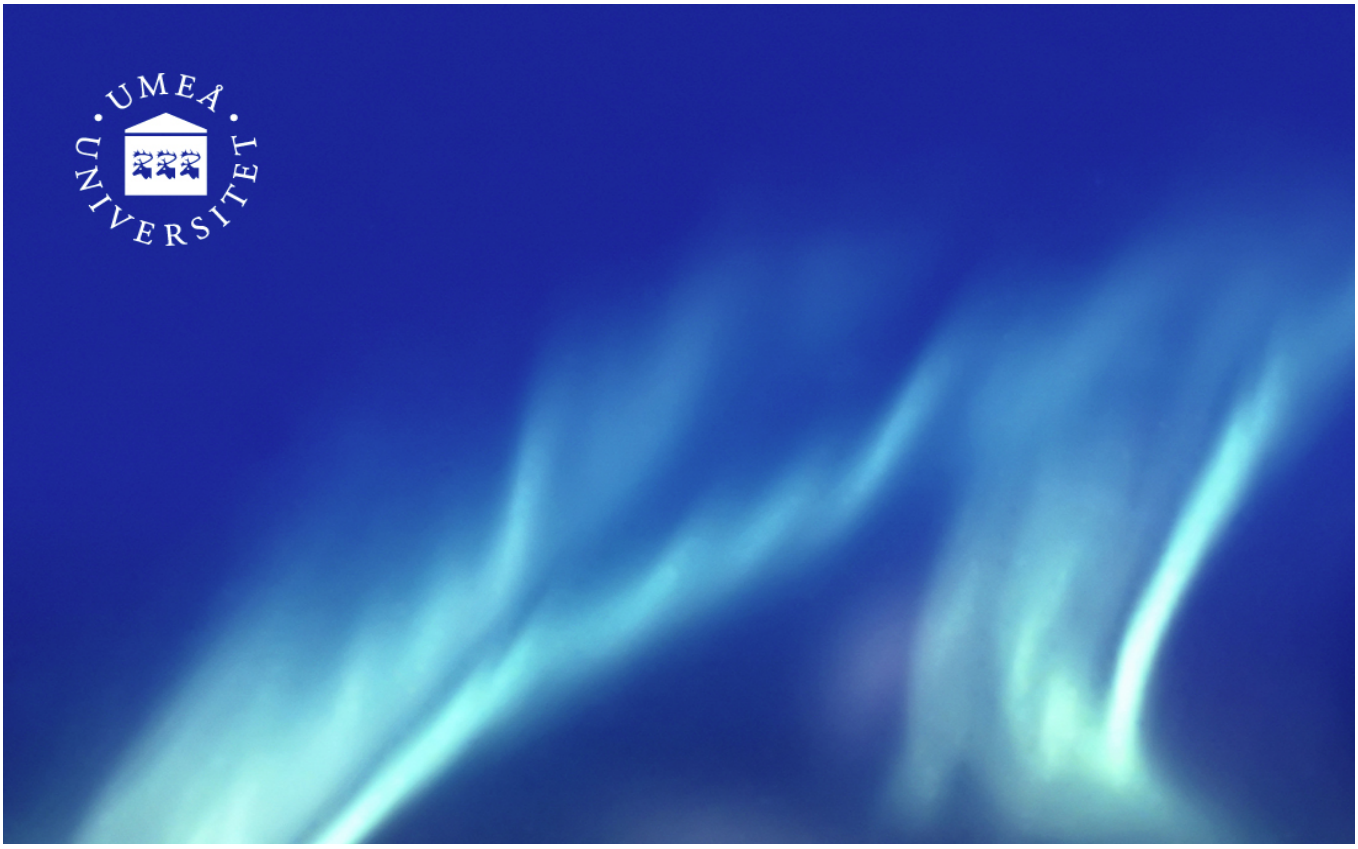


Coupled Sylvester-type Matrix Equations and Block Diagonalization

by

Andrii Dmytryshyn and Bo Kågström

UMINF-15/04

UMEÅ UNIVERSITY
DEPARTMENT OF COMPUTING SCIENCE
SE-901 87 UMEÅ
SWEDEN

COUPLED SYLVESTER-TYPE MATRIX EQUATIONS AND BLOCK DIAGONALIZATION*

ANDRII DMYTRYSHYN† AND BO KÅGSTRÖM†

Abstract. We prove Roth's type theorems for systems of matrix equations including an arbitrary mix of Sylvester and $\star$-Sylvester equations, in which also the transpose or conjugate transpose of the unknown matrices appear. In full generality, we derive consistency conditions by proving that such a system has a solution if and only if the associated set of 2×2 block matrix representations of the equations are block diagonalizable by (linked) equivalence transformations. Various applications leading to several particular cases have already been investigated in the literature, some recently and some long ago. Solvability of these cases follow immediately from our general consistency theory. We also show how to apply our main result to systems of Stein-type matrix equations.

Key words. matrix equation, Sylvester equation, Stein equation, Roth's theorem, consistency, block diagonalization

AMS subject classifications. 15A24, 15A21, 15A63, 65F99

1. Introduction. Let $\mathbb{F}$ be a field of characteristic different from two and let $A_i, B_i, C_i, F_{i'}, G_{i'}, H_{i'}$, and X_k be matrices over $\mathbb{F}$. This includes matrices over real or complex numbers which appear frequently in various applications. We investigate the system of $n_1 + n_2$ matrix equations with m unknowns

$$\begin{aligned} A_i X_k \pm X_j B_i &= C_i, \qquad i = 1, \ldots, n_1, \\ F_{i'} X_{k'} \pm X_{j'}^{\star} G_{i'} &= H_{i'}, \quad\ i' = 1, \ldots, n_2, \end{aligned}$$

where $k, j, k', j' \in \{1, \ldots, m\}$, each unknown X_l is $r_l \times c_l$, $l = 1, \ldots, m$, and all other matrices are of compatible sizes. The $(\cdot)^{\star}$-operator used in $X_l^{\star}$ denotes the matrix transpose X_l^T and, for the field of complex numbers, also the matrix conjugate transpose X_l^H. In the system above with $n_1 + n_2$ matrix equations, each equation has one or at most two different unknown matrices, and thus the total number of unknowns m is obviously bounded by $2n_1 + 2n_2$. In general, m can take any value from 1 to $2n_1 + 2n_2$. The indices k and j depend on i (are integer functions of i) as well as k' and j' depend on i' which reflect that different equations may share the same or have

*The work was supported by the Swedish Research Council (VR) under grant E0485301, and by eSSENCE, a strategic collaborative e-Science programme funded by the Swedish Research Council. A preprint appears as Report UMINF 15.04.

†Department of Computing Science and HPC2N, Umeå University, SE-901 87 Umeå, Sweden. E-mails: andrii@cs.umu.se, bokg@cs.umu.se

different unknown matrices. In general, we will not repeat these index dependencies in the statements of our results, but they will be obvious in some of the examples discussed in Section 2. Without loss of generality, we may consider only the minus signs in the first type of matrix equations and the plus signs in the second type of matrix equations.

