Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

ELA

CONSISTENCY AND EFFICIENT SOLUTION OF THE SYLVESTER EQUATION FOR $\star$-CONGRUENCE*

FERNANDO DE TERÁN† AND FROILÁN M. DOPICO‡

Abstract. In this paper, the matrix equation $AX + X^{\star}B = C$ is considered, where the matrices A and B have sizes $m \times n$ and $n \times m$, respectively, the size of the unknown X is $n \times m$, and the operator $(\cdot)^{\star}$ denotes either the transpose or the conjugate transpose of a matrix. In the first part of the paper, necessary and sufficient conditions for the existence and uniqueness of solutions are reviewed. These conditions were obtained previously by Wimmer [H.K. Wimmer. Roth's theorems for matrix equations with symmetry constraints. *Linear Algebra Appl.*, 199:357–362, 1994.], by Byers and Kressner [R. Byers and D. Kressner. Structured condition numbers for invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 28:326–347, 2006.], and by Kressner, Schröder and Watkins [D. Kressner, C. Schröder, and D.S. Watkins. Implicit QR algorithms for palindromic and even eigenvalue problems. *Numer. Algorithms*, 51:209–238, 2009.]. This review generalizes to fields of characteristic different from two the existence condition that Wimmer originally proved for the complex field. In the second part, an algorithm is developed, in the real or complex square case $m = n$, to solve the equation in $O(n^3)$ flops when the solution is unique. This algorithm is based on the generalized Schur decomposition of the matrix pencil $A - \lambda B^{\star}$. The equation $AX + X^{\star}B = C$ is connected with palindromic eigenvalue problems and, as a consequence, the square complex case has attracted recently the attention of several authors.

Key words. Generalized Schur decomposition, Matrix equations, Sylvester equation, Palindromic eigenvalue problems, Congruence of matrices.

AMS subject classifications. 65F05, 65F15, 15A24.

1. Introduction.

The Sylvester equation $AX - XB = C$, where $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, $C \in \mathbb{C}^{m\times n}$ are given and $X \in \mathbb{C}^{m\times n}$ is to be determined, is one of the most important matrix equations in theory and applications. Let us recall some of its well-known properties that may be found in standard references on matrix analysis as [17, Chapter 16] or [19, Section 4.4]. The Sylvester equation has a unique solution for each C if and only if A and B have no eigenvalues in common. In 1952, Roth proved in [24] that the Sylvester equation has some solution (perhaps nonunique) if

*Received by the editors on March 12, 2011. Accepted for publication on August 14, 2011. Handling Editor: Christian Mehl. This work was partially supported by the Ministerio de Ciencia e Innovación of Spain through grant MTM-2009-09281.

†Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (fteran@math.uc3m.es).

‡Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM and Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (dopico@math.uc3m.es).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

ELA

and only if

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad \text{are similar.} \tag{1.1}$$

Roth proved this result in any field $\mathbb{F}$ through a rather indirect argument that uses the Smith canonical form of matrices with entries in the polynomial ring $\mathbb{F}[x]$. A more direct proof was presented 25 years later in [12], by using certain linear mappings and dimensional arguments. The proof in [12] may be found in modern references, as for instance in [13, Theorem S2.1] or [19, Theorem 4.4.22]. The relationship of the Sylvester equation with the block-diagonalization of block triangular matrices shown in (1.1) is the reason of its importance in invariant subspace computations [15, Section 7.6.3]. When the solution is unique for every C, the classical numerical method for solving the Sylvester equation is the Bartels-Stewart algorithm [2] (see also [15, Algorithm 7.6.2]) that makes use of the Schur decompositions of A and B and requires $O(m^3 + n^3)$ flops. A more efficient modification of the Bartels-Stewart algorithm was proposed in [14].

We consider in this paper the following matrix equations

$$AX + X^\star B = C, \tag{1.2}$$

where $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$, the unknown is $X \in \mathbb{C}^{n\times m}$, and the operator $(\cdot)^\star$ denotes either the transpose $((\cdot)^T)$ or the conjugate transpose $((\cdot)^*)$ of a matrix. In contrast to the Sylvester equation, there are not many references in the literature for the equations (1.2) and the existing ones seem to be scattered and not well-known in the Linear Algebra community. However, equations (1.2) have attracted recently the attention of several researchers as a consequence of their relationship with palindromic eigenvalue problems. Our purpose in this note is to gather some results published in the literature on the existence of solutions for these equations, and then to develop a new efficient numerical algorithm to compute the solution.

The first reference we know on equation (1.2) is [29]. In this paper, Wimmer provides necessary and sufficient conditions for the existence of solutions over the complex field in the case $\star = *$. After this work, which, despite of its relevance, seems to have passed quite unnoticed, equations (1.2) have been considered by several authors. Braden refers in [4] to the existence of a simple explicit expression for the solution of (1.2) for the case $\star = T$ as an open problem. More recently, equation (1.2) has been solved in [7] for bounded linear g-invertible operators (where $\star$ now denotes the adjoint operator) and for a very particular case in which the operators A, B, and C, and their Moore-Penrose generalized inverses satisfy certain specific identities. The solution is given in terms of the operators A, B, C, their adjoints and their Moore-Penrose generalized inverses. Also, equations (1.2) for $m = n$ have appeared in [5,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

