Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

SOLVING THE SYLVESTER EQUATION $AX - XB = C$ WHEN $\sigma(A) \cap \sigma(B) \neq \emptyset$*

NEBOJŠA Č. DINČIĆ†

Abstract. The method for solving the Sylvester equation $AX - XB = C$ in the complex matrix case, when $\sigma(A) \cap \sigma(B) \neq \emptyset$, by using Jordan normal form, is given. Also, the approach via the Schur decomposition is presented.

Key words. Sylvester equation, Jordan normal form, Schur decomposition.

AMS subject classifications. 15A24.

1. An introduction.

The Sylvester equation

$$AX - XB = C, \tag{1.1}$$

where $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$ and $C \in \mathbb{C}^{m\times n}$ are given, and $X \in \mathbb{C}^{m\times n}$ is an unknown matrix, is a very popular topic in linear algebra, with numerous applications in control theory, signal processing, image restoration, engineering, solving ordinary and partial differential equations, etc. It is named for famous mathematician J. J. Sylvester, who was the first to prove that this equation in matrix case has unique solution if and only if $\sigma(A) \cap \sigma(B) = \emptyset$ [17]. One important special case of the Sylvester equation is the continuous-time Lyapunov matrix equation ($AX + XA^* = C$).

In 1952, Roth [15] proved that for operators A, B on finite-dimensional spaces $\begin{bmatrix} A & C \\ 0 & B \end{bmatrix}$ is similar to $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ if and only if the equation $AX - XB = C$ has a solution X. Rosenblum [14] proved that if A and B are operators such that $\sigma(A) \cap \sigma(B) = \emptyset$, then the equation $AX - XB = Y$ has a unique solution X for every operator Y. If we define the linear operator $\mathcal{T}$ on the space of operators by $\mathcal{T}(X) = AX - XB$, the conclusion of the Rosenblum theorem can be restated: $\mathcal{T}$ is invertible if $\sigma(A) \cap \sigma(B) = \emptyset$. Kleinecke showed that when A and B are operators on the same space and $\sigma(A) \cap \sigma(B) \neq \emptyset$, then the operator $\mathcal{T}$ is not invertible, see [14]. For the exhaustive survey on these topics, please see [2] and [18] and the references therein. In recent paper of Li and Zhou [12], an extensive review of literature can be found, with a brief classification of existing methods for solving the Sylvester equation.

Drazin [6] recently gave another equivalent condition.

THEOREM 1.1. [6] *For any field F, any $r, s \in \mathbb{N}$ and any square matrices $A \in M_r(F)$, $B \in M_s(F)$ whose eigenvalues all lie in F, the following three properties of the pair A, B are equivalent:*

i) for any given polynomials $f, g \in F[t]$, there exists $h \in F[t]$ such that $f(A) = h(A)$ and $g(B) = h(B)$;
ii) A and B share no eigenvalue in common;
iii) for $r \times s$ matrices X over F, $AX = XB \Rightarrow X = 0$.

*Received by the editors on January 10, 2018. Accepted for publication on December 2, 2018. Handling Editor: Froilán Dopico.
†Faculty of Sciences and Mathematics, University of Niš, PO Box 224, 18000 Niš, Serbia (ndincic@hotmail.com). The author was supported by the Ministry of Science, Republic of Serbia, grant no. 174007.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

We recall some results where the solution to the Sylvester equation is given in various forms.

Proposition 1.2. ([2, p. 9]) *Let A and B be operators such that $\sigma(B) \subset \{z : |z| < \rho\}$ and $\sigma(A) \subset \{z : |z| > \rho\}$ for some $\rho > 0$. Then the solution of the equation $AX - XB = Y$ is*

$$X = \sum_{n=0}^{\infty} A^{-n-1} Y B^n.$$

We will often use the following special form of the previous result.

Proposition 1.3. *In the complex matrix case, the Sylvester equation $AX - XB = C$ is consistent if and only if $\sigma(A) \cap \sigma(B) = \emptyset$ and the solution is given by*

$$X = A^{-1} \cdot \sum_{k=0}^{\infty} A^{-k} C B^k$$

for invertible A, and by

$$X = -\sum_{k=0}^{\infty} A^k C B^{-k} \cdot B^{-1}$$

for invertible matrix B.

Proposition 1.4. [9] *Let A and B be operators whose spectra are contained in the open right half plane and the open left half plane, respectively. Then the solution of the equation $AX - XB = Y$ can be expressed as*

$$X = \int_0^{\infty} e^{-tA} Y e^{tB} dt.$$

Proposition 1.5. [14] *Let Γ be a union of closed contours in the plane, with total winding numbers 1 around $\sigma(A)$ and 0 around $\sigma(B)$. Then the solution of the equation $AX - XB = Y$ can be expressed as*

$$X = \frac{1}{2\pi i} \int_{\Gamma} (A - \zeta)^{-1} Y (B - \zeta)^{-1} d\zeta.$$

Solving the Sylvester equation in the case when $\sigma(A) \cap \sigma(B) \neq \emptyset$ is rather complicated and is not thoroughly studied yet. According to [3], only Ma [13], and Datta and Datta [4] treated the particular cases when $B(= A)$ has only simple eigenvalues or $C = 0$ and $B(= A^T)$ are unreduced Hessenberg matrices, respectively. In the paper [3], two special cases, when both A and B are symmetric or skew-symmetric matrices, were considered, by using the notion of the eigenprojection.

In [11], it is shown that in the case $\sigma(A) \cap \sigma(B) = \emptyset$ ($A \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{n\times n}$), the solution of the Sylvester equation is $X = q(A)^{-1}\eta(A, C, B)$, which is a polynomial in the matrices A, B and C. Also, it is shown that when $\sigma(A) \cap \sigma(B) \neq \emptyset$, the solution set is contained in that of the linear algebraic equation $q(A)X = \eta(A, C, B)$. Recall that $q(s) = s^n + \beta_{n-1}s^{n-1} + \cdots + \beta_1 s + \beta_0$ is the characteristic polynomial of B and $\eta(A, C, B) = \sum_{k=1}^{n} \beta_k \eta(k - 1, A, C, B)$, where $\eta(k, A, C, B) = \sum_{i=0}^{k} A^{k-i} C B^i$.

In [5], the equation $AX - X^\star B = 0$ is considered by using the Kronecker canonical form of the matrix pencil $A + \lambda B^\star$, where $X^\star$ denotes either transpose or the conjugate transpose of X.

Recently, Li and Zhou [12] used the method based on spectral decompositions of the matrices A and B. Note that this method also works in the case when spectra of A and B are not disjoint.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

In the paper [18], homogeneous and nonhomogeneous Sylvester-conjugate matrix equation of the forms $A\overline{X} + BY = XF$ and $A\overline{X} + BY = XF + R$ are considered.

In this paper, we are dealing with the Sylvester equation in complex matrix case, when $\sigma(A) \cap \sigma(B) \neq \emptyset$. Because the operator $\mathcal{T}(X) = AX - XB$ is not invertible in this case, as Kleinecke showed, we must find the consistency condition, and general solution in the case of consistency. We will use the Jordan normal form for the matrices A and B (similar method as Gantmacher used in [7, Ch. VIII], but for homogeneous case only!), so the Sylvester equation will become the set of the simpler equations of the form $J_m(\lambda)Z - ZJ_n(\mu) = W$, where $W \in \mathbb{C}^{m\times n}$ is given matrix, $Z \in \mathbb{C}^{m\times n}$ is unknown matrix, and λ, μ are complex numbers. In the paper, we are dealing mainly with such simpler equations, discussing the consistency condition and describing the algorithm for finding the general solution. Finally, we give the result characterizing the most general case, as well as the approach via Schur decomposition.

2. Reducing the problem to the most elementary equations. Let the matrices $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be given, and suppose that $W = \sigma(A) \cap \sigma(B) = \{\lambda_1, \ldots, \lambda_s\} \neq \emptyset$. It is a well-known fact (see e.g., [10]) that any square complex matrix can be reduced to the Jordan canonical form. Therefore, there exist regular matrices $S \in \mathbb{C}^{m\times m}$ and $T \in \mathbb{C}^{n\times n}$ such that $A = SJ_AS^{-1}$ and $B = TJ_BT^{-1}$, or, more precisely,

$$A = S \begin{bmatrix} J_{A_1} & 0 \\ 0 & J_{A_2} \end{bmatrix} S^{-1}, \quad B = T \begin{bmatrix} J_{B_1} & 0 \\ 0 & J_{B_2} \end{bmatrix} T^{-1}, \tag{2.2}$$

where J_{A_1} (respectively, J_{B_1}) consists of just those Jordan matrices corresponding to the eigenvalues from the $\sigma(A) \setminus W$ (respectively, $\sigma(B) \setminus W$), and J_{A_2} and J_{B_2} are those Jordan matrices corresponding to the eigenvalues from the set W:

$$\begin{aligned} J_{A_2} &= diag\{J(\lambda_1; p_{11}, p_{12}, \ldots, p_{1,k_1}), \ldots, J(\lambda_s; p_{s1}, p_{s2}, \ldots, p_{s,k_s})\}, \\ J_{B_2} &= diag\{J(\lambda_1; q_{11}, q_{12}, \ldots, q_{1,\ell_1}), \ldots, J(\lambda_s; q_{s1}, q_{s2}, \ldots, q_{s,\ell_s})\}. \end{aligned}$$

