX^*. \tag{3.17}$$

A closer look at the proof of Theorem 3.1 shows that the operator $X$ given by (3.16) solves the Riccati equation (1.1). Then applying Theorem 3.2 and using (3.13), (3.15), and (3.17) we obtain by inspection

$$\begin{aligned} I &= \frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta (Z - \zeta I)^{-1} \\ &= \frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta \mathsf{P}_A \begin{pmatrix} (Z - \zeta I)^{-1} & 0 \\ 0 & (Z_C - \zeta I)^{-1} \end{pmatrix} \mathsf{P}_A^* \\ &= \frac{1}{2\pi \mathrm{i}} \int_\Gamma d\zeta \mathsf{P}_A \mathbf{V}^{-1} (\mathbf{H} - \zeta I)^{-1} \mathbf{V} \mathsf{P}_A^* \\ &= \mathsf{P}_A \mathbf{V}^{-1} \mathsf{E}_{\mathbf{H}}\big(\sigma(Z)\big) \mathbf{V} \mathsf{P}_A^* \\ &= E(I + X^* X), \end{aligned} \tag{3.18}$$

where $\widehat{Z} = C - B^* X^*$ with $\mathrm{Dom}(\widehat{Z}) = \mathrm{Dom}(C)$,

$$\mathbf{V} = \begin{pmatrix} I & -X^* \\ X & I \end{pmatrix},$$

and $\mathsf{P}_A$ is the canonical projection from $\mathcal{H}$ to $\mathcal{H}_A$, i.e., $\mathsf{P}_A(f_A \oplus f_C) = f_A$ for any $f_A \in \mathcal{H}_A$ and $f_C \in \mathcal{H}_C$. It follows from (3.18) that

$$E = (I + X^* X)^{-1} \tag{3.19}$$

since $I + X^* X$ has a bounded inverse. Combining, (3.15) (3.17), and (3.19) proves that the spectral projection $\mathsf{E}_{\mathbf{H}}(\Delta)$ admits the following representation

$$\mathsf{E}_{\mathbf{H}}(\Delta) = \begin{pmatrix} (I + X^* X)^{-1} & (I + X^* X)^{-1} X^* \\ X(I + X^* X)^{-1} & X(I + X^* X)^{-1} X^* \end{pmatrix}. \tag{3.20}$$

Observing that the r.h.s. of (3.20) represents the orthogonal projection in $\mathcal{H} = \mathcal{H}_A \oplus \mathcal{H}_C$ onto the graph of the operator $X$ completes the proof. $\square$

The proof of the first principal result of the paper is now straightforward.

*Proof of Theorem 1.* By Theorem 3.1 the Riccati equation (1.1) has a bounded solution $X$ given by (3.1) (see the proof of this theorem). By Theorem 3.4 the graph of the operator $X$ is the spectral subspace $\operatorname{Ran} \mathsf{E}_{\mathbf{H}}\big(\Delta\big)$ of the block operator matrix $\mathbf{H}$ given by (3.5). Now, applying Theorem 3.2 shows that $X$ possesses the properties (1.6). Corollary 3.3 then proves that $X$ is the unique bounded solution to (1.1) satisfying (1.6) completing the proof. $\square$

**Remark 3.5.** Under hypothesis of Theorem 1 one obviously concludes that $|\Delta| \geq 2d$ with the equality sign occurring only if the spectrum of the operator $A$ is a one point set. Hence, (1.5) implies the solvability of the Riccati equation (1.1) under the condition $\|B\| < \sqrt{2}d$. Therefore, in the case where the operator $C$ has a finite spectral gap $\Delta$ containing the spectrum of $A$ the constant $c_{\mathrm{best}} \geq \sqrt{2}$. On the other hand in [3, Lemma 3.11 and Remark 3.12] it is shown that $c_{\mathrm{best}} \leq \sqrt{2}$. Thus, $c = \sqrt{2}$ is best possible in inequality (1.3).

## 4 Norm Estimates of Solutions. The tan $\Theta$ Theorem

We start by recalling the concept of the operator angle between two subspaces in a Hilbert space going back to the works by Friedrichs [9], M. Krein, Kransnoselsky, and Milman [14], [15], Halmos [10], and Davis and Kahan [8]. A comprehensive discussion of this notion can be found in [11].