ELA

Lemma 5.10] in connection with structured condition numbers of deflating subspaces of regular palindromic pencils $G + \lambda G^T$. Reference [5] only considers the case $\star = T$ and establishes necessary and sufficient conditions for the existence of a unique solution for every right-hand side C. These conditions are modified to cover the case $\star = *$ in [23, Lemma 8], where the equation (1.2) arises in the context of a structure-preserving QR algorithm for computing the eigenvalues of regular palindromic pencils. The following particular case of (1.2)

$$AX + X^\star A = 0 \tag{1.3}$$

has been considered in [9, 10], where the authors present a (non-numerical) method to find the set of solutions of (1.3) through the use of the canonical form of the matrix A under $\star$-congruence [20]. References [9, 10] pay special attention to the relationship between (1.3) and the orbit of A under the action of $\star$-congruence. More precisely, the dimension of the solution space of (1.3) is shown to be equal to the codimension of this orbit. Hence, since the authors determine the dimension of the solution space of (1.3), they also obtain the dimension of the $\star$-congruence orbit of A. The much simpler version $A^T X \pm X^T A = B$ of (1.2) was solved in [4]. In this case, the fact that $(A^T X)^T = X^T A$ simplifies considerably the analysis. The main result in [4] has been extended in [11] to the equation $A^* X + X^* A = C$, where A, C, X are linear bounded operators and A is of closed range (here $*$ stands for the adjoint operator). Reference [4] is related to the much older references [18] and [27]. In [18] the author considers the equation $X^\star A + A^\star X = C$ over finite fields, and with C being symmetric, skew-symmetric or Hermitian. He obtains explicit formulas for the number of solutions and provides also conditions for the solvability. In [27] the eigenvalues of the linear transformation $g(X) = A^T X + X^T A$ are determined. This allows the establishment of necessary and sufficient conditions for the existence of a unique solution of $A^T X + X^T A = C$ for every C. Somewhat connected to [4], [27], and equation (1.3), we mention [1, Theorem 2] that gives necessary and sufficient conditions for the consistency of $AX + X^* A = C$ with $A = A^*$ and positive definite. The results discussed in this paragraph are the only ones that have been published for the equations (1.2), as far as we know. We want to mention also the recent manuscript [6], which includes results related to the ones in the present work.

The necessary and sufficient conditions on the existence and uniqueness of solutions of equations (1.2) developed in [5, 23, 29] are stated and reviewed in Section 2, with the goal of bringing these results to the attention of researchers interested in the solution of this equation. In this reviewing process, we have extended Wimmer's necessary and sufficient condition for consistency in the complex field and $\star = *$. More precisely, we provide a necessary and sufficient condition for the existence of solutions of $AX + X^\star B = C$ in a much more general case, that is, *for rectangular matrices with entries in any field* $\mathbb{F}$ *of characteristic different from two and* $\star = T$ *or* $\star = *$. This

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

result is presented in Theorem 2.3 below. The proof uses different techniques than the ones used in [29]. The condition has the same flavor as Roth's criterion for the standard Sylvester equation, although a very important difference must be observed: Roth's criterion involves block-diagonalization through similarity, while Theorem 2.3 involves *block-antidiagonalization through* $\star$*-congruence*. This fact has motivated us to call $AX + X^\star B = C$ the *Sylvester equation for* $\star$*-congruence*.

In Section 3, we focus on real or complex square equations (1.2) that satisfy the conditions of Lemma 8 in [23] for the existence of a unique solution for every right-hand side C. We present an efficient numerical method to find this solution, and this is our main original contribution. The cost of this algorithm is $O(n^3)$ flops and is in the spirit of the Bartels-Stewart algorithm for the standard Sylvester equation. The method we propose uses the generalized Schur form of the pencil $A - \lambda B^\star$ [15, Theorem 7.7.1], something natural once the conditions in [23, Lemma 8] are known, and is also related to solution methods of generalized Sylvester equations [22]. In addition, we will discuss briefly the rounding errors committed by this procedure.

The paper is organized as follows. Section 2 deals with the existence and uniqueness of solutions of (1.2) and Section 3 presents the numerical algorithm mentioned above for computing the solution. Finally, some conclusions and lines of future research are discussed in Section 4.

2. Existence and uniqueness of solutions of Sylvester equation for $\star$-congruence.

2.1. Uniqueness of solutions. We start with Lemma 8 in [23], which deals with the existence of a unique solution of (1.2) when A and B are both complex square matrices with the same size. This result solves completely the question of uniqueness for every right-hand side C in the complex field, since we will discuss in Section 2.2 that if A and B are rectangular matrices, then equation (1.2) never has a unique solution for every C. First, we need to define that a set of complex numbers $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is $\star$*-reciprocal free* if $\lambda_i \neq 1/\lambda_j^\star$ for any $1 \leq i, j \leq n$. This definition admits 0 and/or ∞ as elements of $\{\lambda_1, \ldots, \lambda_n\}$. Note that for numbers the $(\cdot)^\star$-operator is simply $\lambda_j^T = \lambda_j$ or $\lambda_j^* = \bar{\lambda}_j$. We will denote by $\Lambda(A, B^\star)$ the set of eigenvalues of the pencil $A - \lambda B^\star$. Recall also that the pencil $A - \lambda B^\star$ is regular if $\det(A - \lambda B^\star) \not\equiv 0$.

Lemma 2.1 (Lemma 8 in [23]). *Let $A, B \in \mathbb{C}^{n\times n}$ be given. The matrix equation $AX + X^\star B = C$ has a unique solution X for every right-hand side $C \in \mathbb{C}^{n\times n}$ if and only if the following conditions hold:*

1) *The pencil $A - \lambda B^\star$ is regular, and*
2a) *if $\star = T$, $\Lambda(A, B^T)\backslash\{1\}$ is T-reciprocal free and if $1 \in \Lambda(A, B^T)$, then it has*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

ELA

algebraic multiplicity 1*, or*
2b) *if* $\star = *$*,* $\Lambda(A, B^*)$ *is* $*$*-reciprocal free.*

2.2. Consistency of the equation. To prove Theorem 2.3 we will use a result obtained by Wimmer in [28] on the consistency of pairs of generalized Sylvester equations. Before describing this result, we need to introduce some notation and basic definitions. Given an arbitrary field $\mathbb{F}$, we denote by $\mathbb{F}^{m\times n}$ the space of $m \times n$ matrices with entries in $\mathbb{F}$. Two matrix pencils $E_1 - \lambda F_1$ and $E_2 - \lambda F_2$, with $E_1, F_1, E_2, F_2 \in \mathbb{F}^{m\times n}$ are *strictly equivalent* if there exist two nonsingular matrices $P \in \mathbb{F}^{m\times m}$ and $Q \in \mathbb{F}^{n\times n}$ such that $P(E_1 - \lambda F_1)Q = E_2 - \lambda F_2$. Next theorem appeared in [28, Theorem 1.1]. It was proved independently in [26, Theorem 2.3], and also in [3, Theorem 5.1] for the complex field $\mathbb{F} = \mathbb{C}$.