For simplicity, we call the two types of matrix equations in the system Sylvester and $\star$-Sylvester equations, respectively, although coupled (or generalized) Sylvester and $\star$-Sylvester equations would be a more accurate description. The same unknown matrices may appear in the matrix equations of the same type, as well as in the matrix equations of different types. This couples all equations on the intra- and inter-type levels, respectively. Depending on the choice of n_1, n_2, and m as well as the positioning (or indexing) of the unknown matrices, there exist several special cases already studied in the literature. Later, we will review some of these particular cases. But first, the simplicity in the statement of our main result allows us to present it already now.

THEOREM 1.1. *The system of $n_1 + n_2$ matrix equations with m unknown matrices*

$$A_i X_k - X_j B_i = C_i, \qquad i = 1, \dots, n_1, \tag{1.1}$$

$$F_{i'} X_{k'} + X_{j'}^{\star} G_{i'} = H_{i'}, \quad i' = 1, \dots, n_2, \tag{1.2}$$

where $k, j, k', j' \in \{1, \dots, m\}$, has a solution $(X_1, X_2, \dots, X_m)$ if and only if there exist nonsingular matrices $P_1, P_2, \dots, P_m$ such that

$$P_j^{-1} \begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} P_k = \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}, \qquad i = 1, \dots, n_1, \tag{1.3}$$

$$P_{j'}^{\star} \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & H_{i'} \end{bmatrix} P_{k'} = \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & 0 \end{bmatrix}, \quad i' = 1, \dots, n_2. \tag{1.4}$$

In Theorem 1.1 we may have $n_1 = 0$ or $n_2 = 0$ which would mean that matrix equations (1.1) or (1.2) are absent as well as the conditions (1.3) or (1.4), respectively. Theorem 1.1 relates the consistency, i.e., the property of having a solution, of a system of matrix equations (1.1)–(1.2) to the block diagonalization and the block anti-diagonalization of the corresponding set of block-triangular matrices (1.3)–(1.4). Notably, Theorem 1.1 also covers the cases of existence of (skew-)hermitian or (skew-)symmetric solutions since we can add the equations $X_l \pm X_l^{\star} = 0$ for the variables we want to satisfy the corresponding condition (see also [39] for this result for one equation).

Since the first paper on the consistency of the matrix equations $AX - XB = C$ and $AX - YB = C$, published by Roth in 1952 [29], several particular cases of Theorem 1.1 have been proven, e.g., see [2, 12, 21, 22, 28, 29, 32, 38, 39]. Roth's results and their generalizations are often referred in the literature as Roth's theorems.

A great number of problems lead to Sylvester and $\star$-Sylvester matrix equations; examples include robust, optimal, and singular system control, signal processing, filtering techniques, feedback, model reduction, numerical solution of differential equations, e.g., see [3, 4, 10, 12, 34, 35] and references therein; see also Section 2. Besides applications, our motivation for developing this theory in full generality is the close relationship to the problem of reduction of block-triangular matrices (by a particular equivalence transformation) to the corresponding block diagonal and anti-diagonal forms, see (1.3)–(1.4).

The consistency conditions for some systems of matrix equations have been stated not only through the corresponding equivalence relations of the block matrices but also via, e.g., ranks or generalized inverses [1, 34, 35]. Nevertheless, many of these conditions can be derived from the corresponding equivalence relations of the block triangular matrices in Theorem 1.1.

In order to appreciate and grasp the generality of Theorem 1.1 we use a tool from the representation theory as follows; any set of sesqui- or bilinear, and linear mappings (and thus any system of Sylvester and $\star$-Sylvester equations (1.1)–(1.2)) is associated to a graph with undirected and directed edges [23]. We illustrate the graph representation for several particular cases.

Summarizing, Theorem 1.1 generalizes Roth's theorem to systems consisting of Sylvester and $\star$-Sylvester equations. In Section 2, we discuss a few particular cases and motivations from the literature. Section 3 illustrates how we associate each possible corollary of Theorem 1.1 with a graph representation. The next two sections are dedicated to the proofs: In Section 4, we prove the result for systems of Sylvester equations; Section 5 solves the general case of mixed systems of Sylvester and $\star$-Sylvester equations. Finally, in Section 6, we apply our main result and derive a similar consistency theorem for systems of Stein and $\star$-Stein equations.

2. Particular cases and applications. In applications, we typically deal with a particular system of matrix equations or a set of matrices. Here, we discuss a selection of particular cases, some of them already addressed in the literature, and formulate them as corollaries to Theorem 1.1. Note that not only n_1, n_2, and m define the settings but also the positions of the unknowns in every equation of the system (1.1)–(1.2).

First, we consider a system of n standard (continuous-time) Sylvester equations, i.e. $n_2 = 0$ and $m = 1$ in Theorem 1.1, leading to the following Corollary, which also has been proved in [28], and in [21] for matrices over commutative rings.

COROLLARY 2.1. *The system of matrix equations*

$$A_i X - X B_i = C_i, \quad i = 1, \dots, n \tag{2.1}$$

has a solution X if and only if there exists a nonsingular matrix P such that

$$P^{-1}\begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} P = \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}, \quad i = 1, \dots, n. \tag{2.2}$$

As mentioned, Corollary 2.1 for only one equation ($n = 1$) is done in [15, 29] and relates the consistency (existence of the solution) of the matrix equation with the similarity (having the same Jordan canonical form) of the two associated 2×2 block matrices. A comprehensive review of the various results about the matrix equation $AX - XB = C$ as well as an extensive reference list are presented in [4].

Setting $n_2 = 0, m = 2$, and fixing the positions of the two unknown matrices in Theorem 1.1 we get the following corollary, which is also proved in [21, 28].

Corollary 2.2. *The system of matrix equations*

$$A_i X_1 - X_2 B_i = C_i, \quad i = 1, \dots, n \tag{2.3}$$

has a solution (X_1, X_2) if and only if there exist nonsingular matrices P_1 and P_2 such that

$$P_2^{-1}\begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} P_1 = \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}, \quad i = 1, \dots, n. \tag{2.4}$$

Two important partial cases of Corollary 2.2 are those with one and two equations. The case with only one equation ($n = 1$) is the initial result by Roth [29] (see also [1, 15]). For $n = 2$ this result is important for the investigation of matrix pencils as is done in [2, 32, 38]. In particular, these systems arise in computing stable eigen-decompositions of matrix pencils [7]. Robust and efficient algorithms and software for solving these generalized Sylvester equations have been developed, e.g., see [27], RECSY [24, 25], SCASY [17, 18], and [26] for a perturbation analysis.

Maybe, a more exotic partial case of Theorem 1.1 ($n_1 = 2, n_2 = 0$, and $m = 3$), see [28], is as follows.

Corollary 2.3. *The system of matrix equations*

$$\begin{aligned} A_1 X_1 - X_2 B_1 &= C_1, \\ A_2 X_3 - X_2 B_2 &= C_2 \end{aligned} \tag{2.5}$$

has a solution (X_1, X_2, X_3) if and only if there exist nonsingular matrices P_1, P_2, and P_3 such that

$$P_2^{-1}\begin{bmatrix} A_1 & C_1 \\ 0 & B_1 \end{bmatrix} P_1 = \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix} \quad \textit{and} \quad P_2^{-1}\begin{bmatrix} A_2 & C_2 \\ 0 & B_2 \end{bmatrix} P_3 = \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}. \tag{2.6}$$

Let us state another particular case of Theorem 1.1 ($n_1 = m, n_2 = 0$, and the "cyclic" positioning of the unknowns) which is of interest due to its relations to the periodic eigenvalue problem [5, 16] and its deflating subspaces [19, 20]. Surprisingly, the consistency of this particular case does not seem to be explicitly stated and published before.