Given a subspace $\mathcal{Q}$ of the Hilbert space $\mathcal{H} = \mathcal{H}_A \oplus \mathcal{H}_C$, introduce the operator angle $\Theta$ between the subspaces $\mathcal{H}_A \oplus \{0\}$ and $\mathcal{Q}$ by

$$\Theta = \arcsin(I_{\mathcal{H}_A} - \mathsf{P}_A \mathsf{Q} \mathsf{P}_A^*), \tag{4.1}$$

where $\mathsf{P}_A$ is the canonical projection from $\mathcal{H}$ onto $\mathcal{H}_A$ and $\mathsf{Q}$ the orthogonal projection in $\mathcal{H}$ onto $\mathcal{Q}$. If the subspace $\mathcal{Q}$ is the graph $\mathcal{G}(X)$ of a bounded operator $X$ from $\mathcal{H}_A$ to $\mathcal{H}_C$, then (see [11]; cf. [8] and [10])

$$\tan\Theta = \sqrt{X^*X} \tag{4.2}$$

and

$$\|\sin\Theta\| = \|\mathsf{Q} - \mathsf{P}\|, \tag{4.3}$$

where $\mathsf{P} = \mathsf{P}_A^* \mathsf{P}_A$ denotes the orthogonal projection in $\mathcal{H}$ onto the subspace $\mathcal{H}_A \oplus \{0\}$.

Note that the common definition of the operator angle (see, e.g., [11]) slightly differs from (4.1). Usually, the operator angle is defined as the restriction of (4.1) onto the maximal subspace of $\mathcal{H}_A$ where it has a trivial kernel. Obviously, for both definitions $\|\tan\Theta\|$ is the same.

Now we are ready to prove the second principal result of the paper.

*Proof of Theorem 2.* By Theorem 3.2 the operator $Z = A + BX$ is similar to the bounded self-adjoint operator $\Lambda$ given by (3.7) and hence the spectrum of $Z$ is a subset of the real axis,

$$\sigma(Z) = \sigma(\Lambda) \subset \mathbb{R}. \tag{4.4}$$

Rewrite the Riccati equation (1.1) in the form

$$X(Z - aI) - (C - aI)X = B^*, \tag{4.5}$$

where

$$a = \frac{1}{2}\Big(\sup\sigma(Z) + \inf\sigma(Z)\Big) = \frac{1}{2}\Big(\sup\sigma(\Lambda) + \inf\sigma(\Lambda)\Big). \tag{4.6}$$

By hypothesis $\operatorname{dist}\Big(\sigma(\mathbf{H}|_{\mathcal{G}(X)}), \sigma(C)\Big) = \delta > 0$. Theorem 3.2 implies that

$$\sigma(\mathbf{H}|_{\mathcal{G}(X)}) = \sigma(Z) = \sigma(\Lambda).$$

Therefore,

$$\sigma(C) \subset \Big(-\infty, \inf\sigma(\Lambda) - \delta\Big] \cup \Big[\sup\sigma(\Lambda) + \delta, \infty\Big). \tag{4.7}$$

Hence $C - aI$ has a bounded inverse, and combining (4.4), (4.6), and (4.7) yields

$$\|(C - aI)^{-1}\| = \frac{1}{\|\Lambda - aI\| + \delta}. \tag{4.8}$$

Multiplying both sides of (4.5) by $(C - aI)^{-1}$ from the left proves the representation

$$X = (C - aI)^{-1}\Big(X(Z - aI) - B^*\Big). \tag{4.9}$$

Using the claim (ii) of Theorem 3.2 one obtains

$$\begin{aligned}\|X(Z - aI)\| =& \|X(I + X^*X)^{-1/2}(\Lambda - aI)(I + X^*X)^{1/2}\| \\ &\le \|X(I + X^*X)^{-1/2}\|\, \|\Lambda - aI\|(1 + \|X\|^2)^{1/2}.\end{aligned} \tag{4.10}$$