Theorem 2.2 (Theorem 1.1 in [28]). *Given* $A_1, A_2 \in \mathbb{F}^{m\times n}$*,* $B_1, B_2 \in \mathbb{F}^{p\times k}$*, and* $C_1, C_2 \in \mathbb{F}^{m\times k}$*, the pair of generalized Sylvester equations*

$$\begin{aligned} A_1X + YB_1 &= C_1\,, \\ A_2X + YB_2 &= C_2 \end{aligned}$$

has a solution (X, Y) *if and only if the matrix pencils*

$$\begin{bmatrix} A_1 - \lambda A_2 & C_1 - \lambda C_2 \\ 0 & B_1 - \lambda B_2 \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} A_1 - \lambda A_2 & 0 \\ 0 & B_1 - \lambda B_2 \end{bmatrix}$$

are strictly equivalent.

Given an arbitrary field $\mathbb{F}$, the operator $(\cdot)^\star$ on $\mathbb{F}^{m\times n}$ denotes the transpose of a matrix, except in the particular case $\mathbb{F} = \mathbb{C}$, where it may denote either the transpose or the conjugate transpose of a matrix. Two matrices $A, B \in \mathbb{F}^{n\times n}$ are $\star$*-congruent* if there exists a nonsingular matrix $P \in \mathbb{F}^{n\times n}$ such that $P^\star AP = B$. Theorem 2.3 extends the equivalence (a) $\Leftrightarrow$ (b) of Theorem 2 in [29], which is stated only for matrices over the complex field $\mathbb{C}$ and for the case $\star = *$. Theorem 2.3 establishes a necessary and sufficient condition for the consistency of the Sylvester equation for $\star$-congruence for rectangular matrices with entries in any field of characteristic different from two.

Theorem 2.3. *Let* $\mathbb{F}$ *be a field of characteristic different from two and let* $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$, $C \in \mathbb{F}^{m\times m}$ *be given. There is some* $X \in \mathbb{F}^{n\times m}$ *such that*

$$AX + X^\star B = C \tag{2.1}$$

if and only if

$$\begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \quad \textit{are } \star\textit{-congruent.} \tag{2.2}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

Proof. Let us first prove the necessary condition. Let $X \in \mathbb{F}^{n\times m}$ be a solution of the equation (2.1). Then we have

$$\begin{bmatrix} I_m & -X^\star \\ 0 & I_n \end{bmatrix} \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \begin{bmatrix} I_m & 0 \\ -X & I_n \end{bmatrix} = \begin{bmatrix} C - AX - X^\star B & A \\ B & 0 \end{bmatrix} = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix},$$

so the matrices in (2.2) are $\star$-congruent, with $P = \left[\begin{smallmatrix} I_m & 0 \\ -X & I_n \end{smallmatrix}\right]$ as a congruency matrix.

Let us prove the sufficient condition. Assume that the matrices in (2.2) are $\star$-congruent. Then, there is a nonsingular matrix P such that

$$P^\star \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} P = \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}. \tag{2.3}$$

The $(\cdot)^\star$ operator applied on (2.3) gives

$$P^\star \begin{bmatrix} C^\star & B^\star \\ A^\star & 0 \end{bmatrix} P = \begin{bmatrix} 0 & B^\star \\ A^\star & 0 \end{bmatrix}, \tag{2.4}$$

and equation (2.3) minus λ times (2.4) produces

$$P^\star \begin{bmatrix} C - \lambda C^\star & A - \lambda B^\star \\ B - \lambda A^\star & 0 \end{bmatrix} P = \begin{bmatrix} 0 & A - \lambda B^\star \\ B - \lambda A^\star & 0 \end{bmatrix}.$$

A permutation of the block columns of previous equation allows us to see that the matrix pencils

$$\begin{bmatrix} A - \lambda B^\star & C - \lambda C^\star \\ 0 & B - \lambda A^\star \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A - \lambda B^\star & 0 \\ 0 & B - \lambda A^\star \end{bmatrix}$$

are strictly equivalent. Now, Theorem 2.2 implies that the system

$$\begin{aligned} AY + ZB &= C \\ B^\star Y + ZA^\star &= C^\star \end{aligned} \tag{2.5}$$

has a solution (Y, Z). Apply the $(\cdot)^\star$ operator to the second equation in (2.5), sum the result to the first equation, and get

$$A(Y + Z^\star) + (Z + Y^\star)B = 2C.$$

So, if the characteristic of $\mathbb{F}$ is not two, then $X = \frac{1}{2}(Y + Z^\star)$ satisfies (2.1). Hence, the sufficiency follows. ◻

Observe that if $m \neq n$, then the equation (2.1) never has a unique solution for every right-hand side C, that is, the operator $X \mapsto AX + X^\star B$ is never invertible.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

ELA

This follows from the fact that $X \in \mathbb{F}^{n\times m}$, while $AX + X^\star B \in \mathbb{F}^{m\times m}$. Therefore the domain and the codomain of the operator have different dimensions and the operator cannot be invertible. To make this argument fully precise, observe that the Sylvester equation for congruence, and the corresponding operator, is linear in $\mathbb{F}$ if $\star = T$, but not if $\star = *$. If $\star = *$, then equation (2.1) is equivalent to a real linear system of two matrix equations having as unknowns the real and imaginary parts of X.