Corollary 2.4. *The cyclic system of matrix equations*

$$\begin{aligned} A_1X_1 - X_2B_1 &= C_1, \\ A_2X_2 - X_3B_2 &= C_2, \\ &\cdots \\ A_{n-1}X_{n-1} - X_nB_{n-1} &= C_{n-1}, \\ A_nX_n - X_1B_n &= C_n \end{aligned} \tag{2.7}$$

has a solution $(X_1, X_2, \dots, X_n)$ if and only if there exist nonsingular matrices $P_1, P_2, \dots, P_n$ such that

$$\begin{aligned} P_2^{-1}\begin{bmatrix} A_1 & C_1 \\ 0 & B_1 \end{bmatrix}P_1 &= \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}, \\ P_3^{-1}\begin{bmatrix} A_2 & C_2 \\ 0 & B_2 \end{bmatrix}P_2 &= \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}, \\ &\cdots \\ P_n^{-1}\begin{bmatrix} A_{n-1} & C_{n-1} \\ 0 & B_{n-1} \end{bmatrix}P_{n-1} &= \begin{bmatrix} A_{n-1} & 0 \\ 0 & B_{n-1} \end{bmatrix}, \\ P_1^{-1}\begin{bmatrix} A_n & C_n \\ 0 & B_n \end{bmatrix}P_n &= \begin{bmatrix} A_n & 0 \\ 0 & B_n \end{bmatrix}. \end{aligned} \tag{2.8}$$

We remark that Corollary 2.4 for $n = 2$ gives Roth's theorem for contragredient matrix pencils.

So far, the systems in Corollaries 2.1–2.4 only involve Sylvester equations of type (1.1). Notwithstanding, $\star$-Sylvester equations recently enjoyed considerable attention [10, 11, 13, 14], partially, due to their relations with the congruence orbit of palindromic matrix pencils [12], e.g., in the analysis of associated deflating subspaces [6]. Therefore, the following corollary (Theorem 1.1 for $n_1 = 0$ and $n_2 = m = 1$) is already known [12, 39].

Corollary 2.5. *The matrix equation*

$$FX - X^{\star}G = H,$$

has a solution X if and only if there exists a nonsingular matrix P such that

$$P^{\star}\begin{bmatrix} 0 & G \\ F & H \end{bmatrix} P = \begin{bmatrix} 0 & G \\ F & 0 \end{bmatrix}. \tag{2.9}$$

Notably, Corollary 2.5 has recently been generalized to several $\star$-Sylvester equations [9] (Theorem 1.1 for $n_1 = 0$ and $m = 1$).

Additionally, in [39] a particular case of Theorem 1.1 ($n_1 = n_2 = m = 1$, and $F = I, G = -I$, and $H = 0$) with both Sylvester and $\star$-Sylvester equations is addressed.

Corollary 2.6. *The system of matrix equations*

$$\begin{aligned} AX - XB &= C, \\ X - X^{\star} &= 0 \end{aligned} \tag{2.10}$$

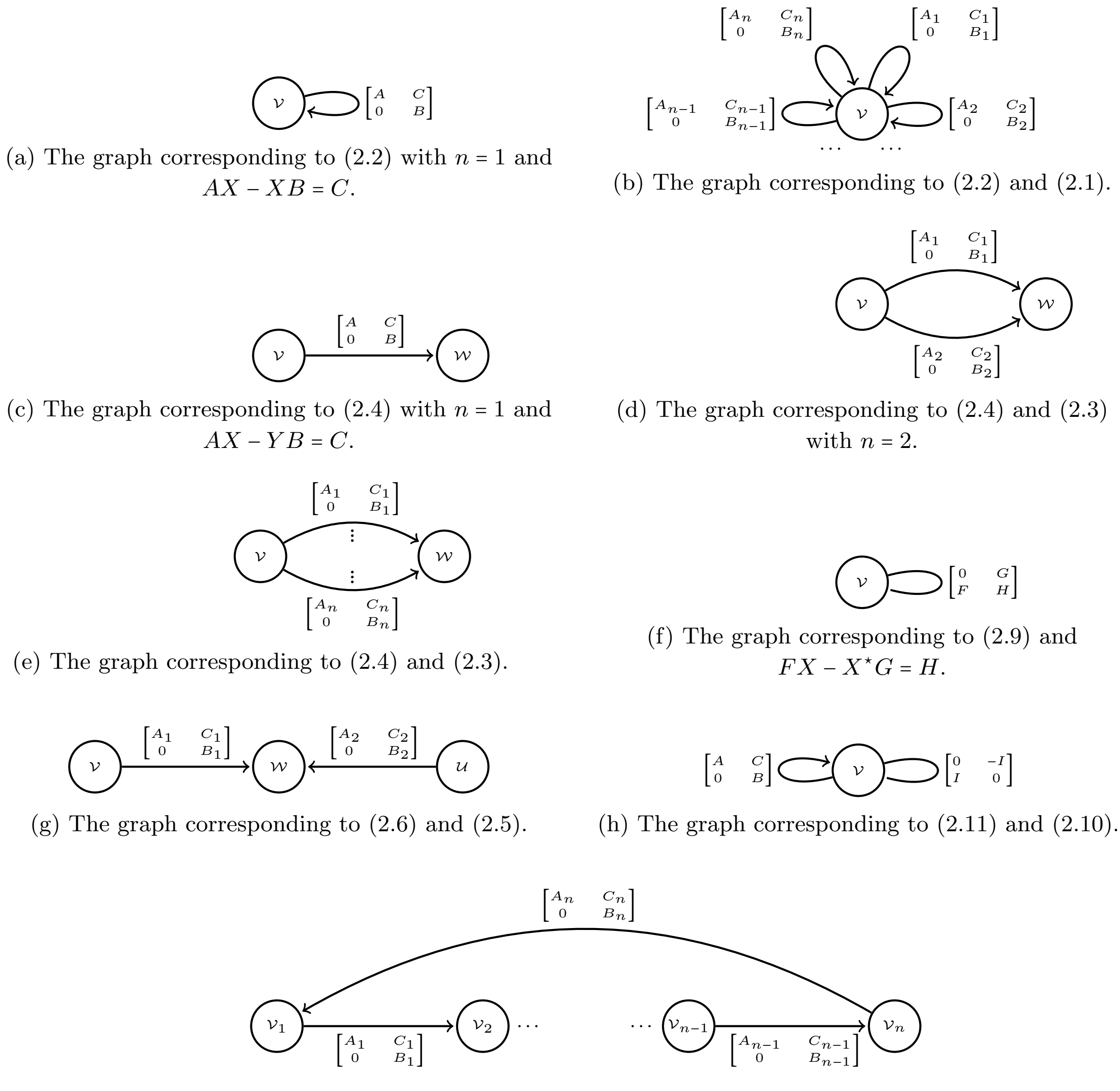
has a solution X if and only if there exists a nonsingular matrix P such that

$$\begin{aligned} P^{-1}\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} P &= \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}, \\ P^{\star}\begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} P &= \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}. \end{aligned} \tag{2.11}$$

Essentially, Corollary 2.6 covers the existence of a hermitian solution for the standard Sylvester matrix equation.