Clearly,

$$\|X(I + X^*X)^{-1/2}\| = \sqrt{\|X^*X(I + X^*X)^{-1}\|} = \frac{\|X\|}{(1 + \|X\|^2)^{1/2}}$$

applying the spectral theorem on the last step. Hence (4.10) implies the estimate

$$\|X(Z-aI)\| \leq \|X\|\,\|(\Lambda-aI)\|,$$

which together with (4.9) proves the norm inequality

$$\|X\| \leq \|(C-aI)^{-1}\|\Big(\|\Lambda-aI\|\|X\|+\|B\|\Big). \tag{4.11}$$

Solving (4.11) with respect to $\|X\|$ and taking into account (4.8) proves (1.7). Finally, since $\left\|\sqrt{X^*X}\right\| = \|X\|$, using (4.2) one concludes that

$$\|\tan\Theta\| = \|X\|.$$

Hence, (1.7) is equivalent to (1.8). □

The estimates (1.7) and (1.8) depend on the spectral properties of the perturbed operator $\mathbf{H}$. Under additional assumptions one can also get an *a priori* estimate on the norm of the solution $X$.

**Corollary 4.1.** *Assume Hypothesis 2.1 and suppose that*

$$\|B\| < \sqrt{d|\Delta|}.$$

*If $X$ is the unique solution to the Riccati equation* (1.1) *referred to in Theorem 2, then*

$$\|X\| \leq \frac{\|B\|}{\min\{d^- - \delta_B^-, d^+ - \delta_B^+\}}, \tag{4.12}$$

*where $d^- = \inf\sigma(A) - \alpha$, $d^+ = \beta - \sup\sigma(A)$, and $d_B^\pm$ are given by* (2.6) *and* (2.7). *Equivalently,*

$$\|\tan\Theta\| \leq \frac{\|B\|}{\min\{d^- - \delta_B^-, d^+ - \delta_B^+\}}, \tag{4.13}$$

*where $\Theta$ is the operator angle between the subspace $\mathcal{H}_A \oplus \{0\}$ and spectral subspace of the* $2\times 2$ *operator matrix* (3.5) *associated with the interval $\Delta$. Moreover, the spectral projections $\mathsf{E}_{\mathbf{H}_0}(\Delta)$ and $\mathsf{E}_{\mathbf{H}}(\Delta)$ of the operators $\mathbf{H}_0 = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}$ and $\mathbf{H} = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}$, respectively, satisfy the estimate*

$$\|\mathsf{E}_{\mathbf{H}}(\Delta) - \mathsf{E}_{\mathbf{H}_0}(\Delta)\| \leq \sin\arctan\left(\frac{\|B\|}{\min\{d^- - \delta_B^-, d^+ - \delta_B^+\}}\right). \tag{4.14}$$

*Proof.* Let $X$ be the solution to the Riccati equation (1.1) referred to in Theorem 1. By Theorem (3.1)

$$\sigma(A + BX) = \sigma(Z), \tag{4.15}$$

where $Z$ is the operator root of $M$ referred to in Theorem 2.4. Theorem 2.4 also states that

$$\sigma(Z) \subset [\inf \sigma(A) - \delta_B^-, \sup \sigma(A) + \delta_B^+]. \tag{4.16}$$

Recall that by Hypothesis 2.1

$$\sigma(C) \subset \Big(-\infty, \alpha\Big] \cup \Big[\beta, \infty\Big). \tag{4.17}$$

Combining (4.15), (4.16), and (4.17) yields

$$\operatorname{dist}\Big(\sigma(A + BX), \sigma(C)\Big) \geq \min\{d^- - \delta_B^-, d^+ - \delta_B^+\},$$

which proves (4.12) applying Theorem 2.

By Theorem 3.4 the spectral projection of the operator (3.5) associated with the interval $\Delta$ is the orthogonal projection onto the graph of the operator $X$. Then it follows from Theorem 2 that (4.12) is equivalent to (4.13).

The last estimate (4.14) is an immediate consequence of (4.13) by using (4.2) and (4.3). $\square$

## Acknowledgements

V. Kostrykin in grateful to V. Enss, A. Knauf, H. Leschke, and R. Schrader for useful discussions. K. A. Makarov is indebted to Graduiertenkolleg “Hierarchie und Symmetrie in mathematischen Modellen” for kind hospitality during his stay at RWTH Aachen in summer 2002. A. K. Motovilov acknowledges the kind hospitality and support by the Department of Mathematics, University of Missouri, Columbia, MO, USA. He was also supported in part by the Russian Foundation for Basic Research within Project RFBR 01-01-00958.

## References

[1] V. M. Adamjan and H. Langer, Spectral properties of a class of rational operator valued functions, *J. Operator Theory* **33** (1995), 259 – 277.