It is worth to compare the block structure of the matrices in (2.2) with the ones appearing in Roth's criterion (1.1) for the standard Sylvester equation. We want to remark in this respect that the $\star$-congruence of the matrices in (2.2) does not imply in general the $\star$-congruence of

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

As a counterexample, consider, for instance, $A = B = C = 1$. We have that

$$H = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad \text{and} \qquad G = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

are T-congruent, because $P^T HP = G$ with $P = \left[\begin{smallmatrix} 1 & 0 \\ -1/2 & 1 \end{smallmatrix}\right]$. However,

$$E = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad F = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

are not T-congruent, since $P^T EP$ is never symmetric for nonsingular P.

3. Solution of the equation $\mathbf{AX + X^\star B = C}$ via the generalized Schur decomposition of the pair $(A, B^\star)$. Throughout this section, we consider the Sylvester equation for $\star$-congruence only for square real or complex matrices, that is, we assume that $A, B, C \in \mathbb{F}^{n\times n}$ with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. In addition, we will assume that the conditions of Lemma 2.1 hold, that is, we assume that the equation $AX + X^\star B = C$ has a unique solution for every C. In this context, the reader should note that if $\mathbb{F} = \mathbb{R}$, then the unique solution of $AX + X^\star B = C$ is necessarily real both for $\star = T$ and $\star = *$. This is obvious for $AX + X^T B = C$, because nonsingular linear systems with real matrix coefficient and real right-hand side have a unique real solution. For $AX + X^* B = C$, if X is a solution, then by conjugating the equation, $\overline{X}$ is also a solution and, by the uniqueness assumption, $X = \overline{X}$, implying that X is real. Therefore, if $\mathbb{F} = \mathbb{R}$, then one only needs to consider $\star = T$. For brevity, we deal simultaneously with the real and complex cases, and with $\star = T$ and $\star = *$.

As in the study of the standard Sylvester equation, well-known properties of the Kronecker product [19, Chapter 4] can be used to write the matrix equation $AX + X^T B = C$ as a standard linear system for the unknown $\mathrm{vec}(X) \in \mathbb{F}^{n^2}$, where

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

ELA

the vec operator stacks the columns of a matrix into one long column vector. This system is

$$\left[\,(I_n \otimes A) + (B^T \otimes I_n)\Pi\,\right]\ \mathrm{vec}(X) = \mathrm{vec}(C), \tag{3.1}$$

where $\otimes$ denotes the Kronecker product, $\Pi \in \mathbb{R}^{n^2\times n^2}$ is a permutation matrix that satisfies $\mathrm{vec}(X^T) = \Pi\,\mathrm{vec}(X)$ for every $X \in \mathbb{F}^{n\times n}$ [19, Theorem 4.3.8], and I_n is the identity matrix. One may apply directly Gaussian elimination with partial pivoting (GEPP) to solve (3.1) with a cost of $O(n^6)$ flops, which is prohibitive except for very small n. Similar techniques allow us to write $AX+X^*B = C$, in the complex case, as a standard real linear system for the unknown $\left[\,(\mathrm{vec}(\mathrm{Re}\,X))^T\ \ (\mathrm{vec}(\mathrm{Im}\,X))^T\,\right]^T \in \mathbb{F}^{2\,n^2}$, where $\mathrm{Re}\,X$ and $\mathrm{Im}\,X$ are the real and imaginary parts of X. GEPP on this linear system leads again to a prohibitive cost of $O(n^6)$ flops.

Next, we present an algorithm for computing the unique solution of $AX+X^\star B = C$ with a cost of $O(n^3)$ flops. This algorithm is based on the generalized Schur decomposition of the pair $(A, B^\star)$, and involves four steps, as also happens for generalized Schur algorithms for other types of linear matrix equations [22]. Only Step 3 in this procedure requires a careful development, that will be presented in detail in Algorithm 3.2.

Algorithm 3.1. (**Algorithm to solve $\mathbf{AX + X^\star B = C}$**) Given $A, B, C \in \mathbb{F}^{n\times n}$, with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, such that A and B satisfy the conditions 1) and 2) in Lemma 2.1, this algorithm computes the unique solution $X \in \mathbb{F}^{n\times n}$ of $AX + X^\star B = C$ in $O(n^3)$ flops.

Step 1 Compute the generalized Schur decomposition of the pair $(A, B^\star)$ using the QZ algorithm [15, Section 7.7]

$$A = URV, \qquad B^\star = USV. \tag{3.2}$$

In general, $U, V \in \mathbb{C}^{n\times n}$ are unitary matrices and $R, S \in \mathbb{C}^{n\times n}$ are upper triangular matrices. However, if $A, B \in \mathbb{R}^{n\times n}$, then one can use only real arithmetic and compute the generalized *real* Schur decomposition, for which $U, V \in \mathbb{R}^{n\times n}$ are real orthogonal matrices, $S \in \mathbb{R}^{n\times n}$ is upper triangular, but $R \in \mathbb{R}^{n\times n}$ is *upper quasi-triangular*, that is, block upper triangular with 1×1 or 2×2 diagonal blocks.

Step 2 Compute

$$E = U^* C\, (U^\star)^*\,.$$

Observe that $(U^\star)^* = U$ if $\star = *$, and that $(U^\star)^* = \overline{U}$ if $\star = T$. In addition, if $U \in \mathbb{R}^{n\times n}$, then $U^* = U^T$ and $\overline{U} = U$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

Step 3 Use Algorithm 3.2 below to solve the transformed equation

$$RW + W^\star S^\star = E \tag{3.3}$$

for the unknown $W \in \mathbb{F}^{n\times n}$. Equation (3.3) is obtained from $AX + X^\star B = C$ with the decompositions (3.2) and the change of variable $W = VX\,(U^\star)^*$. The pencils $R - \lambda S$ and $A - \lambda B^\star$ are strictly equivalent, so Lemma 2.1 guarantees that the Sylvester equation for $\star$-congruence (3.3) has a unique solution W for every right-hand side E.

Step 4 Compute $X = V^* W U^\star$.