Here $p_{ij},\ j = \overline{1, k_i},\ i = \overline{1, s}$, and $q_{ij},\ j = \overline{1, \ell_i},\ i = \overline{1, s}$, are natural numbers, and k_i and ℓ_i are geometric multiplicities of the eigenvalue $\lambda_i,\ i = \overline{1, s}$, of A and B, respectively; their algebraic multiplicities are given by $m_i = p_{i1} + p_{i2} + \cdots + p_{i,k_i}$ and $n_i = q_{i1} + q_{i2} + \cdots + q_{i,\ell_i}$, where $i = \overline{1, s}$. We will use the notation

$$J(\lambda; t_1, \ldots, t_k) = diag\{J_{t_1}(\lambda), \ldots, J_{t_k}(\lambda)\} = J_{t_1}(\lambda) \oplus \cdots \oplus J_{t_k}(\lambda),$$

where $J_{t_i}(\lambda),\ i = \overline{1, k}$, is the Jordan block

$$J_{t_i}(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}_{t_i\times t_i}.$$

Recall that a matrix A is called *non-derogatory* if every eigenvalue of A has geometric multiplicity 1, if and only if corresponding to each distinct eigenvalue is exactly one Jordan block [10, p. 135].

If we substitute (2.2) in (1.1), and denote $S^{-1}XT = Y$ and $S^{-1}CT = D$, the equation becomes

$$\begin{bmatrix} J_{A_1} & 0 \\ 0 & J_{A_2} \end{bmatrix} Y - Y \begin{bmatrix} J_{B_1} & 0 \\ 0 & J_{B_2} \end{bmatrix} = D. \tag{2.3}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

We will partition the matrices $Y = [Y_{ij}]_{2\times 2}$ and $D = [D_{ij}]_{2\times 2}$ in accordance with the partition of J_A and J_B. Now (2.3) becomes:

$$\begin{bmatrix} J_{A_1} & 0 \\ 0 & J_{A_2} \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} - \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} J_{B_1} & 0 \\ 0 & J_{B_2} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix},$$

which reduces to the following four simpler Sylvester equations:

$$\begin{aligned} J_{A_1}Y_{11} - Y_{11}J_{B_1} &= D_{11}, \\ J_{A_1}Y_{12} - Y_{12}J_{B_2} &= D_{12}, \\ J_{A_2}Y_{21} - Y_{21}J_{B_1} &= D_{21}, \\ J_{A_2}Y_{22} - Y_{22}J_{B_2} &= D_{22}. \end{aligned}$$

By Proposition 1.3, the first three equations have the unique solutions, because $\sigma(A_1)\cap\sigma(B_1) = \emptyset$, $\sigma(A_1)\cap\sigma(B_2) = \emptyset$, $\sigma(A_2)\cap\sigma(B_1) = \emptyset$, given by the appropriate power series. The fourth equation does not have a unique solution (when it is consistent), because $\sigma(A_2)\cap\sigma(B_2) = W \neq \emptyset$, so we turn our attention to this case.

Therefore, throughout the paper, we can, without loss of generality, consider the Sylvester equation in the form $AX - XB = C$, where A and B are square complex matrices (in general of different size) already reduced to their Jordan forms and such that $\sigma(A) = \sigma(B)$.

If we partition $X = [X_{ij}]_{s\times s}$ and $C = [C_{ij}]_{s\times s}$ in accordance with already given decompositions for A and B, and put them in the equation $AX - XB = C$, we have s^2 simpler Sylvester equations of the form:

$$J(\lambda_i; p_{i1}, \dots, p_{i,k_i})X_{ij} - X_{ij}J(\lambda_j; q_{j1}, \dots, q_{j,\ell_j}) = C_{ij}, \quad i, j = \overline{1, s}. \tag{2.4}$$

For $i \neq j$ we have $\sigma(J(\lambda_i; p_{i1}, \dots, p_{i,k_i})) = \{\lambda_i\} \neq \{\lambda_j\} = \sigma(J(\lambda_j; q_{j1}, \dots, q_{j,\ell_j}))$, and by Proposition 1.3, this case also has the unique solution. Therefore, among s^2 equations in (2.4) there are $s^2 - s$ of them which are uniquely solvable, and remaining s equations (for $i = j$) after appropriate translation reduce to $\sum_{i=1}^{s}(k_i \cdot \ell_i)$ simple Sylvester equations of the form

$$J_{p_{i,u}}(0)X_{u,v}^{(ii)} - X_{u,v}^{(ii)}J_{q_{i,v}}(0) = C_{u,v}^{(ii)}, \ \ u = \overline{1, k_i}, \ v = \overline{1, \ell_i}, \ i = \overline{1, s}.$$

Because of that, we will investigate this important particular case in detail.

2.1. Solving the equation $J_m(0)X - XJ_n(0) = C$. From the previous consideration, we see that Jordan matrices play important role in the paper. Recall that multiplying some matrix $C \in \mathbb{C}^{m\times n}$ by the Jordan block (or transposed Jordan block) corresponding to 0 acts like shifting the rows (or columns) and inserting a row (or a column) of all zeros where it is necessary: $J_m(0)C$ shifts up, $CJ_n(0)$ shifts to the right, $J_m^T(0)C$ shifts down, while $CJ_n^T(0)$ shifts to the left. More precisely, the following Lemma holds.

LEMMA 2.1. *For given matrix $C = [c_{ij}] \in \mathbb{C}^{m\times n}$ we have:*

i) $[J_m(0)C]_{ij} = \begin{cases} c_{i+1,j}, & i = \overline{1, m-1}, \\ 0, & i = m. \end{cases}$

ii) $[CJ_n(0)]_{ij} = \begin{cases} c_{i,j-1}, & j = \overline{2, n}, \\ 0, & j = 1. \end{cases}$

iii) $[J_m(0)^TC]_{ij} = \begin{cases} c_{i-1,j}, & i = \overline{2, m}, \\ 0, & i = 1. \end{cases}$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

iv) $[CJ_n(0)^T]_{ij} = \begin{cases} c_{i,j+1}, & j = \overline{1, n-1}, \\ 0, & j = n. \end{cases}$

Moreover, for $\alpha = \overline{0, m-1}$ *and* $\beta = \overline{0, n-1}$,

v) $[J_m(0)^\alpha C J_n(0)^\beta]_{ij} = \begin{cases} c_{i+\alpha, j-\beta}, & i = \overline{1, m-\alpha} \text{ and } j = \overline{\beta+1, n}, \\ 0, & i = \overline{m-\alpha+1, m} \text{ or } j = \overline{1, \beta}. \end{cases}$

vi) $[(J_m(0)^T)^\alpha C J_n(0)^\beta]_{ij} = \begin{cases} c_{i-\alpha, j-\beta}, & i = \overline{\alpha+1, m} \text{ and } j = \overline{\beta+1, n}, \\ 0, & i = \overline{1, \alpha} \text{ or } j = \overline{1, \beta}. \end{cases}$

Example 2.2. We consider the case $A = J_4(0)$, $B = J_3(0)$, i.e., the equation

$$J_4(0)X - XJ_3(0) = C. \tag{2.5}$$

Because of $\sigma(A) = \sigma(B) = \{0\}$ we expect no unique solution. Let us see for which C there are solutions, and then let us characterize all of them. The matrix equation (2.5) (with $C = [c_{ij}]$, $X = [x_{ij}] \in \mathbb{C}^{4\times 3}$) gives us the following linear system:

$$\begin{aligned} x_{21} &= c_{11}, & x_{22} - x_{11} &= c_{12}, & x_{23} - x_{12} &= c_{13}, \\ x_{31} &= c_{21}, & x_{32} - x_{21} &= c_{22}, & x_{33} - x_{22} &= c_{23}, \\ x_{41} &= c_{31}, & x_{42} - x_{31} &= c_{32}, & x_{43} - x_{32} &= c_{33}, \\ 0 &= c_{41}, & -x_{41} &= c_{42}, & -x_{42} &= c_{43}. \end{aligned}$$

We see that some conditions must be imposed to the entries of the matrix C: $c_{41} = 0$, $c_{31} + c_{42} = 0$, $c_{21} + c_{32} + c_{43} = 0$. Therefore, any matrix C for which the equation (2.5) is consistent is of the form:

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \\ 0 & -c_{31} & -c_{21} - c_{32} \end{bmatrix}.$$

The solutions of (2.5) are

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ c_{11} & x_{11} + c_{12} & x_{12} + c_{13} \\ c_{21} & c_{11} + c_{22} & x_{11} + c_{12} + c_{23} \\ c_{31} & c_{21} + c_{32} & c_{11} + c_{22} + c_{33} \end{bmatrix},$$

where x_{11}, x_{12}, x_{13} are arbitrary complex numbers. We can rewrite this general solution X in the following more informative form:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ 0 & x_{11} & x_{12} \\ 0 & 0 & x_{11} \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & c_{11} & c_{12} \\ 0 & c_{21} & c_{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c_{11} \end{bmatrix}.$$

By Lemma 2.1, we can express this as:

$$X = \left[\frac{p_2(J_3(0))}{0_{1\times 3}} \right] + J_4(0)^T \sum_{k=0}^{2} \left(J_4(0)^T\right)^k C J_3(0)^k,$$

where $p_2(J_3(0)) = x_{11} I_3 + x_{12} J_3(0) + x_{13} J_3(0)^2$ is arbitrary polynomial of the matrix $J_3(0)$, with complex coefficients in general, and the degree at most 2.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

This example motivates us to prove general result concerning this topic.