3. Using graphs to represent linear mappings and matrix equations. We use a tool from representation theory and associate a graph with each particular case of Theorem 1.1. This method of visualization is inspired by the representation theory of mixed graphs, developed by Sergeichuk [31] (see also [23]). A *mixed graph* may have both directed and undirected edges. Its *representation* is given by assigning to each vertex a vector space, to each directed edge a linear mapping, and to each undirected edge a bilinear or sesquilinear form on the corresponding vector spaces. We can express these mappings and forms by their matrices choosing bases in the spaces. Reselection the bases reduces by equivalence transformations all matrices that are assigned to directed edges, and by (*)congruence transformations all matrices that are assigned to undirected edges (see [23] for details).

For example, for $n = 1$, (2.2) essentially means that the corresponding 2×2 block matrices are similar (i.e., have the same Jordan canonical form) and the block diagonalization can be interpreted as changing the matrix representing the mapping from some vector space $\mathcal{V}$ to itself when the basis of $\mathcal{V}$ is modified (multiplied) by a transformation matrix P^{-1}. Thus, with $n = 1$, the similarity (2.2) can be associated



(a) The graph corresponding to (2.2) with $n = 1$ and $AX - XB = C$.

(b) The graph corresponding to (2.2) and (2.1).

(c) The graph corresponding to (2.4) with $n = 1$ and $AX - YB = C$.

(d) The graph corresponding to (2.4) and (2.3) with $n = 2$.

(e) The graph corresponding to (2.4) and (2.3).

(f) The graph corresponding to (2.9) and $FX - X^{\star}G = H$.

(g) The graph corresponding to (2.6) and (2.5).

(h) The graph corresponding to (2.11) and (2.10).

(i) The graph corresponding to (2.8) and (2.7).

FIG. 3.1. *Graph representations associated with Corollaries 2.1–2.5.*

with graph (a) in Figure 3.1: the node represents the vector space $\mathcal{V}$ and the loop represents the mapping with the matrices (in different bases)

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \tag{3.1}$$

acting from $\mathcal{V}$ to itself. Similarly, n mappings acting from $\mathcal{V}$ to itself will result at graph (b) in Figure 3.1.

Like in the previous cases, condition (2.4) for $n = 1$ can be interpreted as changing the matrix representing the mapping from some vector space $\mathcal{V}$ to another vector space $\mathcal{W}$ when the bases of the vector spaces are modified by the transformation matrices P_1^{-1} and P_2^{-1}, respectively, and the associated graph is Figure 3.1(c). In Figure 3.1 we also present the graphs associated with the other corollaries in Section 2.

Now through Theorem 1.1 we essentially associate a graph to a system of matrix ($\star$-)Sylvester equations: The number of nodes in the graph is equal to the number of unknown matrices X_i in the system of matrix equations. The number of directed and undirected edges are equal to the number of Sylvester and $\star$-Sylvester equations, respectively. The edges are placed in between the nodes corresponding to the unknowns involved in the respective matrix equation. The directions of the directed edges are defined by the order of the unknowns involved in the respective matrix equations. Applying the $(\cdot)^\star$-operator to any $\star$-Sylvester equation, we get an equivalent $\star$-Sylvester equation with interchanged order (roles) of the unknowns, which also explains their association with the undirected edges. Summing up, Theorem 1.1 shows that Roth's theorem remains true for the systems (and the matrix equivalence relations) associated with any graphs.

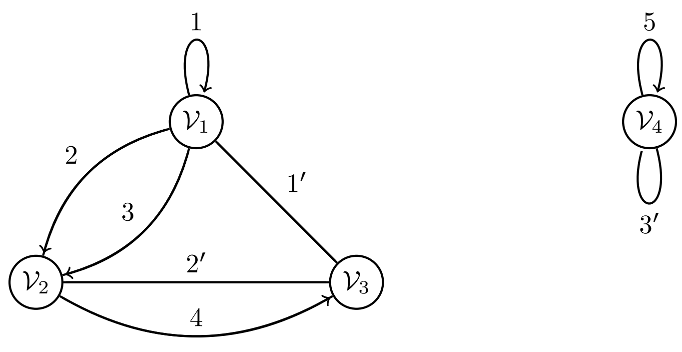


Fig. 3.2. *A disconnected graph with two connected components. Left: Component with 3 nodes, 4 directed and 2 undirected edges associated with 4 Sylvester and 2 $\star$-Sylvester matrix equations with 3 unknown matrices in system* (3.2). *Right: Component with 1 node, 1 directed edge and 1 undirected edge associated with 1 Sylvester and 1 $\star$-Sylvester matrix equation with the unknown matrix X_4 in* (3.2).

For example, we associate the graph in Figure 3.2 with the system of matrix equations as well as with the set of (bi)linear mappings with the corresponding transformations in the left and right columns of (3.2), respectively. Each edge of the graph in Figure 3.2 is labeled with the equation (and mapping) number to which it corresponds:

$$
\begin{array}{llll}
1: & A_1X_1 - X_1B_1 = C_1, & P_1^{-1}\begin{bmatrix} A_1 & C_1 \\ 0 & B_1 \end{bmatrix} P_1 & = \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}, \\
2: & A_2X_1 - X_2B_2 = C_2, & P_2^{-1}\begin{bmatrix} A_2 & C_2 \\ 0 & B_2 \end{bmatrix} P_1 & = \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}, \\
3: & A_3X_1 - X_2B_3 = C_3, & P_2^{-1}\begin{bmatrix} A_3 & C_3 \\ 0 & B_3 \end{bmatrix} P_1 & = \begin{bmatrix} A_3 & 0 \\ 0 & B_3 \end{bmatrix}, \\
4: & A_4X_2 - X_3B_4 = C_4, & P_3^{-1}\begin{bmatrix} A_4 & C_4 \\ 0 & B_4 \end{bmatrix} P_2 & = \begin{bmatrix} A_4 & 0 \\ 0 & B_4 \end{bmatrix}, \\
5: & A_5X_4 - X_4B_5 = C_5, & P_4^{-1}\begin{bmatrix} A_5 & C_5 \\ 0 & B_5 \end{bmatrix} P_4 & = \begin{bmatrix} A_5 & 0 \\ 0 & B_5 \end{bmatrix}, \\
1': & F_1X_3 + X_1^\star G_1 = H_1, & P_1^\star\begin{bmatrix} 0 & G_1 \\ F_1 & H_1 \end{bmatrix} P_3 & = \begin{bmatrix} 0 & G_1 \\ F_1 & 0 \end{bmatrix}, \\
2': & F_2X_2 + X_3^\star G_2 = H_2, & P_3^\star\begin{bmatrix} 0 & G_2 \\ F_2 & H_2 \end{bmatrix} P_2 & = \begin{bmatrix} 0 & G_2 \\ F_2 & 0 \end{bmatrix}, \\
3': & F_3X_4 + X_4^\star G_3 = H_3, & P_4^\star\begin{bmatrix} 0 & G_3 \\ F_3 & H_3 \end{bmatrix} P_4 & = \begin{bmatrix} 0 & G_3 \\ F_3 & 0 \end{bmatrix}.
\end{array}
\tag{3.2}
$$

We remark that the equations numbered 5 and $3'$ in (3.2) are independent from the other matrix equations. As a consequence there is a disconnected graph with two connected components associated with (3.2) in Figure 3.2.