Let us explain how to solve the transformed equation (3.3). To cover the possible case of generalized real Schur decompositions in (3.2) when $\mathbb{F} = \mathbb{R}$ (recall that in this case $\star = T$), we consider R and S partitioned into $p \times p$ blocks as

$$R = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1p} \\ & R_{22} & & \vdots \\ & & \ddots & R_{p-1,p} \\ & & & R_{pp} \end{bmatrix}, \qquad S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ & S_{22} & & \vdots \\ & & \ddots & S_{p-1,p} \\ & & & S_{pp} \end{bmatrix}, \tag{3.4}$$

where $R_{ij}, S_{ij} \in \mathbb{F}^{n_i \times n_j}$ for $1 \le i, j \le p$, and $n_k = 1$ or 2 for $1 \le k \le p$. The diagonal blocks S_{ii} are always upper triangular matrices, but the diagonal blocks R_{ii} may be not if $A, B \in \mathbb{R}^{n\times n}$. If complex generalized Schur decompositions are computed in (3.2), then $p = n$ and $n_k = 1$ for $1 \le k \le n$. We also partition into $p \times p$ blocks the unknown W and the right-hand side E as

$$W = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1p} \\ W_{21} & W_{22} & & W_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ W_{p1} & W_{p2} & \cdots & W_{pp} \end{bmatrix}, \qquad E = \begin{bmatrix} E_{11} & E_{12} & \cdots & E_{1p} \\ E_{21} & E_{22} & & E_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ E_{p1} & E_{p2} & \cdots & E_{pp} \end{bmatrix}, \tag{3.5}$$

where the sizes of the blocks are $W_{ij}, E_{ij} \in \mathbb{F}^{n_i \times n_j}$, that is, the same sizes as in the partitions (3.4). As strategy to solve (3.3), we propose to determine first simultaneously the last block column and the last block row of W, then to determine simultaneously the last block column and the last block row of $W(1:p-1, 1:p-1) := [W_{ij}]_{i,j=1}^{p-1}$, then to determine simultaneously the last block column and the last block row of $W(1:p-2, 1:p-2)$, and, so on until we determine W_{11}. Observe that we have extended in the previous discussion standard MATLAB notation for submatrices from indices of entries to block-indices, since $W(1:p-1, 1:p-1)$ denotes the submatrix of W consisting of block rows 1 through $p-1$ and block columns 1 through $p-1$. Let us show the procedure for the last block column and the last block row of W. From the (p,p) block-entry of equation (3.3) we obtain

$$R_{pp} W_{pp} + W_{pp}^\star S_{pp}^\star = E_{pp}, \tag{3.6}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

that has a unique solution W_{pp} in the conditions of Lemma 2.1, because these conditions are inherited by the matrix pencil $R_{pp} - \lambda S_{pp}$ and (3.6) is again a Sylvester equation for $\star$-congruence. Equation (3.6) can be transformed into a standard linear system for $\mathrm{vec}(W_{pp})$, if $\star = T$, or for $[\,\mathrm{Re}\, W_{pp} \quad \mathrm{Im}\, W_{pp}\,]^T$, if $\star = *$ (recall that in this case all blocks are 1×1). This linear system can be solved by GEPP, since it has at most 4 unknowns when $R_{pp}, S_{pp}, E_{pp} \in \mathbb{R}^{2\times 2}$. Assume now that we have computed $W_{pp}, W_{p,p-1}, W_{p-1,p}, W_{p,p-2}, W_{p-2,p}, \ldots, W_{p,k+1}, W_{k+1,p}$. Then, from the block-entries (p,k) and (k,p) of (3.3) we obtain, after applying $(\cdot)^\star$ to the equation coming from (p,k) and performing some algebraic manipulations,

$$S_{kk}W_{kp} + W_{pk}^\star R_{pp}^\star = E_{pk}^\star - \sum_{j=k+1}^{p} S_{kj}W_{jp}\,, \tag{3.7}$$

$$R_{kk}W_{kp} + W_{pk}^\star S_{pp}^\star = E_{kp} - \sum_{j=k+1}^{p} R_{kj}W_{jp}\,. \tag{3.8}$$

The right-hand sides of equations (3.7)-(3.8) are known by our assumptions, so (3.7)-(3.8) are a pair of generalized Sylvester equations that have a unique solution for W_{kp} and $W_{pk}^\star$. The uniqueness follows again from the conditions of Lemma 2.1, that guarantee that the regular pencils $R_{kk} - \lambda S_{kk}$ and $S_{pp}^\star - \lambda R_{pp}^\star$ have no common eigenvalues (see [25, Theorem 1.11, Chapter VI]). Using the properties of the Kronecker product, equations (3.7)-(3.8) can be transformed into a standard linear system for $\left[(\mathrm{vec}(W_{kp}))^T \quad \left(\mathrm{vec}(W_{pk}^\star)\right)^T \right]^T$ that can be solved with GEPP, since it has at most 8 unknowns when $W_{kp}, W_{pk}^\star \in \mathbb{R}^{2\times 2}$. We have just shown that solving first (3.6) and then the system (3.7)-(3.8) for $k = p-1, p-2, \ldots, 1$ gives a procedure to compute the last block column and the last block row of W. The next step is to compute the last block column and last block row of $W(1:p-1, 1:p-1)$. To this purpose we introduce the notation $\mathcal{W}_{11} := W(1:p-1,1:p-1)$, $\mathcal{R}_{11} = R(1:p-1,1:p-1)$, $\mathcal{S}_{11} := S(1:p-1,1:p-1)$, and $\mathcal{E}_{11} = E(1:p-1,1:p-1)$, and partition the matrices W, R, S and E in (3.3) as follows:

$$W = \begin{bmatrix} \mathcal{W}_{11} & \mathcal{W}_{12} \\ \mathcal{W}_{21} & W_{pp} \end{bmatrix},\; R = \begin{bmatrix} \mathcal{R}_{11} & \mathcal{R}_{12} \\ 0 & R_{pp} \end{bmatrix},\; S = \begin{bmatrix} \mathcal{S}_{11} & \mathcal{S}_{12} \\ 0 & S_{pp} \end{bmatrix},\; W = \begin{bmatrix} \mathcal{E}_{11} & \mathcal{E}_{12} \\ \mathcal{E}_{21} & E_{pp} \end{bmatrix}. \tag{3.9}$$