THEOREM 2.3. *The Sylvester equation*

$$J_m(0)X - XJ_n(0) = C, \tag{2.6}$$

when $m \geq n$, is consistent if and only if

$$\sum_{k=0}^{n-1} J_m(0)^{m-1-k} C J_n(0)^k = 0, \tag{2.7}$$

or, equivalently,

$$\sum_{k=0}^{p-1} c_{m-k,p-k} = 0, \quad p = \overline{1,n}, \tag{2.8}$$

and its general solution (which depends on n complex parameters) is given as

$$X = \left[\frac{p_{n-1}(J_n(0))}{0_{(m-n)\times n}}\right] + J_m(0)^T \sum_{k=0}^{n-1} \left(J_m(0)^T\right)^k C J_n(0)^k, \tag{2.9}$$

where p_{n-1} is an arbitrary polynomial of the degree at most $n-1$.

Proof. First, we show that the conditions (2.7) and (2.8) are equivalent. Let us denote $S = \sum_{k=0}^{n-1} S_k$ where $S_k = J_m(0)^{m-1-k} C J_n(0)^k$, $k = \overline{0, n-1}$. By Lemma 2.1.v), for $\alpha = m-k-1$ and $\beta = k$, we have

$$[S_k]_{ij} = \begin{cases} c_{i+m-1-k,j-k}, & i = \overline{1,k+1} \text{ and } j = \overline{k+1,n}, \\ 0, & i = \overline{k+2,m} \text{ or } j = \overline{1,k}. \end{cases}$$

It is easy to see that $[S_k]_{ij} = 0$ for $i > j$ and $[S_k]_{ij} = [S_{k+1}]_{i+1,j+1}$; therefore, it is enough to consider only the first row (or $n-$th column). We have

$$[S_k]_{1j} = \begin{cases} c_{m-k,j-k}, & j = \overline{k+1,n}, \\ 0, & j = \overline{1,k}. \end{cases}$$

Therefore,

$$[S]_{1j} = \sum_{k=0}^{n-1} [S_k]_{1j} = \sum_{k=0}^{j-1} c_{m-k,j-k}, \quad k = \overline{1,n}.$$

Note that the summation index is changed to $k = \overline{0, j-1}$ because of $k+1 \leq j \leq n$ and $0 \leq k \leq n-1$. Hence, we proved (2.7) $\Leftrightarrow$ (2.8).

($\Rightarrow$) : Suppose that the equation (2.6) is consistent, which means for some $\widehat{C}$ there is $\widehat{X}$ such that $J_m(0)\widehat{X} - \widehat{X}J_n(0) = \widehat{C}$. We have:

$$\begin{aligned}
&\sum_{k=0}^{n-1} J_m(0)^{m-1-k}\widehat{C}J_n(0)^k = \sum_{k=0}^{n-1} J_m(0)^{m-1-k}(J_m(0)\widehat{X} - \widehat{X}J_n(0))J_n(0)^k \\
&= \sum_{k=0}^{n-1} J_m(0)^{m-k}\widehat{X}J_n(0)^k - \sum_{k=0}^{n-1} J_m(0)^{m-1-k}\widehat{X}J_n(0)^{k+1} \\
&= J_m(0)^m\widehat{X} + J_m(0)^{m-1}\widehat{X}J_n(0) + \cdots + J_m(0)^{m-n+1}\widehat{X}J_n(0)^{n-1} \\
&\quad - J_m(0)^{m-1}\widehat{X}J_n(0) - \cdots - J_m(0)^{m-n+1}\widehat{X}J_n(0)^{n-1} - J_m(0)^{m-n}\widehat{X}J_n(0)^n \\
&= J_m(0)^m\widehat{X} - J_m(0)^{m-n}\widehat{X}J_n(0)^n = 0,
\end{aligned}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

so the consistency condition (2.7) holds.

$(\Leftarrow)$: Now we prove, by immediate checking, that X given by (2.9) is indeed the solution under the consistency condition (2.7). Because of

$$\begin{aligned}
&J_m(0)\left[\frac{p_{n-1}(J_n(0))}{0_{(m-n)\times n}}\right] - \left[\frac{p_{n-1}(J_n(0))}{0_{(m-n)\times n}}\right] J_n(0)\\
&= \begin{bmatrix} J_n(0) & * \\ 0 & J_{m-n}(0) \end{bmatrix} \left[\frac{p_{n-1}(J_n(0))}{0_{(m-n)\times n}}\right] - \left[\frac{p_{n-1}(J_n(0))}{0_{(m-n)\times n}}\right] J_n(0)\\
&= \left[\frac{J_n(0)\, p_{n-1}(J_n(0))}{0_{(m-n)\times n}}\right] - \left[\frac{p_{n-1}(J_n(0))\, J_n(0)}{0_{(m-n)\times n}}\right] = 0_{m\times n},
\end{aligned}$$

(by "$*$" we denoted some submatrix which is not important now) we have:

$$\begin{aligned}
&J_m(0)X - XJ_n(0) - C = J_m(0)J_m(0)^T \sum_{k=0}^{n-1} \left(J_m(0)^T\right)^k CJ_n(0)^k\\
&\quad - J_m(0)^T \sum_{k=0}^{n-1} \left(J_m(0)^T\right)^k CJ_n(0)^k J_n(0) - C\\
&= J_m(0)J_m(0)^T C + \left(J_m(0)(J_m(0)^T)^2 - J_m(0)^T\right) CJ_n(0)\\
&\quad + \cdots + \left(J_m(0)(J_m(0)^T)^n - (J_m(0)^T)^{n-1}\right) CJ_n^{n-1}(0)\\
&= \left(J_m(0)J_m(0)^T - I_m\right) \left(\sum_{k=0}^{n-1} (J_m(0)^T)^k CJ_n(0)^k\right) + C - C = 0.
\end{aligned}$$

Indeed, let us denote $T_k = (J_m(0)^T)^k CJ_n(0)^k,\ k = \overline{0, n-1}$, and $T = \sum_{k=0}^{n-1} T_k$.

By Lemma 2.1.vi), for $\alpha = \beta = k$, we have:

$$[T_k]_{ij} = \begin{cases} c_{i-k,j-k}, & i = \overline{k+1, m} \text{ and } j = \overline{k+1, n},\\ 0, & i = \overline{1,k} \text{ or } j = \overline{1,k}. \end{cases}$$

If we obtain that all entries in the $m-$th row of matrix T are zero, we have the proof. We proceed:

$$[T_k]_{mj} = [(J_m(0)^T)^k CJ_n(0)^k]_{mj} = \begin{cases} c_{m-k,j-k}, & j = \overline{k+1, n},\\ 0, & j = \overline{1,k}, \end{cases}$$

and therefore,

$$[T]_{mj} = \sum_{k=0}^{n-1} [T_k]_{mj} = \sum_{k=0}^{j-1} c_{m-k,j-k}, \quad j = \overline{1,n}.$$

Note that again the summation index is changed to $k = \overline{0, j-1}$ because of $k+1 \leq j \leq n$ and $0 \leq k \leq n-1$.

The condition (2.7), or equivalently (2.8), ensures that $[T]_{mj} = 0,\ j = \overline{1,n}$, so X given by (2.9) is, under the condition (2.7), or equivalently (2.8), indeed the solution, so the equation (2.6) is consistent. ∎

REMARK 2.4. For the given matrix $C \in \mathbb{C}^{m\times n},\ m \geq n$, the set of its entries c_{ij} such that $i - j = m - p$ for given $p = \overline{1,n}$ will be called $p-$th small subdiagonal of matrix C. The condition (2.8) actually means that

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

the sum over each of n small subdiagonals must be zero, which is easy to check. However, for constructing the solutions of the more general Sylvester equations from the most simple ones, the condition (2.7) is more appropriate.