Although the graphs are not used for the proofs they are convenient for the problem description and identification, e.g., the system (2.7) and the linear mappings (2.8) in Corollary 2.4 can be represented as the cycle graph (i) in Figure 3.1.

4. Systems of Sylvester equations. We will first prove Roth's theorem for the system of Sylvester matrix equations

$$A_iX_k \pm X_jB_i = C_i, \quad i = 1,\dots,n \ \text{ and } \ j,k \in \{1,\dots,m\}. \tag{4.1}$$

In this section all the matrices are of compatible sizes over a field $\mathbb{F}$ of any characteristic; assuming that each X_l is $r_l \times c_l$, where $l = 1,\dots,m$, (i.e., the unknown matrices may be rectangular and of different sizes) we have that A_i is $r_j \times r_k$, B_i is $c_j \times c_k$, and C_i is $r_j \times c_k$. From now on we usually skip writing the sizes explicitly but assume all matrices being of compatible sizes. Clearly, it is enough to consider only the minus sign in (4.1).

THEOREM 4.1. *The system of matrix equations*

$$A_iX_k - X_jB_i = C_i, \quad i = 1,\dots,n \ \textit{ and } \ j,k \in \{1,\dots,m\} \tag{4.2}$$

has a solution $X_1, X_2, \dots, X_m$ *if and only if there exist nonsingular matrices* $P_1, P_2, \dots, P_m$ *such that*

$$(4.3)\qquad P_j^{-1}\begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} P_k = \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}, \quad i = 1, \dots, n \ \ \textit{and}\ \ j, k \in \{1, \dots, m\}.$$

Our proof below extends and generalizes on the proofs and the techniques used in [15, 38].

Proof. First, let $X_1, X_2, \dots, X_m$ be a solution of (4.2). Then for each i

$$\begin{bmatrix} I & X_j \\ 0 & I \end{bmatrix}\begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix}\begin{bmatrix} I & -X_k \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_i & -A_iX_k + X_jB_i + C_i \\ 0 & B_i \end{bmatrix}, \quad j, k \in \{1, \dots, m\}.$$

Now, for every $l = 1, \dots, m$ we can choose the transformation matrices in (4.3) as

$$P_l = \begin{bmatrix} I & -X_l \\ 0 & I \end{bmatrix} \quad \text{and} \quad P_l^{-1} = \begin{bmatrix} I & X_l \\ 0 & I \end{bmatrix},$$

where the identity matrices are of conformable sizes.

Vice versa: Now, we assume the existence of nonsingular $P_1, P_2, \dots, P_m$ that fulfil (4.3). Consider the two systems of matrix equations

$$(4.4)\qquad \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix} U_k - U_j \begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} = 0,$$

and

$$(4.5)\qquad \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix} U_k - U_j \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix} = 0,$$

where $i = 1, \dots, n$ and $k, j \in \{1, \dots, m\}$ in both systems. Define the solution spaces of (4.4) and (4.5) by $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively. Note that $(U_1, U_2, \dots, U_m) \in \mathcal{W}_1$ if and only if $(U_1P_1, U_2P_2, \dots, U_mP_m) \in \mathcal{W}_2$—follows directly from (4.3)—and therefore we have

$$(4.6)\qquad \dim \mathcal{W}_1 = \dim \mathcal{W}_2.$$

By a conforming block-partition of the solution in (4.4) each matrix equation in the system looks like

$$\begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}\begin{bmatrix} U_{11}^k & U_{12}^k \\ U_{21}^k & U_{22}^k \end{bmatrix} - \begin{bmatrix} U_{11}^j & U_{12}^j \\ U_{21}^j & U_{22}^j \end{bmatrix}\begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} = 0,$$

and after performing the block matrix multiplications we obtain

$$(4.7)\qquad \begin{bmatrix} A_iU_{11}^k - U_{11}^jA_i & A_iU_{12}^k - U_{11}^jC_i - U_{12}^jB_i \\ B_iU_{21}^k - U_{21}^jA_i & B_iU_{22}^k - U_{21}^jC_i - U_{22}^jB_i \end{bmatrix} = 0, \ i = 1, \dots, n \text{ and } k, j \in \{1, \dots, m\}.$$

We complete the proof by showing that there is a solution $(U_1, U_2, \ldots, U_m)$ of (4.7) (i.e., of (4.4)) such that the blocks $(U_{11}^1, U_{11}^2, \ldots, U_{11}^m)$ are all identity matrices of compatible sizes, since then the $(1,2)$ blocks of (4.7) are zero and $(U_{12}^1, U_{12}^2, \ldots, U_{12}^m)$ is a solution of the system of matrix equations (4.2).

Let us introduce an operator that picks the fist block columns of $(U_1, U_2, \ldots, U_m)$:

$$\varphi(U_1, U_2, \ldots, U_m) \coloneqq \left(\begin{bmatrix} U_{11}^1 \\ U_{21}^1 \end{bmatrix}, \begin{bmatrix} U_{11}^2 \\ U_{21}^2 \end{bmatrix}, \ldots, \begin{bmatrix} U_{11}^m \\ U_{21}^m \end{bmatrix} \right).$$

Moreover, let $\varphi_{\mathcal{W}_1} \coloneqq \varphi_{|\mathcal{W}_1}$ and $\varphi_{\mathcal{W}_2} \coloneqq \varphi_{|\mathcal{W}_2}$ be the restrictions of φ onto the solution spaces $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively. Note that

$$\left(\begin{bmatrix} 0 & U_{12}^1 \\ 0 & U_{22}^1 \end{bmatrix}, \begin{bmatrix} 0 & U_{12}^2 \\ 0 & U_{22}^2 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & U_{12}^m \\ 0 & U_{22}^m \end{bmatrix} \right), \tag{4.8}$$

with zero blocks in the first block columns, belongs to $\mathcal{W}_1$ if and only if (4.8) belongs to $\mathcal{W}_2$. Therefore the kernels of the restricted mappings are the same:

$$\operatorname{Ker} \varphi_{\mathcal{W}_1} = \operatorname{Ker} \varphi_{\mathcal{W}_2}. \tag{4.9}$$

We also have that

$$(U_1, U_2, \ldots, U_m) \in \mathcal{W}_1 \text{ then } \left(\begin{bmatrix} U_{11}^1 & 0 \\ U_{21}^1 & 0 \end{bmatrix}, \begin{bmatrix} U_{11}^2 & 0 \\ U_{21}^2 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} U_{11}^m & 0 \\ U_{21}^m & 0 \end{bmatrix} \right) \in \mathcal{W}_2.$$

Therefore $\operatorname{Im} \varphi_{\mathcal{W}_1} \subseteq \operatorname{Im} \varphi_{\mathcal{W}_2}$. Using (4.6), (4.9), and the rank-nullity identities

$$\dim \operatorname{Ker} \varphi_{\mathcal{W}_1} + \dim \operatorname{Im} \varphi_{\mathcal{W}_1} = \dim \mathcal{W}_1,$$

$$\dim \operatorname{Ker} \varphi_{\mathcal{W}_2} + \dim \operatorname{Im} \varphi_{\mathcal{W}_2} = \dim \mathcal{W}_2,$$

we have that $\dim \operatorname{Im} \varphi_{\mathcal{W}_1} = \dim \operatorname{Im} \varphi_{\mathcal{W}_2}$. Thus the image of the restrictions are the same: $\operatorname{Im} \varphi_{\mathcal{W}_1} = \operatorname{Im} \varphi_{\mathcal{W}_2}$.