Here $\mathcal{W}_{21} = W(p, 1:p-1)$, $\mathcal{W}_{12} = W(1:p-1,p)$ and W_{pp} are known. With the partitions (3.9), the block entries $(1:p-1, 1:p-1)$ of equation (3.3) can be written

$$\mathcal{R}_{11}\mathcal{W}_{11} + \mathcal{W}_{11}^\star \mathcal{S}_{11}^\star = \mathcal{E}_{11} - \mathcal{R}_{12}\mathcal{W}_{21} - \mathcal{W}_{21}^\star \mathcal{S}_{12}^\star. \tag{3.10}$$

Observe that equation (3.10) for the unknown $\mathcal{W}_{11}$ is of the same type as equation (3.3). Therefore, the last block column and the last block row of $\mathcal{W}_{11}$ can be computed in the same way as the last block column and the last block row of W. This discussion leads us to Algorithm 3.2.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

ELA

ALGORITHM 3.2. (**Solution of $\mathbf{RW} + \mathbf{W}^\star\mathbf{S}^\star = \mathbf{E}$ for (quasi) triangular coefficient matrices**) Given $E = [E_{ij}]_{i,j=1}^p \in \mathbb{F}^{n\times n}$, $R = [R_{ij}]_{i,j=1}^p \in \mathbb{F}^{n\times n}$ upper triangular if $\mathbb{F} = \mathbb{C}$ and upper quasi-triangular if $\mathbb{F} = \mathbb{R}$, and $S = [S_{ij}]_{i,j=1}^p \in \mathbb{F}^{n\times n}$ upper triangular, with $E_{ij}, R_{ij}, S_{ij} \in \mathbb{F}^{n_i\times n_j}$ for $1 \le i, j \le p$ and $n_k = 1$ or 2 for $1 \le k \le p$, such that the pencil $R - \lambda S$ satisfies the conditions 1) and 2) in Lemma 2.1, this algorithm computes the unique solution $W \in \mathbb{F}^{n\times n}$ of $RW + W^\star S^\star = E$ in $O(n^3)$ flops. The solutions of the (matrix) equations appearing in the algorithm are computed by GEPP applied to the corresponding vectorized linear systems.

for $\quad j = p : -1 : 1$
  solve $\quad R_{jj}W_{jj} + W_{jj}^\star S_{jj}^\star = E_{jj}$ to get W_{jj}
  for $\quad i = j-1 : -1 : 1$
    solve $\left\{ \begin{array}{l} S_{ii}W_{ij} + W_{ji}^\star R_{jj}^\star = E_{ji}^\star - \sum_{k=i+1}^{j} S_{ik}W_{kj} \\ R_{ii}W_{ij} + W_{ji}^\star S_{jj}^\star = E_{ij} - \sum_{k=i+1}^{j} R_{ik}W_{kj} \end{array} \right\}$ to get W_{ij}, W_{ji}
  end
  $E(1:j-1, 1:j-1) = E(1:j-1, 1:j-1) - R(1:j-1, j)W(j, 1:j-1)$
  $\qquad -(S(1:j-1, j)W(j, 1:j-1))^\star$
end

Note that in the last line of Algorithm 3.2, we have used again MATLAB's notation for submatrices through block-indices, as it was explained above.

Let us analyze the computational costs of Algorithms 3.1 and 3.2. Assume first that $\mathbb{F} = \mathbb{R}$. The cost of Algorithm 3.2 is $2n^3 + O(n^2)$ flops, if $R_{ii} \in \mathbb{R}^{1\times 1}$ for all i. The cost of the QZ algorithm in Step 1 of Algorithm 3.1 is $66n^3 + O(n^2)$ flops (see [15, p. 385]). In addition, Steps 2 and 4 in Algorithm 3.1 amount to 4 matrix multiplications of $n \times n$ matrices. Therefore the total cost of Algorithm 3.1 is $76n^3 + O(n^2)$ flops. If $\mathbb{F} = \mathbb{C}$, this cost is multiplied by a factor up to 6.

The way Algorithm 3.2 is written shows clearly that is of the same type as the classical Bartels-Stewart algorithm for the standard Sylvester equation with quasi-triangular coefficients [2]. However, it is known that the Bartels-Stewart algorithm may perform poorly in modern computer architectures, due to the dominance of level-2 BLAS operations. This has motivated the development of recursive blocked algorithms for the Sylvester equation that take advantage of level-3 BLAS operations [21]. Therefore, it might be also more efficient to use a recursive blocked formulation to solve the Sylvester equation for $\star$-congruence with quasi-triangular coefficients.

3.1. Rounding error analysis of Algorithm 3.1. The rounding error analysis of Algorithm 3.1 is standard and very similar to the one of the classical Bartels-Stewart algorithm [2] for the Sylvester equation $AX - XB = C$. As a consequence, we only sketch the main ideas in the style of [16, Section 2] or [17, Section 16.1].

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

The QZ algorithm used in Step 1 of Algorithm 1 is normwise backward stable [15, pp. 385-386]. In addition, floating point multiplication by unitary (orthogonal) matrices that are products of Householder and/or Givens transformations is also a normwise backward stable process [8, Section 3.4.3], [17, Lemmas 19.3 and 19.9]. Therefore, Steps 2 and 4 of Algorithm 3.1 are also normwise backward stable. It only remains to analyze Step 3, that is, Algorithm 3.2. For brevity, we focus only in the case $\star = T$. The case $\star = *$ is similar, although somewhat more complicated since in order to get a linear equation, it is necessary to separate the Sylvester equation for $*$-congruence into its real and imaginary parts.