We remark that the particular solution in (2.9),

$$X_p = J_m(0)^T \sum_{k=0}^{n-1} \left(J_m(0)^T\right)^k C J_n(0)^k,$$

is completely determined by the matrix C, more precisely, by the "independent" part of C (upper $(m-1)\times n$ submatrix), and it can be expressed in equivalent, but for computational purposes more practical form as:

$$[X_p]_{ij} = \begin{cases} 0, & i=1, \\ \sum\limits_{k=0}^{\min\{i-1,j\}-1} c_{i-1-k,j-k}, & i=\overline{2,m}. \end{cases}$$

Let us step away a bit, and consider solving the equation (2.6) by the well-known Kronecker product method. We recall the vectorization operator vec given by

$$vec : \mathbb{C}^{m\times n} \to \mathbb{C}^{mn}, \;\; vec(A) = vec([a_{ij}]) = [a_{\bullet 1}\; a_{\bullet 2}\; \cdots\; a_{\bullet n}]^T,$$

where $a_{\bullet k}$ denotes the $k-$th column of the matrix A. Also, recall that the Kronecker product of $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$ is $mp \times nq$ block matrix

$$A \otimes B = [a_{ij}B]_{m\times n}.$$

If we solve the equation $J_m(0)X - XJ_n(0) = C,\; m \geq n$, by using the well-known method with the operator vec and the Kronecker product, we have

$$(I_n \otimes J_m(0) - J_n(0)^T \otimes I_m)\, vec(X) = vec(C); \tag{2.10}$$

this linear system is consistent if and only if

$$r([I_n \otimes J_m(0) - J_n(0)^T \otimes I_m,\; C]) = r(I_n \otimes J_m(0) - J_n(0)^T \otimes I_m).$$

The system (2.10) in matrix form (zero submatrices are omitted) is:

$$\begin{bmatrix} J_m(0) & & & & & \\ -I_m & J_m(0) & & & & \\ & -I_m & & & & \\ & & \ddots & \ddots & & \\ & & & -I_m & J_m(0) & \\ & & & & -I_m & J_m(0) \end{bmatrix} \begin{bmatrix} X_{\bullet 1} \\ X_{\bullet 2} \\ \vdots \\ X_{\bullet n} \end{bmatrix} = \begin{bmatrix} C_{\bullet 1} \\ C_{\bullet 2} \\ \vdots \\ C_{\bullet n} \end{bmatrix},$$

i.e.,

$$J_m(0)X_{\bullet 1} = C_{\bullet 1},\; J_m(0)X_{\bullet 2} = C_{\bullet 2} + X_{\bullet 1},\; \ldots,\; J_m(0)X_{\bullet n} = C_{\bullet n} + X_{\bullet (n-1)}.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

We can infer the consistency condition in the following way:

$$\begin{aligned} 0 &= J_m(0)^m X_{\bullet n} = J_m(0)^{m-1}(J_m(0) X_{\bullet n}) = J_m(0)^{m-1}(C_{\bullet n} + X_{\bullet (n-1)}) \\ &= \cdots = \sum_{k=1}^{n} J_m(0)^{m-n-1+k} C_{\bullet k}. \end{aligned}$$

We shall not pursue this approach further, because the obtained matrix of the system is an $mn \times mn$ sparse matrix, and for the consistency of the equation we need to check rank conditions and to use some generalized inverses.

For Example 2.2, where $m = 4,\ n = 3$, the previous formula gives:

$$0 = \sum_{k=1}^{3} J_4(0)^k C_{\bullet k} = \begin{bmatrix} c_{21} \\ c_{31} \\ c_{41} \\ 0 \end{bmatrix} + \begin{bmatrix} c_{32} \\ c_{42} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} c_{43} \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

which is in accordance with the conclusion from Example 2.2.

We return to the main flow of the paper.

THEOREM 2.5. *The Sylvester equation*

$$J_m(0)X - XJ_n(0) = C, \tag{2.11}$$

when $m \leq n$, is consistent if and only if

$$\sum_{k=0}^{m-1} J_m(0)^k C J_n(0)^{n-1-k} = 0, \tag{2.12}$$

and its general solution is

$$X = \begin{bmatrix} 0_{m\times(n-m)} & q_{m-1}(J_m(0)) \end{bmatrix} - \sum_{k=0}^{m-1} J_m(0)^k C \left(J_n(0)^T\right)^k J_n(0)^T, \tag{2.13}$$

where q_{m-1} is an arbitrary polynomial of the degree at most $m-1$.

Proof. If we take transpose of the original equation (2.11), and then multiply it by -1, we have

$$J_n(0)^T X^T - X^T J_m(0)^T = -C^T,$$

which is similar, but not the same as the equation (2.6), so we cannot directly apply Theorem 2.3. But if we notice that for any $k \in \mathbb{N}$

$$J_k(0)^T = P_k^{-1} J_k(0) P_k, \ \text{ where } P_k = P_k^{-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}_{k\times k}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

is so-called exchange matrix, the equation becomes

$$\begin{aligned}&P_nJ_n(0)^TX^TP_m-P_nX^TJ_m(0)^TP_m=-P_nC^TP_m,\\&\Leftrightarrow P_nJ_n(0)^TP_nP_nX^TP_m-P_nX^TP_mP_mJ_m(0)^TP_m=-P_nC^TP_m,\\&\Leftrightarrow J_n(0)Y-YJ_m(0)=D,\end{aligned}$$

where $P_nX^TP_m=Y$ and $-P_nC^TP_m=D$. Only now we can apply Theorem 2.3 to the equation $J_n(0)Y-YJ_m(0)=D$. We have the consistency condition (2.12) because of

$$\begin{aligned}0=\sum_{k=0}^{m-1}J_n(0)^{n-k-1}DJ_m(0)^k&=-\sum_{k=0}^{m-1}J_n(0)^{n-1-k}P_nC^TP_mJ_m(0)^k\\&=-P_n\left(\sum_{k=0}^{m-1}(P_nJ_n(0)P_n)^{n-1-k}C^T(P_mJ_m(0)P_m)^k\right)P_m\\&=-P_n\left(\sum_{k=0}^{m-1}(J_n(0)^T)^{n-1-k}C^T(J_m(0)^T)^k\right)P_m\\&=-P_n\left(\sum_{k=0}^{m-1}J_m(0)^kCJ_n(0)^{n-k-1}\right)^TP_m.\end{aligned}$$

General solution of $J_n(0)Y-YJ_m(0)=D$ is, by Theorem 2.3:

$$Y=\left[\frac{p_{m-1}(J_m(0))}{0_{(n-m)\times m}}\right]+\sum_{k=0}^{n-1}\left(J_n(0)^T\right)^{k+1}DJ_m(0)^k.$$

By the substitutions $X=(P_nYP_m)^T,\ D=-P_nC^TP_m$, we have the solution (2.13):

$$\begin{aligned}X&=P_m[p_{m-1}(J_m(0)^T)\quad 0_{m\times(n-m)}]P_n+P_m\sum_{k=0}^{m-1}(J_m(0)^T)^kD^TJ_n(0)^{k+1}P_n\\&=P_m[p_{m-1}(J_m(0)^T)\quad 0_{m\times(n-m)}]\begin{bmatrix}0&P_m\\P_{n-m}&0\end{bmatrix}-\sum_{k=0}^{m-1}P_m(J_m(0)^T)^k(P_mCP_n)J_n(0)^{k+1}P_n\\&=P_m[0_{m\times(n-m)}\quad p_{m-1}(J_m(0)^T)]P_m-\sum_{k=0}^{m-1}(P_mJ_m(0)^TP_m)^kC(P_nJ_n(0)P_n)^{k+1}\\&=[0_{m\times(n-m)}\ p_{m-1}(J_m(0))]-\sum_{k=0}^{m-1}J_m(0)^kC(J_n(0)^T)^{k+1}.\qquad\square\end{aligned}$$

The condition (2.12) actually means that the sum of elements over each of m small subdiagonals ($c_{ij},\ i-j=p-1$, where $p=\overline{1,m}$) must be zero. We also remark that the particular solution in (2.13),

$$X_p=-\sum_{k=0}^{m-1}J_m(0)^kC\left(J_n(0)^T\right)^kJ_n(0)^T,$$

is completely determined by the matrix C, more precisely, by the "independent" part of C (rightmost $m\times(n-1)$ submatrix), and it can be expressed in equivalent, but for computational purposes, a more

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

practical form as:

$$[X_p]_{ij} = \begin{cases} -\sum\limits_{k=0}^{\min\{m-i,n-1-j\}} c_{i+k,j+1+k}, & j = \overline{1, n-1}, \\ 0, & j = n. \end{cases}$$

Corollary 2.6. *The homogeneous Sylvester equation*

$$J_m(0)X = XJ_n(0)$$

is consistent and its general solution is

$$X = \begin{cases} \left[\dfrac{p_{n-1}(J_n(0))}{0_{(m-n)\times n}} \right], & m \geq n, \\ \left[\begin{array}{cc} 0_{m\times(n-m)} & p_{m-1}(J_m(0)) \end{array} \right], & m \leq n. \end{cases}$$

where p_{n-1} is an arbitrary polynomial of the degree at most $n-1$. The rank of the solution X is given by

$$r(X) = r(p_{s-1}(J_s(0))) = s - \min_{k=\overline{0,s-1}} \{k : p_{s-1}^{(k)}(0) \neq 0\}, \quad s = \min\{m, n\}.$$

Corollary 2.7. *All matrices commuting with $J_m(0)$ are given by $X = p_{m-1}(J_m(0))$. Any polynomial such that $p_{m-1}(0) \neq 0$ gives a nonsingular X. Also*

$$\sigma(X) = \{p_{m-1}(0)\}.$$

From Theorem 2.3 and Theorem 2.5 in the case $m = n$, we have the following result.