Since $(U_1, U_2, \ldots, U_m) = (I, I, \ldots, I) \in \mathcal{W}_2$, we have

$$\left(\begin{bmatrix} I \\ 0 \end{bmatrix}, \begin{bmatrix} I \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} I \\ 0 \end{bmatrix} \right) \in \varphi(\mathcal{W}_2),$$

where the unit matrices I in the two expressions (in general) have different sizes. This fact together with $\operatorname{Im} \varphi_{\mathcal{W}_1} = \operatorname{Im} \varphi_{\mathcal{W}_2}$ imply that

$$\left(\begin{bmatrix} I \\ 0 \end{bmatrix}, \begin{bmatrix} I \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} I \\ 0 \end{bmatrix} \right) \in \varphi(\mathcal{W}_1). \tag{4.10}$$

□

5. Systems of Sylvester and $\star$-Sylvester equations. In the investigation of Roth's theorem for systems of Sylvester and $\star$-Sylvester equations, i.e. Theorem 1.1, we consider (1.3)–(1.4) as well as their conjugate transpose variants. Indeed, this makes it possible to apply Theorem 4.1 presented in Section 4 which in turn lead to the existence of a solution to a certain system of type (4.2). Similar techniques were used for some $\star$-Sylvester equations in [9, 36, 39]. In the following, we present the details of our derivations and further manipulations.

Proof. [Proof of Theorem 1.1] Let $X_1, X_2, \dots, X_m$ be a solution of (1.1)–(1.2). Then for each i

$$\begin{bmatrix} I & X_j \\ 0 & I \end{bmatrix} \begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} \begin{bmatrix} I & -X_k \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_i & -A_iX_k + X_jB_i + C_i \\ 0 & B_i \end{bmatrix},$$
$$\begin{bmatrix} I & 0 \\ -X_{j'}^\star & I \end{bmatrix} \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & H_{i'} \end{bmatrix} \begin{bmatrix} I & -X_{k'} \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & -X_{j'}^\star G_{i'} - F_{i'}X_{k'} + H_{i'} \end{bmatrix},$$

where $j, k, j', k' \in \{1, \dots, m\}$. Now, for every $l = 1, \dots, m$ we can choose the transformation matrices in (1.3)–(1.4) as

$$P_l := \begin{bmatrix} I & -X_l \\ 0 & I \end{bmatrix}, \quad P_l^{-1} = \begin{bmatrix} I & X_l \\ 0 & I \end{bmatrix}, \quad \text{and } P_l^\star = \begin{bmatrix} I & 0 \\ -X_l^\star & I \end{bmatrix}.$$

Vice versa: Assume that there exist nonsingular matrices $P_1, P_2, \dots, P_m$ such that (1.3)–(1.4) hold. We transpose or transpose and take the conjugate of (1.3) and (1.4) and altogether we get the following four equalities:

$$P_j^{-1} \begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} P_k = \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}, \quad i = 1, \dots, n_1,$$
$$P_k^\star \begin{bmatrix} A_i^\star & 0 \\ C_i^\star & B_i^\star \end{bmatrix} P_j^{-\star} = \begin{bmatrix} A_i^\star & 0 \\ 0 & B_i^\star \end{bmatrix}, \quad i = 1, \dots, n_1,$$
$$P_{j'}^\star \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & H_{i'} \end{bmatrix} P_{k'} = \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & 0 \end{bmatrix}, \quad i' = 1, \dots, n_2,$$
$$P_{k'}^\star \begin{bmatrix} 0 & F_{i'}^\star \\ G_{i'}^\star & H_{i'}^\star \end{bmatrix} P_{j'} = \begin{bmatrix} 0 & F_{i'}^\star \\ G_{i'}^\star & 0 \end{bmatrix}, \quad i' = 1, \dots, n_2,$$

where $j, k, j', k' \in \{1, \dots, m\}$. Multiplying some equalities above from left (2nd, 3rd, and 4th) by the block-flip matrix

$$V_l := \begin{bmatrix} 0 & I_{c_l \times c_l} \\ I_{r_l \times r_l} & 0 \end{bmatrix}$$

and from right (2nd) by its (conjugate) transpose, as well as taking into account that $V_l^\star V_l = V_l V_l^\star = I$, we obtain

$$\begin{aligned}
P_j^{-1}\begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} P_k &= \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}, & i &= 1,\dots,n_1,\\
V_k P_k^\star V_k^\star V_k \begin{bmatrix} A_i^\star & 0 \\ C_i^\star & B_i^\star \end{bmatrix} V_j^\star V_j P_j^{-\star} V_j^\star &= V_k \begin{bmatrix} A_i^\star & 0 \\ 0 & B_i^\star \end{bmatrix} V_j^\star, & i &= 1,\dots,n_1,\\
V_{j'} P_{j'}^\star V_{j'}^\star V_{j'} \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & H_{i'} \end{bmatrix} P_{k'} &= V_{j'} \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & 0 \end{bmatrix}, & i' &= 1,\dots,n_2,\\
V_{k'} P_{k'}^\star V_{k'}^\star V_{k'} \begin{bmatrix} 0 & F_{i'}^\star \\ G_{i'}^\star & H_{i'}^\star \end{bmatrix} P_{j'} &= V_{k'} \begin{bmatrix} 0 & F_{i'}^\star \\ G_{i'}^\star & 0 \end{bmatrix}, & i' &= 1,\dots,n_2.
\end{aligned}$$

Defining $Q_l \coloneqq V_l P_l^{-\star} V_l^\star$, where $l = 1,\dots,m$, (thus $Q_l^{-1} \coloneqq V_l P_l^\star V_l^\star$) and performing the matrix multiplications of the remaining V with the 2×2 block matrices we have

$$\begin{aligned}
P_j^{-1}\begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} P_k &= \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}, & i &= 1,\dots,n_1,\\
Q_k^{-1}\begin{bmatrix} B_i^\star & C_i^\star \\ 0 & A_i^\star \end{bmatrix} Q_j &= \begin{bmatrix} B_i^\star & 0 \\ 0 & A_i^\star \end{bmatrix}, & i &= 1,\dots,n_1,\\
Q_{j'}^{-1}\begin{bmatrix} F_{i'} & H_{i'} \\ 0 & G_{i'} \end{bmatrix} P_{k'} &= \begin{bmatrix} F_{i'} & 0 \\ 0 & G_{i'} \end{bmatrix}, & i' &= 1,\dots,n_2,\\
Q_{k'}^{-1}\begin{bmatrix} G_{i'}^\star & H_{i'}^\star \\ 0 & F_{i'}^\star \end{bmatrix} P_{j'} &= \begin{bmatrix} G_{i'}^\star & 0 \\ 0 & F_{i'}^\star \end{bmatrix}, & i' &= 1,\dots,n_2.
\end{aligned}$$