Let $\widehat{R}$ and $\widehat{S}$ be the matrices computed in Step 1 of Algorithm 3.1, and $\widehat{E}$ the matrix computed in Step 2. Recall that the equation $\widehat{R}W + W^T\widehat{S}^T = \widehat{E}$ can be written as the standard linear system

$$\left[(I_n \otimes \widehat{R}) + (\widehat{S} \otimes I_n)\Pi \right] \operatorname{vec}(W) = \operatorname{vec}(\widehat{E}).$$

Suppose that we permute the entries of $\operatorname{vec}(W) \in \mathbb{C}^{n^2}$ to put, *starting from the bottom*, the vectors $\operatorname{vec}(W_{ij})$, $i, j = 1, \ldots, p$, corresponding to the blocks in (3.5) *in the order that they are computed by Algorithm 3.2* (we insist again in the fact that in the complex case $\mathbb{F} = \mathbb{C}$ all blocks are 1×1, and $\operatorname{vec}(W_{ij})$ are simply equal to the entries w_{ij} of W). Let us denote the vector so obtained by $\Pi_2\operatorname{vec}(W)$, where $\Pi_2 \in \mathbb{C}^{n^2\times n^2}$ is a certain permutation matrix. Observe now that Algorithm 3.2 is equivalent in floating point arithmetic to solve the (block) upper triangular linear system

$$P\,(\,\Pi_2\operatorname{vec}(W)\,) = \operatorname{vec}(\widehat{E}), \quad \text{where} \quad P = \left[(I_n \otimes \widehat{R}) + (\widehat{S} \otimes I_n)\Pi \right] \Pi_2^T, \tag{3.11}$$

by (block) backward substitution. The matrix P is (block) upper triangular since we can compute each $\operatorname{vec}(W_{ij})$ after computing the entries below it in $(\,\Pi_2\operatorname{vec}(W)\,)$, *for every upper (quasi) triangular matrix $\widehat{R}$ and every upper triangular matrix $\widehat{S}$*. The solution of the system (3.11) by (block) backward substitution is again a normwise backward stable process, under the very mild assumption that GEPP computes in a backward stable way all $\operatorname{vec}(W_{jj})$ and $\left[(\operatorname{vec}(W_{ij}))^T \quad (\operatorname{vec}(W_{ji}))^T \right]^T$, for $i \neq j$. This follows from well-known results on solving linear systems by block algorithms (apply [17, Theorem 13.6] noting that in our case it is not necessary to compute a block LU factorization since the system (3.11) is already block upper triangular). Therefore, Algorithm 3.2 computes a solution $\operatorname{vec}(\widehat{W})$ of (3.11) such that

$$(P + \Delta P)\,(\,\Pi_2\operatorname{vec}(\widehat{W})\,) = \operatorname{vec}(\widehat{E}), \tag{3.12}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

with

$$\|\Delta P\|_F \leq \beta\, u\, n^2 \, \|P\|_F \leq \beta\, u\, n^2 \, \left\| \left[\, (I_n \otimes \widehat{R}) + (\widehat{S} \otimes I_n)\Pi \, \right] \Pi_2^T \right\|_F$$

$$\leq \beta\, u\, n^{5/2} \, \left(\|\widehat{R}\|_F + \|\widehat{S}\|_F \right), \tag{3.13}$$

where u denotes the unit roundoff, β a small integer constant and $\|\cdot\|_F$ the Frobenius norm. The backward error bound $\|\Delta P\|_F \leq \beta\, u\, n^2 \, \|P\|_F$ comes essentially from the traditional error analysis of backward substitution in [17, Theorem 8.5], taking into account that the size of the system is in this case $n^2 \times n^2$. The fact that the system is *block* upper triangular does not change the dependence n^2 on the size of the error, but it may change the numerical constants. Now, let $\|\cdot\|_2$ be the Euclidean vector norm, then from (3.12)-(3.13), we obtain for the residual

$$\|\widehat{R}\widehat{W} + \widehat{W}^T \widehat{S}^T - \widehat{E}\|_F = \|P(\,\Pi_2 \mathrm{vec}(\widehat{W})\,) - \mathrm{vec}(\widehat{E})\|_2 = \|(\Delta P)\,(\,\Pi_2 \mathrm{vec}(\widehat{W})\,)\|_2$$

$$\leq \|(\Delta P)\|_F \, \|\mathrm{vec}(\widehat{W})\|_2$$

$$\leq \beta\, u\, n^{5/2} \, \left(\|\widehat{R}\|_F + \|\widehat{S}\|_F \right) \, \|\widehat{W}\|_F. \tag{3.14}$$

The residual bound (3.14) can be combined with the backward errors of the QZ algorithm and the multiplication by unitary matrices to show that the solution $\widehat{X}$ computed by Algorithm 3.1 satisfies

$$\|A\widehat{X} + \widehat{X}^T B - C\|_F \leq \alpha\, u\, n^{5/2} \, (\|A\|_F + \|B\|_F) \, \|\widehat{X}\|_F, \tag{3.15}$$

where α is a small integer constant independent of the size of the matrices.

Equation (3.15) proves that Algorithm 3.1 computes solutions with tiny relative residual of order unit roundoff. However this does not guarantee a small backward error in the input matrices A, B and C. In this respect, Algorithm 3.1 for the Sylvester equation for $\star$-congruence has a similar behavior to the Bartels-Stewart algorithm for the standard Sylvester equation [16], [17, Section 16.2]. We plan to study in near future the backward error for the Sylvester equation for $\star$-congruence.