Corollary 2.8. *The Sylvester equation*

$$J_n(0)X - XJ_n(0) = C,$$

is consistent if and only if

$$\sum_{k=0}^{n-1} J_n(0)^k C J_n(0)^{n-1-k} = \sum_{k=0}^{n-1} J_n(0)^{n-1-k} C J_n(0)^k = 0,$$

and its general solution is

$$\begin{aligned} X &= p_{n-1}(J_n(0)) - \sum_{k=0}^{n-2} J_n(0)^k C \left(J_n(0)^T\right)^{k+1} \\ &= q_{n-1}(J_n(0)) + \sum_{k=0}^{n-2} \left(J_n(0)^T\right)^{k+1} C J_n(0)^k, \end{aligned}$$

where p_{n-1} and q_{n-1} are arbitrary polynomials of the degree at most $n-1$.

2.2. Solving the equation $J_m(\lambda)X - XJ_n(\mu) = C$ for $\lambda, \mu \in \mathbb{C}$. If we observe that $J_k(\lambda) = \lambda I_k + J_k(0),\ k \in \mathbb{N}$, the equation

$$J_m(\lambda)X - XJ_n(\mu) = C \tag{2.14}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

becomes

$$(\lambda-\mu)X + J_m(0)X - XJ_n(0) = C. \tag{2.15}$$

Case $\lambda=\mu$*:* Reduces to already solved equation (2.6) or (2.11).

Case $\lambda\neq\mu$*:* Suppose, without loss of generality, that $\lambda\neq 0$. By Proposition 1.3, we have the solution

$$X = J_m(\lambda)^{-1}\sum_{k=0}^{\infty} J_m(\lambda)^{-k} C J_n(\mu)^k.$$

We will try to avoid this infinite sum, as well as finding all $k-$th powers of the inverse of a Jordan matrix.

If we rewrite (2.15) as $(J_m(0)+(\lambda-\mu)I_m)\,X - XJ_n(0) = C$, i.e.,

$$J_m(\lambda-\mu)X - XJ_n(0) = C,$$

$J_m(\lambda-\mu)$ is nonsingular, so we have the unique solution by Proposition 1.3.

Theorem 2.9. *The Sylvester equation* (2.14) *for* $\lambda\neq\mu$ *has the unique solution*

$$X = \sum_{k=0}^{n-1} (J_m(\lambda-\mu))^{-(k+1)} C J_n(0)^k = -\sum_{k=0}^{m-1} J_m(0)^k C (J_n(\mu-\lambda))^{-(k+1)}.$$

Remark 2.10. Finding the inverses of the powers of nonsingular Jordan blocks appearing in previous theorem is not hard. Indeed, by using some matrix functional calculus, it is easy to prove that

$$J_m(\alpha)^{-k} = \sum_{s=0}^{m-1}(-1)^s \binom{k+s-1}{s}\alpha^{-(k+s)} J_m(0)^s, \quad k\in\mathbb{N},\ \alpha\neq 0.$$

3. The case when $\sigma(A)=\sigma(B)=\{0\}$. In this section, we solve the Sylvester equation $AX-XB=C$ in the case $A=diag[J_{m_1}(0), J_{m_2}(0),\dots,J_{m_p}(0)]$ and $B=diag[J_{n_1}(0),J_{n_2}(0),\dots,J_{n_q}(0)]$, where $m_1\geq m_2\geq \cdots\geq m_p>0$ and $n_1\geq n_2\geq\cdots\geq n_q>0$ and $m_1+\cdots+m_p=m$, $n_1+\cdots+n_q=n$.

Lemma 3.1. *Let* $A=J_{m_1}(0)\oplus\cdots\oplus J_{m_p}(0),\ m_1\geq\cdots\geq m_p>0,$ *and define*

$$A^{\langle k\rangle} := J_{m_1}(0)^{m_1-1-k}\oplus\cdots\oplus J_{m_p}(0)^{m_p-1-k}, \quad k=\overline{-1,m_p-1}.$$

We have:

i) $A^{\langle -1\rangle}=0$,
ii) $A^{\langle k+1\rangle}A = AA^{\langle k+1\rangle} = A^{\langle k\rangle},\ k=\overline{-1,m_p-2}$,
iii) $A^{\langle k\rangle}=I \Leftrightarrow (\forall i=\overline{1,p})\ m_i=k+1$,
iv) $(I\pm A^{\langle k\rangle})^{-1} = I\mp A^{\langle k\rangle},\ k=\overline{-1,\lfloor m_p/2\rfloor-1}$.

The proof follows by the definition of $A^{\langle k\rangle}$. Note that from (ii) it particularly follows $A^{\langle 0\rangle}A=AA^{\langle 0\rangle}=0$ and $A^{\langle 1\rangle}A=AA^{\langle 1\rangle}=A^{\langle 0\rangle}$.

Remark 3.2. Recall that $J_1(0)^0=I_1$, because of $0^0=1$.

In the sequel, $\mathbb{N}_m$ denotes the set $\{1,2,\dots,m\}$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

Definition 3.3. Let M and N be two disjoint subsets of $\mathbb{N}_m \times \mathbb{N}_n$ such that $M \cup N = \mathbb{N}_m \times \mathbb{N}_n$. The matrix mask given by the set M is mapping

$$\Pi_M : \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}, \;\; \Pi_M(A) = A_M,$$

which maps any matrix $A = [a_{ij}] \in \mathbb{C}^{m\times n}$ to the matrix A_M whose entries are given by

$$[A_M]_{ij} = \begin{cases} a_{ij}, & (i,j) \in M, \\ 0, & (i,j) \in N. \end{cases}$$

If $A \in \mathbb{C}^{m\times n}$ is partitioned on submatrices as $A = [A_{ij}]_{p\times q}$, and the set $M \subset \mathbb{N}_p \times \mathbb{N}_q$ is given (therefore, N is also known), then $\Pi_M(A) = A_M$, where

$$[A_M]_{ij} = \begin{cases} A_{ij}, & (i,j) \in M, \\ 0, & (i,j) \in N. \end{cases}$$

Note that the mapping Π_M is linear and idempotent. Also, when M is known, it means we know N too, because N is the complement of M. The binary relation described by the set M will frequently be represented by a matrix also denoted by M, with entries $M = [m_{ij}]$ given as $m_{ij} = 1$ if $(i,j) \in M$, $m_{ij} = 0$ if $(i,j) \in N$. Note that M by construction has upper quasitriangular structure.

The matrix mask method is the subject of another my paper which is still not finished, so we will not pursue the topic here more than it is necessary.

We will be interested in the case when $A = diag[J_{m_1}(0), J_{m_2}(0), \ldots, J_{m_p}(0)]$ and $B = diag[J_{n_1}(0), J_{n_2}(0), \ldots, J_{n_q}(0)]$ are given and then the sets M and N depend only on the sizes of appropriate Jordan blocks:

$$M = \{(i,j) \in \mathbb{N}_p \times \mathbb{N}_q : m_i \geq n_j\}, \;\; N = \{(i,j) \in \mathbb{N}_p \times \mathbb{N}_q : m_i < n_j\}. \tag{3.16}$$

Lemma 3.4. *Suppose that $A = diag[J_{m_1}(0), J_{m_2}(0), \cdots, J_{m_p}(0)] \in \mathbb{C}^{m\times m}$ and $B = diag[J_{n_1}(0), J_{n_2}(0), \cdots, J_{n_q}(0)] \in \mathbb{C}^{n\times n}$, where $m_1 \geq m_2 \geq \cdots \geq m_p > 0$ and $n_1 \geq n_2 \geq \cdots \geq n_q > 0$ and $C \in \mathbb{C}^{m\times n}$. For the set $M \subset \mathbb{N}_p \times \mathbb{N}_q$ (equivalently, matrix $M \in \mathbb{C}^{p\times q}$) given by (3.16), we have:*

i) $((A^T)^s C B^t)_M = (A^T)^s C_M B^t$, $(A^s C (B^T)^t)_M = A^s C_M (B^T)^t$, $s,t \in \mathbb{N}_0$,
ii) $(AC - CB)_M = AC_M - C_M B$.