By Theorem 4.1 the system

$$\begin{aligned}
A_i Y_k - Y_j B_i &= C_i, & i &= 1,\dots,n_1,\\
B_i^\star Z_j - Z_k A_i^\star &= C_i^\star, & i &= 1,\dots,n_1,\\
F_{i'} Y_{k'} - Z_{j'} G_{i'} &= H_{i'}, & i' &= 1,\dots,n_2,\\
G_{i'}^\star Y_{j'} - Z_{k'} F_{i'}^\star &= H_{i'}^\star, & i' &= 1,\dots,n_2,
\end{aligned}$$

where $j,k,j',k' \in \{1,\dots,m\}$, with the unknown matrices $(Y_1,\dots,Y_m,Z_1,\dots,Z_m)$ has a solution. After transposing (as well as possibly conjugating) all equations and then adding the first two and the last two subsystems of matrix equations we obtain

$$\begin{aligned}
A_i(Y_k - Z_k^\star) - (Y_j - Z_j^\star)B_i &= 2C_i, & i &= 1,\dots,n_1,\\
F_{i'}(Y_{k'} - Z_{k'}^\star) + (Y_{j'}^\star - Z_{j'})G_{i'} &= 2H_{i'}, & i' &= 1,\dots,n_2,
\end{aligned}$$

where $j,k,j',k' \in \{1,\dots,m\}$. Therefore, $X_l \coloneqq \frac{1}{2}(Y_l - Z_l^\star),\ l = 1,\dots,m$ is a solution to the system of Sylvester and $\star$-Sylvester equations (1.1)–(1.2). □

Remark 5.1. (Connections to the representation theory) In the proof of Theorem 1.1 we essentially "substitute" the $\star$-congruence of each bilinear (or sesquilinear) mapping by the equivalence of the two linear mappings. This (and vice versa) substitution for bilinear mappings over complex fields is a corollary of the proposition in [30]. In the language of the representation theory it means that over the field of complex numbers with trivial (or identity) involution the two representations of mixed graphs are equivalent if and only if the two representations of quivers with involution are equivalent; for more details see [23]. We recall that in Section 3 we use mixed graphs to represent linear mappings and matrix equations.

6. Systems of generalized Stein and $\star$-Stein equations. Every generalized Stein equation can be rewritten as a system of three Sylvester equations by introducing two new unknown matrices, e.g., $AXK - LXB = C$ has a solution X if and only if the following system does

$$\begin{aligned} AZ - YB &= C,\\ LX - Y &= 0,\\ Z - XK &= 0, \end{aligned}$$

where Y and Z are the introduced unknown matrices. The same trick is possible for the generalized $\star$-Stein equations. As a consequence we have the following theorem.

THEOREM 6.1. *The system of matrix equations with m unknown matrices*

$$A_i X_k K_i - L_i X_j B_i = C_i, \qquad i = 1, \dots, n_1, \tag{6.1}$$

$$F_{i'} X_{k'} M_{i'} + N_{i'} X_{j'}^{\star} G_{i'} = H_{i'}, \quad i' = 1, \dots, n_2, \tag{6.2}$$

where $j, k, j', k' \in \{1, \dots, m\}$, has a solution $X_1, X_2, \dots, X_m$ if and only if there exist nonsingular matrices $P_1, P_2, \dots, P_m$, $R_1, R_2, \dots, R_{n_1}$, $Q_1, Q_2, \dots, Q_{n_1}$, $\widetilde{R}_1, \widetilde{R}_2, \dots, \widetilde{R}_{n_2}$, and $\widetilde{Q}_1, \widetilde{Q}_2, \dots, \widetilde{Q}_{n_2}$ such that

$$\begin{aligned}
R_i^{-1} \begin{bmatrix} A_i & C_i \\ 0 & B_i \end{bmatrix} Q_i &= \begin{bmatrix} A_i & 0 \\ 0 & B_i \end{bmatrix}, &\qquad \widetilde{R}_{i'}^{\star} \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & H_{i'} \end{bmatrix} \widetilde{Q}_{i'} &= \begin{bmatrix} 0 & G_{i'} \\ F_{i'} & 0 \end{bmatrix},\\
R_i^{-1} \begin{bmatrix} L_i & 0 \\ 0 & I_i \end{bmatrix} P_j &= \begin{bmatrix} L_i & 0 \\ 0 & I_i \end{bmatrix}, &\qquad P_{j'}^{-1} \begin{bmatrix} I_{i'} & 0 \\ 0 & N_{i'}^{\star} \end{bmatrix} \widetilde{R}_{i'} &= \begin{bmatrix} I_{i'} & 0 \\ 0 & N_{i'}^{\star} \end{bmatrix},\\
P_k^{-1} \begin{bmatrix} I_i & 0 \\ 0 & K_i \end{bmatrix} Q_i &= \begin{bmatrix} I_i & 0 \\ 0 & K_i \end{bmatrix}, &\qquad P_{k'}^{-1} \begin{bmatrix} I_{i'} & 0 \\ 0 & M_{i'} \end{bmatrix} \widetilde{Q}_{i'} &= \begin{bmatrix} I_{i'} & 0 \\ 0 & M_{i'} \end{bmatrix},
\end{aligned}$$

where $i = 1, \dots, n_1$, $i' = 1, \dots, n_2$, and $j, k, j', k' \in \{1, \dots, m\}$.

Proof. Let $Y_i := L_i X_j$ and $Z_i := X_k K_i$ then (6.1) can be rewritten as follows

$$\begin{aligned} A_i Z_i - Y_i B_i &= C_i,\\ L_i X_j - Y_i &= 0,\\ Z_i - X_k K_i &= 0, \end{aligned}$$

where $i = 1, \dots, n_1$, $j, k \in \{1, \dots, m\}$, and let $U_{i'} := X_{j'} N_{i'}^{\star}$ and $V_{i'} := X_{k'} M_{i'}$ then (6.2) can be rewritten as follows

$$\begin{aligned} F_{i'} V_{i'} + U_{i'}^{\star} G_{i'} &= H_{i'}, \\ U_{i'} - X_{j'} N_{i'}^{\star} &= 0, \\ V_{i'} - X_{k'} M_{i'} &= 0, \end{aligned}$$

where $i' = 1, \dots, n_2$ and $j', k' \in \{1, \dots, m\}$. Now the result follows from Theorem 1.1. $\square$

Some particular cases of Theorem 6.1 are also known; examples include the matrix Stein equations $X - LXB = C$ [33, 37] and $X - LYB = C$ [28], and the generalisations $A_i X - LXB_i = C_i,\ i = 1, \dots, n$ and $XK_i - L_i XB = C_i,\ i = 1, \dots, n$ [28]. In [8] (see also references therein) an overview of some results on various $\star$-Stein equations is presented. In particular, the problem of finding necessary and sufficient conditions for the consistency of $FX - NX^{\star} = H$ and more generally for $FXM - NX^{\star}G = H$ are stated as being open [8], but are now solved by Theorem 6.1.