4. Conclusions. We have reviewed necessary and sufficient conditions for the existence and uniqueness of solutions of the Sylvester equation for $\star$-congruence. These conditions were proved by Wimmer [29] and Byers, Kressner, Schröder, and Watkins [5, 23]. In this review, we have extended to any field of characteristic different from two, and $\star = T$ or $*$, the original Wimmer's condition, which has required to develop a new proof. Wimmer's characterization of consistency is in the spirit of Roth's criterion for the standard Sylvester equation. However, both criteria are very different, because Roth's criterion involves block diagonalization through similarity transformations, while Wimmer's condition involves block anti-diagonalization through $\star$-congruence transformations. When the solution of the Sylvester equation

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

ELA

for $\star$-congruence is unique for every right-hand side, according to the conditions by Kressner, Schröder, and Watkins [23], we have developed a numerical method to compute efficiently its solution based on the generalized Schur decomposition of the pair $(A, B^\star)$. This method requires to use the QZ algorithm for matrix pencils, which represents a significant difference with respect the classical Bartels-Stewart algorithm to solve the standard Sylvester equation that does not require to deal with matrix pencils. The rounding errors committed by the new algorithm have been analyzed and we have shown that it produces a relative residual of order of the unit roundoff of the computer. In addition, this work may motivate to investigate several open problems as, for instance, to study the set of solutions of the Sylvester equation for $\star$-congruence when the solution is not unique, to develop the perturbation theory for this equation, and the analysis of the backward errors committed by the new algorithm that we have presented.

Acknowledgments. The authors wish to thank Prof. Daniel Kressner for fruitful discussions and Prof. Harald K. Wimmer for pointing out reference [29]. They also thank two anonymous referees for pointing out reference [7] and for helpful comments.

REFERENCES

[1] C.S. Ballantine. A note on the matrix equation $H = AP + PA^*$. *Linear Algebra Appl.*, 2:37–47, 1969.

[2] R.H. Bartels and G.W. Stewart. Algorithm 432: Solution of the matrix equation $AX+XB = C$. *Comm. ACM*, 15:820–826, 1972.

[3] M.A. Beitia and J.-M. Gracia. Sylvester matrix equation for matrix pencils. *Linear Algebra Appl.*, 232:155–197, 1996.

[4] H.W. Braden. The equations $A^T X \pm X^T A = B$. *SIAM J. Matrix Anal. Appl.*, 20:295–302, 1999.

[5] R. Byers and D. Kressner. Structured condition numbers for invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 28:326–347, 2006.

[6] C.-Y. Chiang, E.K.-W. Chu, and W.-W. Lin. On the $\star$-Sylvester equation $AX + X^\star B^\star = C$. Submitted, 2011. Available at `http://oz.nthu.edu.tw/~d947207/colloquium/20081020/TSylvester_081013.pdf`.

[7] D.S. Cvetković-Ilić. The solutions of some operator equations. *J. Korean Math. Soc.*, 45:1417–1425, 2008.

[8] J.W. Demmel. *Applied Numerical Linear Algebra.* SIAM, Philadelphia, PA, 1997.

[9] F. De Terán and F.M. Dopico. The solution of the equation $XA+AX^T = 0$ and its application to the theory of orbits. *Linear Algebra Appl.*, 434:44–67, 2011.

[10] F. De Terán and F.M. Dopico. The equation $XA+AX^* = 0$ and the dimension of *-congruence orbits. *Electron. J. Linear Algebra*, 22:448–465, 2011.

[11] D.S. Djordjević. Explicit solution of the operator equation $A^*X + X^*A = B$. *J. Comput. Appl. Math.*, 200:701–704, 2007.

[12] H. Flanders and H.K. Wimmer. On the matrix equations $AX - XB = C$ and $AX - YB = C$. *SIAM J. Appl. Math.*, 32:707–710, 1977.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 22, pp. 849-863, September 2011

[13] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials.* Academic Press, New York, 1982.

[14] G.H. Golub, S. Nash, and C.F. Van Loan. A Hessenberg-Schur method for the problem $AX + XB = C$. *IEEE Trans. Automat. Control*, 24:909-913, 1979.

[15] G.H. Golub and C.F. Van Loan. *Matrix Computations,* Third Edition. Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[16] N.J. Higham. Perturbation theory and backward error for $AX - XB = C$. *BIT*, 33:124–136, 1993.

[17] N.J. Higham. *Accuracy and Stability of Numerical Algorithms,* Second Edition. SIAM, Philadelphia, PA, 2002.

[18] J.H. Hodges. Some matrix equations over a finite field. *Ann. Mat. Pura Appl.*, 44:245–250, 1957.

[19] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.

[20] R.A. Horn and V.V. Sergeichuk. Canonical forms for complex matrix congruence and *congruence. *Linear Algebra Appl.*, 416:1010–1032, 2006.

[21] I. Jonsson and B. Kågström. Recursive blocked algorithms for solving triangular systems. Part I: One-sided and coupled Sylvester-type matrix equations. *ACM Trans. Math. Software*, 28:392–415, 2002.

[22] B. Kågström and L. Westin. Generalized Schur methods with condition estimators for solving the generalized Sylvester equation. *IEEE Trans. Automat. Control*, 34:745–751, 1989.

[23] D. Kressner, C. Schröder, and D.S. Watkins. Implicit QR algorithms for palindromic and even eigenvalue problems. *Numer. Algorithms*, 51:209–238, 2009.

[24] W.E. Roth. The equations $AX - YB = C$ and $AX - XB = C$ in matrices. *Proc. Amer. Math. Soc.*, 3:392–396, 1952.

[25] G.W. Stewart and J.-G Sun. *Matrix Perturbation Theory.* Academic Press, Inc., Boston. MA, 1990.

[26] V.L. Syrmos and F.L. Lewis. Decomposability and quotient subspaces for the pencil $sL - M$. *SIAM J. Matrix Anal. Appl.*, 15:865–880, 1994.

[27] O. Taussky and H. Wielandt. On the matrix function $AX + X'A'$. Arch. Rational Mech. Anal., 9:93–96, 1962.

[28] H.K. Wimmer. Consistency of a pair of generalized Sylvester equations. *IEEE Trans. Automatic Control*, 39:1014–1016, 1994.

[29] H.K. Wimmer. Roth's theorems for matrix equations with symmetry constraints. *Linear Algebra Appl.*, 199:357–362, 1994.