Proof. Suppose C is decomposed in accordance with A and B as $C = [C_{ij}]_{p\times q}$.

i) The element in the position (i,j) for both left and right hand matrix is the same, namely

$$\begin{cases} (J_{m_i}(0)^T)^s C_{ij} J_{n_j}(0)^t, & (i,j) \in M, \\ 0, & (i,j) \in N, \end{cases}$$

so the equality holds. The second equality is proven analogously.

ii) We have desired equality because the element in the position (i,j) of the left and right hand matrix is the same:

$$\begin{cases} J_{m_i}(0) C_{ij} - C_{ij} J_{n_j}(0), & (i,j) \in M, \\ 0, & (i,j) \in N. \end{cases} \qquad \square$$

Let us split the matrix C as $C = C_M + C_N$, where $C_M = \Pi_M(C)$ and $C_N = \Pi_N(C)$; in the analogous way $X = X_M + X_N$. The reason for such splitting is the fact that solving simple Sylvester equation (2.6) is different whether $m \geq n$ or $m < n$, as shown in Theorem 2.3 and Theorem 2.5.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

THEOREM 3.5. *Let* $A = J_{m_1}(0) \oplus \cdots \oplus J_{m_p}(0),\ m_1 \geq \cdots \geq m_p > 0$, *and* $B = J_{n_1}(0) \oplus \cdots \oplus J_{n_q}(0),\ n_1 \geq \cdots \geq n_q > 0$. *Suppose* M *and* N *to be as in* (3.16), $C_M = \Pi_M(C),\ C_N = \Pi_N(C)$ *and* $d = \min\{m_1, n_1\}$. *The equation* $AX - XB = C$ *is consistent if and only if*

$$\sum_{k=0}^{d-1} A^{\langle k\rangle} C_M B^k = 0, \ \sum_{k=0}^{d-1} A^k C_N B^{\langle k\rangle} = 0, \tag{3.17}$$

or, in more condensed form,

$$\sum_{k=0}^{d-1} \begin{bmatrix} A^{\langle k\rangle} & 0 \\ 0 & A^k \end{bmatrix} \begin{bmatrix} C_M & 0 \\ 0 & C_N \end{bmatrix} \begin{bmatrix} B^k & 0 \\ 0 & B^{\langle k\rangle} \end{bmatrix} = 0.$$

The particular solution is given as $X_p = X_M + X_N$, *where*

$$X_M = \sum_{k=0}^{n_1-1} (A^T)^{k+1} C_M B^k, \ X_N = -\sum_{k=0}^{m_1-1} A^k C_N (B^T)^{k+1}, \tag{3.18}$$

and the solutions of the homogeneous equation $AX - XB = 0$ *are given by*

$$X_h = [X_{ij}^h]_{p\times q}, \ X_{ij}^h = \begin{cases} \left[\dfrac{p_{n_j-1}(J_{n_j}(0))}{0_{(m_i-n_j)\times n_j}}\right], & (i,j) \in M, \\ \begin{bmatrix} 0_{m_i\times(n_j-m_i)} & p_{m_i-1}(J_{m_i}(0)) \end{bmatrix}, & (i,j) \in N. \end{cases}$$

Proof. $(\Rightarrow):$ Let $\widehat{X}$ be a solution, so $A\widehat{X} - \widehat{X}B = C$. By Lemma 3.4, we have:

$$C_M = (A\widehat{X} - \widehat{X}B)_M = A\widehat{X}_M - \widehat{X}_M B, \ C_N = (A\widehat{X} - \widehat{X}B)_N = A\widehat{X}_N - \widehat{X}_N B.$$

Now we have

$$\begin{aligned} \sum_{k=0}^{d-1} A^{\langle k\rangle} C_M B^k &= \sum_{k=0}^{d-1} \left(A^{\langle k\rangle} A \widehat{X}_M B^k - A^{\langle k\rangle} \widehat{X}_M B^{k+1}\right) \\ &= \sum_{k=0}^{d-1} \left(A^{\langle k-1\rangle} \widehat{X}_M B^k - A^{\langle k\rangle} \widehat{X}_M B^{k+1}\right) \\ &= A^{\langle -1\rangle} \widehat{X}_M - A^{\langle d-1\rangle} \widehat{X}_M B^d = 0, \end{aligned}$$

because of Lemma 3.1. In the analogous way one can prove the second equality in (3.17).

$(\Leftarrow)$: We show that $X_p = X_M^p + X_N^p$, where X_M^p and X_N^p are given by (3.18), is the solution of $AX - XB = C$ provided that (3.17) holds. By using Lemma 3.4, we have:

$$\begin{aligned} [X_M^p]_{ij} &= \left[\sum_{k=0}^{n_1-1} (A^T)^{k+1} C_M B^k\right]_{ij} = \left[\sum_{k=0}^{n_1-1} \left((A^T)^{k+1} C B^k\right)_M\right]_{ij} \\ &= \sum_{k=0}^{n_j-1} (J_{m_i}(0)^T)^{k+1} C_{ij} J_{n_j}(0)^k, \ (i,j) \in M. \end{aligned}$$

The first condition in (3.17) says

$$\sum_{k=0}^{n_j-1} J_{m_i}(0)^{\langle k\rangle} C_{ij} J_{n_j}(0)^k = 0, \ (i,j) \in M,$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

which by Theorem 2.3 implies that $J_{m_i}(0)[X_M^p]_{ij} - [X_M^p]_{ij} J_{n_j}(0) = C_{ij},\ (i,j) \in M$, i.e., $AX_M^p - X_M^p B = C_M$. In the analogous way $AX_N^p - X_N^p B = C_N$ can be proven (there, Theorem 2.5 is used).

Therefore, we have $AX_p - X_p B = AX_M^p - X_M^p B + AX_N^p - X_N^p B = C_M + C_N = C$, so the equation is consistent provided that (3.17) holds.

The homogeneous equation $AX = XB$ is equivalent to the set of equations of the form $J_{m_i}(0)X_{ij} - X_{ij}J_{n_i}(0) = 0,\ i = \overline{1,p},\ j = \overline{1,q}$, so by Corollary 2.6, we have the expressions for appropriate submatrices of X_h. ∎

Remark 3.6. It may happen that some expressions for the consistency condition (3.17) include negative powers of $J_k(0)$, but all of them will be multiplied by some zero terms, so there is no actual problem.

The next theorem immediately follows from Proposition 1.3, but we restate it here because of the results from the next section.

Theorem 3.7. *Suppose that* $A(\lambda) = J_{m_1}(\lambda) \oplus \cdots \oplus J_{m_p}(\lambda) \in \mathbb{C}^{m\times m}$ *and* $B(\mu) = J_{n_1}(\mu) \oplus \cdots \oplus J_{n_q}(\mu) \in \mathbb{C}^{n\times n}$ *and* $a = \max\{m_1, \ldots, m_p\}$, $b = \max\{n_1, \ldots, n_q\}$. *The equation* $A(\lambda)X - XB(\mu) = C$, $\lambda \neq \mu$, *is consistent and its unique solution is*

$$X = \sum_{k=0}^{b-1} A(\lambda - \mu)^{-(k+1)} CB(0)^k = -\sum_{k=0}^{a-1} A(0)^k CB(\mu - \lambda)^{-(k+1)}.$$

4. The general case when $\sigma(A) = \sigma(B)$. Let us consider general case $AX - XB = C$ when $\sigma(A) = \sigma(B) = \{\lambda_1, \ldots, \lambda_s\}$, $C = [C_{ij}]_{s\times s}$ and suppose that the matrices are in their Jordan forms:

$$A = diag\{J(\lambda_1; p_{11}, p_{12}, \ldots, p_{1,k_1}), \ldots, J(\lambda_s; p_{s1}, p_{s2}, \ldots, p_{s,k_s})\},$$
$$B = diag\{J(\lambda_1; q_{11}, q_{12}, \ldots, q_{1,\ell_1}), \ldots, J(\lambda_s; q_{s1}, q_{s2}, \ldots, q_{s,\ell_s})\},$$

where $p_{i1} \geq \cdots \geq p_{i,k_i} > 0,\ q_{j1} \geq \cdots \geq q_{j,\ell_j} > 0,\ i,j = \overline{1,s}$.

It is not hard to see that the equations on which solvability depends are precisely of the form

$$J(\lambda_i; p_{i1}, p_{i2}, \ldots, p_{i,k_i})X_{ii} - X_{ii}J(\lambda_i; q_{i1}, q_{i2}, \ldots, q_{i,\ell_i}) = C_{ii}, \quad i = \overline{1,s},$$

and, after translation for $-\lambda_i$, they reduce to the form

$$J(0; p_{i1}, p_{i2}, \ldots, p_{i,k_i})X_{ii} - X_{ii}J(0; q_{i1}, q_{i2}, \ldots, q_{i,\ell_i}) = C_{ii}, \quad i = \overline{1,s},$$

for which Theorem 3.5 is applicable.

The results from the previous two sections can be summed up in the following theorem. We remark the notation $C_{ij} = [C_{uv}^{(ij)}]_{u=\overline{1,k_i}, v=\overline{1,\ell_j}} \in \mathbb{C}^{p_i \times q_j}$.