Acknowledgements. The authors thank Vladimir Sergeichuk for his suggestions that improved presentation of Section 3, and the two anonymous referees for their insightful comments.

REFERENCES

[1] J.K. Baksalary and R. Kala, *The matrix equation $AX - YB = C$*, Linear Algebra Appl., 25, (1979), pp. 41–43.

[2] M.A. Beitia and J.-M. Gracia, *Sylvester matrix equation for matrix pencils*, Linear Algebra Appl., 232, (1996), pp. 157–197.

[3] P. Benner, R.-C. Li, and N. Truhar, *On the ADI method for Sylvester equations*, J. Comput. Appl. Math., 233(4) (2009), pp. 1035–1045.

[4] R. Bhatia and P. Rosenthal, *How and why to solve the operator equation $AX - XB = Y$*, Bull. London Math. Soc., 29 (1997), pp. 1–21.

[5] A.W. Bojanczyk, G.H. Golub, and P. Van Dooren, Periodic Schur decomposition: algorithms and applications, *Proc. SPIE 1770, Advanced Signal Processing Algorithms, Architectures, and Implementations III*, 1992.

[6] R. Byers and D. Kressner, *Structured condition numbers for invariant subspaces*, SIAM J. Matrix Anal. Appl., 28(2) (2006), pp. 326–347.

[7] J. Demmel and B. Kågström, *Computing stable eigendecompositions of matrix pencils*, Linear Algebra Appl., 88 (1987), pp. 139–186.

[8] F. De Terán, *The solution of the equation $AX + BX^T = 0$*, Lin. Multilin. Algebra, 61(12) (2013), pp. 1605–1628.

[9] F. De Terán, *A note on the consistency of a system of $\star$-Sylvester equations*, arXiv: 1411.0420

[10] 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 (2011), pp. 44–67.

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

[12] F. De Terán and F.M. Dopico, *Consistency and efficient solution of the Sylvester equation for $\star$-congruence*, Electron. J. Linear Algebra, 22 (2011), pp. 849–863.

[13] A. Dmytryshyn, B. Kågström, and V. V. Sergeichuk, *Skew-symmetric matrix pencils: Codimension counts and the solution of a pair of matrix equations*, Linear Algebra Appl., 438(8) (2013), pp. 3375–3396.

[14] A. Dmytryshyn, B. Kågström, and V.V. Sergeichuk, *Symmetric matrix pencils: Codimension counts and the solution of a pair of matrix equations*, Electron. J. Linear Algebra, 27 (2014), pp. 1–18.

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

[16] R. Granat, I. Jonsson, and B. Kågström, Recursive Blocked Algorithms for Solving Periodic Triangular Sylvester-Type Matrix Equations. In B. Kågström et al., editor, *Applied Parallel Computing: State of the Art in Scientific Computing*, PARA 2006, Lecture Notes in Computer Science, LNCS 4699, pp. 531–539, Springer, 2007.

[17] R. Granat and B. Kågström, *Parallel Solvers for Sylvester-Type Matrix Equations with Applications in Condition Estimation, Part I: Theory and Algorithms*, ACM Trans. Math. Softw., 37(3) (2010), pp. 32:1–32:32.

[18] R. Granat and B. Kågström, *Algorithm 904: The SCASY Library – Parallel Solvers for Sylvester-Type Matrix Equations with Applications in Condition Estimation, Part II*, ACM Trans. Math. Softw., 37(3) (2010) pp. 33:1–33:4.

[19] R. Granat, B. Kågström, and D. Kressner, *Computing periodic deflating subspaces associated with a specified set of eigenvalues*, BIT Numerical Mathematics, 47 (2007), pp. 763–791.

[20] R. Granat and B. Kågström, *Direct eigenvalue reordering in a product of matrices in extended periodic Schur form*, SIAM J. Matrix Anal. Appl., 28(1) (2006), pp. 285–300.

[21] R.M. Guralnick, *Roth's theorems for sets of matrices*, Linear Algebra Appl., 71 (1985), pp. 113–117.

[22] W.H. Gustafson, *Roth's theorem over commutative rings*, Linear Algebra Appl., 23 (1979), pp. 245–251.

[23] R.A. Horn and V.V. Sergeichuk, Representations of quivers and mixed graphs, Chapter 34 in: L. Hogben (Ed.), *Handbook of Linear Algebra*, 2nd ed., Chapman & Hall/CRC, Boca Raton, FL, 2014.

[24] 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(4) (2002), pp. 392–415.

[25] I. Jonsson and B. Kågström, *Recursive blocked algorithms for solving triangular systems. Part II. Two-sided and generalized Sylvester and Lyapunov matrix equations*, ACM Trans. Math. Software, 28(4) (2002), pp. 416–435.

[26] B. Kågström, *A perturbation analysis of the generalized Sylvester equation $(AR - LB, DR - LE) = (C, F)$*, SIAM J. Matrix Anal. Appl., 15(4) (1994), pp. 1045–1060.

[27] B. Kågström and L. Westin, *Generalized Schur methods with condition estimators for solving the generalized Sylvester equation*, IEEE Transactions on Automatic Control, 34(7) (1989), pp. 745–751.

[28] S.-G. Lee and Q.-P. Vu, *Simultaneous solutions of matrix equations and simultaneous equivalence matrices*, Linear Algebra Appl., 437 (2012), pp. 2325–2339.

[29] W.E. Roth, *The equations $AX - YB = C$ and $AX - XB = C$ in matrices*, Proceed. of the American Math. Soc., 3(3) (1952), pp. 392–396.

[30] A.V. Roĭter, BOCSs with involution, in: *Representations and Quadratic Forms*, Akad. Nauk Ukrain. SSR, Inst. Mat., Kiev, 1979, pp. 124–126 (in Russian).

[31] V.V. Sergeichuk, *Classification problems for systems of forms and linear mappings*, Math. USSR-Izv., 31(3) (1988), pp. 481–501.

[32] V.L. Syrmos and F.L. Lewis, *Decomposability and quotient subspaces for the pencil $sL - M$*, SIAM J. Matrix Anal. Appl., 15(3) (1994), pp. 865–880.

[33] Y.S. Tong, *On the operator equation $AXB - X = C$*, Chinese Annals of Mathematics, 7 (1986), pp. 327–337.

[34] Q.-W. Wang and Z.-H. He, *Solvability conditions and general solution for mixed Sylvester equations*, Automatica, 49 (2013), pp. 2713–2719.

[35] Q.-W. Wang and Z.-H. He, *Systems of coupled generalized Sylvester matrix equations*, Automatica, 50 (2014), pp. 2840–2844.

[36] Q.-W. Wang, J.-H. Sun, and S.-Z. Li, *Consistency for bi(skew) symmetric solutions to systems of generalized Sylvester equations over a finite central algebra*, Lineal Algebra Appl., 353 (2002) 169–182.

[37] H.K. Wimmer, *The matrix equation $X - AXB = C$ and analogue of Roth's theorem*, Linear Algebra Appl., 109 (1988), pp. 145–147.

[38] H.K. Wimmer, *Consistency of a pair of generalised Sylvester equations*, IEEE Trans. on Autom. Control, 39(5) (1994), pp. 1014–1016.

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