[2] V. Adamyan, H. Langer, and C. Tretter, Existence and uniqueness of contractive solutions of some Riccati equations, *J. Funct. Anal.* **179** (2001), 448 – 473.

[3] S. Albeverio, K. A. Makarov, and A. K. Motovilov, Graph subspaces and the spectral shift function, *Canad. J. Math.* (to appear); LANL e-print math.SP/0105142; available from http://arxiv.org/abs/math.SP/0105142.

[4] Yu. L. Daleckiĭ, On the asymptotic solution of a vector differential equation, *Dokl. Akad. Nauk SSSR* **92** (1953), 881 – 884 (Russian).

[5] Ju. L. Daleckiĭ and M. G. Kreĭn, “Stability of Solutions of Differential Equations in Banach Spaces”, Translations of Mathematical Monographs, Vol. 43, AMS, Providence, Rhode Island, 1974.

[6] J. Daughtry, Isolated solutions of quadratic matrix equations, *Linear Algebra Appl.* **21** (1978), 89 – 94.

[7] C. Davis, The rotation of eigenvectors by a perturbation. I and II, *J. Math. Anal. Appl.* **6** (1963), 159 – 173; **11** (1965), 20 – 27.

[8] C. Davis and W. M. Kahan, The rotation of eigenvectors by a perturbation. III, *SIAM J. Numer. Anal.* **7** (1970), 1 – 46.

[9] K. Friedrichs, On certain inequalities and characteristic value problems for analytic functions and for functions of two variables, *Trans. Amer. Math. Soc.* **41** (1937), 321 – 364.

[10] P. R. Halmos, Two subspaces, *Trans. Amer. Math. Soc.* **144** (1969), 381–389.

[11] V. Kostrykin, K. A. Makarov, and A. K. Motovilov, Existence and uniqueness of solutions to the operator Riccati eqution. A geometric approach. LANL e-print math.SP/0207125; available from http://arxiv.org/abs/math.SP/0207125.

[12] V. Kostrykin, K. A. Makarov, and A. K. Motovilov, A generalization of the $\tan 2\Theta$ theorem, in preparation.

[13] V. Kostrykin, K. A. Makarov, and A. K. Motovilov, Perturbation of spectra and spectral subspaces, in preparation.

[14] M. G. Krein and M. A. Krasnoselsky, Fundamental theorems about extensions of Hermite operators and some applications to the theory of orthogonal polynomials and to the moment problem, *Uspekhi Mat. Nauk* **2** (1947), 60 – 106 (Russian).

[15] M. G. Krein, M. A. Krasnoselsky, and D. P. Milman, On defect numbers of linear operators in Banach space and some geometric problems, *Sbornik Trudov Instituta Matematiki Akademii Nauk Ukrainskoy SSR*, **11** (1948), 97–112 (Russian).

[16] A. S. Markus and V. I. Matsaev, Spectral theory of holomorphic operator-functions in Hilbert space, *Funct. Anal. Appl.* **9** (1975), 73 – 74.

[17] R. Mennicken and A. K. Motovilov, Operator interpretation of resonances arising in spectral problems for $2 \times 2$ operator matrices, *Math. Nachr.* **201** (1999), 117 – 181 (LANL e-print funct-an/9708001).

[18] R. Mennicken and A. A. Shkalikov, Spectral decomposition of symmetric operator matrices, *Math. Nachr.* **179** (1996), 259 – 273.

[19] A. K. Motovilov, Potentials appearing after the removal of energy-dependence and scattering by them, in Proc. Intern. Workshop *"Mathematical Aspects of the Scattering Theory and Applications"*, St. Petersburg State University, St. Petersburg, 1991, pp. 101 – 108.

[20] A. K. Motovilov, Removal of the resolvent-like energy dependence from interactions and invariant subspaces of a total Hamiltonian, *J. Math. Phys.* **36** (1995), 6647 – 6664 (LANL e-print funct-an/9606002).

[21] M. Rosenblum, On the operator equation $BX - XA = Q$, *Duke Math. J.*, **23** (1956), 263 – 269.

[22] A. I. Virozub and V. I. Matsaev, The spectral properties of a certain class of self-adjoint operator functions, *Funct. Anal. Appl.* **8** (1974), 1 – 9.