Theorem 4.1. *Suppose that*

$$A = diag\{J(\lambda_1; p_{11}, p_{12}, \ldots, p_{1,k_1}), \ldots, J(\lambda_s; p_{s1}, p_{s2}, \ldots, p_{s,k_s})\},$$
$$B = diag\{J(\lambda_1; q_{11}, q_{12}, \ldots, q_{1,\ell_1}), \ldots, J(\lambda_s; q_{s1}, q_{s2}, \ldots, q_{s,\ell_s})\},$$

where $p_{i1} \geq \cdots \geq p_{i,k_i} > 0$, *and* $q_{j1} \geq \cdots \geq q_{j,\ell_j} > 0,\ i,j = \overline{1,s}$. *Let us denote:* $d_i = \min\{p_{i1}, q_{i1}\}$, $C_{M_i}^{(ii)} = \Pi_{M_i}(C^{(ii)})$, $C_{N_i}^{(ii)} = \Pi_{N_i}(C^{(ii)})$, *where*

$$M_i = \{(u,v) : p_{iu} \geq q_{iv}\}, \quad N_i = \{(u,v) : p_{iu} < q_{iv}\} \subset \mathbb{N}_{k_i} \times \mathbb{N}_{\ell_i}, \quad i = \overline{1,s}.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

The Sylvester equation $AX-XB=C$ is consistent if and only if

$$\sum_{k=0}^{d_i-1}\begin{bmatrix} A_i(0)^{\langle k\rangle} & 0 \\ 0 & A_i(0)^k \end{bmatrix}\begin{bmatrix} C_{M_i}^{(ii)} & 0 \\ 0 & C_{N_i}^{(ii)} \end{bmatrix}\begin{bmatrix} B_i(0)^k & 0 \\ 0 & B_i(0)^{\langle k\rangle} \end{bmatrix}=0,\quad i=\overline{1,s},$$

where we used the notation

$$A_i(0)=diag[J_{p_{i1}}(0),\ldots,J_{p_{i,k_i}}(0)],\ \ B_i(0)=diag[J_{q_{i1}}(0),\ldots,J_{q_{i,\ell_i}}(0)],\ \ i=\overline{1,s}.$$

In that case, the particular solution $X_p=[X_p^{(ij)}]_{s\times s}$ is given by:

$$\begin{aligned} X_p^{(ii)} &= \sum_{k=0}^{d_i-1}(A_i(0)^T)^{k+1}C_{M_i}^{(ii)}B_i(0)^k-\sum_{k=0}^{d_i-1}A_i(0)^kC_{N_i}^{(ii)}(B_i(0)^T)^{k+1},\ \ i=\overline{1,s},\\ X_p^{(ij)} &= \sum_{k=0}^{q_{j1}-1}A_i(\lambda_i-\lambda_j)^{-(k+1)}C^{(ij)}B_j(0)^k\\ &= -\sum_{k=0}^{p_{i1}-1}A_i(0)^kC^{(ij)}B_j(\lambda_j-\lambda_i)^{-(k+1)}\in\mathbb{C}^{k_i\times\ell_j},\ \ i\neq j; \end{aligned}$$

the homogeneous solution $X_h=diag[X_h^{(ii)}],\ i=\overline{1,s}$, is given by

$$\left[X_h^{(ii)}\right]_{uv}=\begin{cases} \begin{bmatrix} p_{q_{iv}-1}(J_{q_{iv}}(0)) \\ 0_{(p_{iu}-q_{iv})\times q_{iv}} \end{bmatrix}, & (u,v)\in M_i,\\ \begin{bmatrix} 0_{p_{iu}\times(q_{iv}-p_{iu})} & p_{p_{iu}-1}(J_{p_{iu}}(0)) \end{bmatrix}, & (u,v)\in N_i. \end{cases}$$

Remark 4.2. The consistency condition from the previous theorem can be written in more condensed form as:

$$\sum_{k=0}^{d-1}\begin{bmatrix} A(0)^{\langle k\rangle} & 0 \\ 0 & A(0)^k \end{bmatrix}\begin{bmatrix} C_M & 0 \\ 0 & C_N \end{bmatrix}\begin{bmatrix} B(0)^k & 0 \\ 0 & B(0)^{\langle k\rangle} \end{bmatrix}=0,$$

where $d=\min\{\max\{p_{i1}:i=\overline{1,s}\},\max\{q_{j1}:j=\overline{1,s}\}\},\ M=M_1\oplus\cdots\oplus M_s,\ N=N_1\oplus\cdots\oplus N_s$, and we used the notation

$$A(0)=diag[A_1(0),\ldots,A_s(0)],\ B(0)=diag[B_1(0),\ldots,B_s(0)].$$

Also, the diagonal blocks in X_p can be given by single explicit formula

$$X_p^{diag}=\sum_{k=0}^{d-1}\left((A(0)^T)^{k+1}C_MB(0)^k-A(0)^kC_N(B(0)^T)^{k+1}\right).$$

The previous theorem has several important corollaries which deal with the cases when matrices A and B are non-derogatory or diagonalizable.

Corollary 4.3. *Suppose that matrices A and B are non-derogatory, i.e.,*

$$A=J_{p_1}(\lambda_1)\oplus\cdots\oplus J_{p_s}(\lambda_s),\ \ B=J_{q_1}(\lambda_1)\oplus\cdots\oplus J_{q_s}(\lambda_s),$$

and $d_i=\min\{p_i,q_i\},\ C_{M_i}^{(ii)}=C^{(ii)}$ if $p_i\geq q_i,\ C_{N_i}^{(ii)}=C^{(ii)}$ if $p_i<q_i,\ i=\overline{1,s}$. The equation $AX-XB=C$ is consistent if and only if

$$\sum_{k=0}^{d_i-1}\begin{bmatrix} J_{p_i}(0)^{p_i-1-k} & 0 \\ 0 & J_{p_i}(0)^k \end{bmatrix}\begin{bmatrix} C_{M_i}^{(ii)} & 0 \\ 0 & C_{N_i}^{(ii)} \end{bmatrix}\begin{bmatrix} J_{q_i}(0)^k & 0 \\ 0 & J_{q_i}(0)^{q_i-1-k} \end{bmatrix}=0,\ \ i=\overline{1,s},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

and its particular solution $X_p = [X_p^{(ij)}]_{s\times s}$ *is given by:*

$$X_p^{(ii)} = \begin{cases} \sum\limits_{k=0}^{d_i-1} (J_{p_i}(0)^T)^{k+1} C^{(ii)} J_{q_i}(0)^k, & p_i \geq q_i, \\ -\sum\limits_{k=0}^{d_i-1} J_{p_i}(0)^k C^{(ii)} (J_{q_i}(0)^T)^{k+1}, & p_i < q_i; \end{cases}$$

$$\begin{aligned} X_p^{(ij)} &= \sum_{k=0}^{q_j-1} J_{p_i}(\lambda_i - \lambda_j)^{-(k+1)} C^{(ij)} J_{q_j}(0)^k \\ &= -\sum_{k=0}^{p_i-1} J_{p_i}(0)^k C^{(ij)} J_{q_j}(\lambda_j - \lambda_i)^{-(k+1)} \in \mathbb{C}^{p_i \times q_j}, \ \ i \neq j; \end{aligned}$$

the homogeneous solution $X_h = diag[X_h^{(ii)}],\ i = \overline{1,s}$, *is given by*

$$X_h^{(ii)} = \begin{cases} \begin{bmatrix} p_{q_i-1}(J_{q_i}(0)) \\ 0_{(p_i-q_i)\times q_i} \end{bmatrix}, & p_i \geq q_i, \\ \begin{bmatrix} 0_{p_i\times(q_i-p_i)} & p_{p_i-1}(J_{p_i}(0)) \end{bmatrix}, & p_i < q_i. \end{cases}$$

Corollary 4.4. *Suppose that A and B are diagonalizable matrices, i.e.,*

$$A = \lambda_1 I_{k_1} \oplus \cdots \oplus \lambda_s I_{k_s}, \ \ B = \lambda_1 I_{\ell_1} \oplus \cdots \oplus \lambda_s I_{\ell_s}.$$

Then the Sylvester equation $AX - XB = C$ is consistent if and only if $C^{(ii)} = 0,\ i = \overline{1,s}$, its homogeneous solution $X_h = X_h^{(11)} \oplus \cdots \oplus X_h^{(ss)}$ is arbitrary block-diagonal matrix and the particular solution is $X_p = [X_p^{(ij)}]_{s\times s}$ with $X_p^{(ii)} = 0,\ i = \overline{1,s}$, and

$$X_p^{(ij)} = (\lambda_i - \lambda_j)^{-1} C^{(ij)}, \ \ i \neq j.$$

The fact that $AX - XA \neq I_n$ for $A, X \in \mathbb{C}^{n\times n}$ is well-known result, usually proven by the trace of a matrix:

$$n = tr(I_n) \neq tr(AX - XA) = tr(AX) - tr(XA) = tr(AX) - tr(AX) = 0.$$

This result appears as the corollary of Theorem 4.1.

Theorem 4.5. *The matrix equation $AX - XA = I$ is inconsistent.*

Proof. According to Theorem 4.1, if $B = A$ then $d_i = p_{i1},\ C_{M_i}^{(ii)} = I,\ i = \overline{1,s}$, and the equation $AX - XA = I$ is consistent if and only if

$$\begin{aligned} &\sum_{k=0}^{p_{i1}-1} diag[J_{p_{i1}}(0)^{p_{i1}-1-k}, \ldots, J_{p_{i,k_i}}(0)^{p_{i,k_i}-1-k}] \cdot diag[J_{p_{i1}}(0)^k, \ldots, J_{p_{i,k_i}}(0)^k] \\ &= \sum_{k=0}^{p_{i1}-1} diag[J_{p_{i1}}(0)^{p_{i1}-1}, \ldots, J_{p_{i,k_i}}(0)^{p_{i,k_i}-1}] = 0, \ \ i = \overline{1,s}, \end{aligned}$$

which is not true. Therefore, the equation $AX - XA = I$ is inconsistent. ∎

Recall that the similar result holds on any unital Banach algebra [16, p. 351].

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

EXAMPLE 4.6. Let $A = J_2(1) \oplus J_1(1) \oplus J_2(0)$, $B = J_2(1) \oplus J_2(1) \oplus J_3(0) \oplus J_1(0)$ (so $s = 2$, $\lambda_1 = 1$, $\lambda_2 = 0$, $p_{11} = 2$, $p_{12} = 1$, $p_{21} = 2$, $q_{11} = 2$, $q_{12} = 2$, $q_{21} = 3$, $q_{22} = 1$). We have: $A_1(0) = J_2(0) \oplus J_1(0)$, $A_2(0) = J_2(0)$, $B_1(0) = J_2(0) \oplus J_2(0)$, $B_2(0) = J_3(0) \oplus J_1(0)$. Also $d_1 = \min\{p_{11}, q_{11}\} = 2$, $d_2 = \min\{p_{21}, q_{21}\} = 2$.

We decompose matrices C and X in accordance with the forms of A and B:

$$C = \begin{bmatrix} C^{(11)} & C^{(12)} \\ C^{(21)} & C^{(22)} \end{bmatrix}_{5\times 8}, \quad X = \begin{bmatrix} X^{(11)} & X^{(12)} \\ X^{(21)} & X^{(22)} \end{bmatrix}_{5\times 8}$$

and then further decompose $C^{(11)}$ and $C^{(22)}$ in accordance with the mask matrices $M_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ (because $p_{11} \geq q_{11}$ and $p_{11} \geq q_{12}$) and $M_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}$ (because $p_{21} \geq q_{22}$):

$$C^{(11)} = \begin{bmatrix} C_{11}^{(11)} & C_{12}^{(11)} \\ C_{21}^{(11)} & C_{22}^{(11)} \end{bmatrix} = \begin{bmatrix} C_{11}^{(11)} & C_{12}^{(11)} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ C_{21}^{(11)} & C_{22}^{(11)} \end{bmatrix} = C_{M_1}^{(11)} + C_{N_1}^{(11)},$$

$$C^{(22)} = \begin{bmatrix} C_{11}^{(22)} & C_{12}^{(22)} \end{bmatrix} = \begin{bmatrix} 0 & C_{12}^{(22)} \end{bmatrix} + \begin{bmatrix} C_{11}^{(22)} & 0 \end{bmatrix} = C_{M_2}^{(22)} + C_{N_2}^{(22)};$$

similar is done for the matrix X. By Theorem 4.1, consistency conditions are (we simply write J_k instead of $J_k(0)$ and omit zero block-matrices):

$$0 = \sum_{k=0}^{1} \begin{bmatrix} J_2^{1-k} & & & \\ & J_1^{-k} & & \\ & & J_2^{k} & \\ & & & J_1^{k} \end{bmatrix} \begin{bmatrix} C_{11}^{(11)} & C_{12}^{(11)} & & \\ 0 & 0 & & \\ & & 0 & 0 \\ & & C_{21}^{(11)} & C_{22}^{(11)} \end{bmatrix} \begin{bmatrix} J_2^{k} & & & \\ & J_2^{k} & & \\ & & J_2^{1-k} & \\ & & & J_2^{1-k} \end{bmatrix},$$

$$0 = \sum_{k=0}^{1} \begin{bmatrix} J_2^{1-k} & \\ & J_2^{k} \end{bmatrix} \begin{bmatrix} 0 & C_{12}^{(22)} & & \\ & & C_{11}^{(22)} & 0 \end{bmatrix} \begin{bmatrix} J_3^{k} & & & \\ & J_1^{k} & & \\ & & J_3^{2-k} & \\ & & & J_1^{-k} \end{bmatrix},$$

i.e., $J_2 C_{11}^{(11)} + C_{11}^{(11)} J_2 = 0$, $J_2 C_{12}^{(11)} + C_{12}^{(11)} J_2 = 0$, $C_{21}^{(11)} J_2 = 0$, $C_{22}^{(11)} J_2 = 0$, $J_2 C_{12}^{(22)} = 0$, $C_{11}^{(22)} J_3^2 + J_2 C_{11}^{(22)} J_3 = 0$.

If we put $C = [c_{ij}]$, $X = [x_{ij}] \in \mathbb{C}^{5\times 8}$, consistency condition says that all matrices C for whom the equation is solvable are of the form

$$C = \begin{bmatrix} C^{(11)} & C^{(12)} \\ C^{(21)} & C^{(22)} \end{bmatrix} = \left[\begin{array}{cc|cc||ccc|c} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} & c_{17} & c_{18} \\ 0 & -c_{11} & 0 & -c_{13} & c_{25} & c_{26} & c_{27} & c_{28} \\ \hline 0 & c_{32} & 0 & c_{34} & c_{35} & c_{36} & c_{37} & c_{38} \\ \hline\hline c_{41} & c_{42} & c_{43} & c_{44} & c_{45} & c_{46} & c_{47} & c_{48} \\ c_{51} & c_{52} & c_{53} & c_{54} & 0 & -c_{45} & c_{57} & 0 \end{array} \right].$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

By Theorem 4.1, the particular solution $X_p = [X_p^{(ij)}]$ is:

$$\begin{aligned}
X_p^{(11)} &= \sum_{k=0}^{1} (A_1(0)^T)^{k+1} C_{M_1}^{(11)} B_1(0)^k - \sum_{k=0}^{1} A_1(0)^k C_{N_1}^{(11)} (B_1(0)^T)^{k+1} \\
&= \begin{bmatrix} J_2(0)^T C_{11}^{(11)} & J_2(0)^T C_{12}^{(11)} \\ -C_{21}^{(11)} J_2(0)^T & -C_{22}^{(11)} J_2(0)^T \end{bmatrix}, \\
X_p^{(22)} &= \sum_{k=0}^{1} (A_2(0)^T)^{k+1} C_{M_2}^{(22)} B_2(0)^k - \sum_{k=0}^{1} A_2(0)^k C_{N_2}^{(22)} (B_2(0)^T)^{k+1} \\
&= \begin{bmatrix} -C_{11}^{(22)} J_3(0)^T - J_2(0) C_{11}^{(22)} (J_3(0)^T)^2 & J_2(0)^T C_{12}^{(22)} \end{bmatrix}, \\
X_p^{(12)} &= \sum_{k=0}^{2} A_1(1)^{-(k+1)} C^{(12)} B_2(0)^k \\
&= \begin{bmatrix} J_2(1)^{-1} C_{11}^{(12)} + J_2(1)^{-2} C_{11}^{(12)} J_3(0) + J_2(1)^{-3} C_{11}^{(12)} J_3(0)^2 & J_2(1)^{-1} C_{12}^{(12)} \\ C_{21}^{(12)} + C_{21}^{(12)} J_3(0) + C_{21}^{(12)} J_3(0)^2 & C_{22}^{(12)} \end{bmatrix}, \\
X_p^{(21)} &= -\sum_{k=0}^{1} A_2(0)^k C^{(21)} B_1(1)^{-(k+1)} \\
&= \begin{bmatrix} -C_{11}^{(21)} J_2(1)^{-1} - J_2(0) C_{11}^{(21)} J_2(1)^{-2} & -C_{12}^{(21)} J_2(1)^{-1} - J_2(0) C_{12}^{(22)} J_2(1)^{-2} \end{bmatrix}.
\end{aligned}$$

Therefore,

$$X_p^{(11)} = \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ c_{11} & c_{12} & c_{13} & c_{14} \\ \hline -c_{32} & 0 & -c_{34} & 0 \end{array}\right], \quad X_p^{(22)} = \left[\begin{array}{ccc|c} -c_{46} - c_{57} & -c_{47} & 0 & 0 \\ c_{45} & -c_{57} & 0 & c_{48} \end{array}\right]$$

$$X_p^{(12)} = \left[\begin{array}{ccc|c} c_{15} - c_{25} & c_{15} + c_{16} - 2c_{25} - c_{26} & c_{15} + c_{16} + c_{17} - 3c_{25} - 2c_{26} - c_{27} & c_{18} - c_{28} \\ c_{25} & c_{25} + c_{26} & c_{25} + c_{26} + c_{27} & c_{28} \\ \hline c_{35} & c_{35} + c_{36} & c_{35} + c_{36} + c_{37} & c_{38} \end{array}\right]$$

$$X_p^{(21)} = \left[\begin{array}{cc|cc} -c_{41} - c_{51} & c_{41} + 2c_{51} - c_{42} - c_{52} & -c_{43} - c_{53} & c_{43} - c_{44} + 2c_{53} - c_{54} \\ -c_{51} & c_{51} - c_{52} & -c_{53} & c_{53} - c_{54} \end{array}\right].$$

The homogeneous solution is given by $X_h = [X_h^{(ii)}]$ where:

$$X_h^{(11)} = \begin{bmatrix} p_1(J_2(0)) & p_1(J_2(0)) \\ \begin{bmatrix} 0_{1\times 1} & p_0(J_1(0)) \end{bmatrix} & \begin{bmatrix} 0_{1\times 1} & p_0(J_1(0)) \end{bmatrix} \end{bmatrix},$$

$$X_h^{(22)} = \begin{bmatrix} \begin{bmatrix} 0_{2\times 1} & p_1(J_2(0)) \end{bmatrix} & \begin{bmatrix} p_0(J_1(0)) \\ 0_{1\times 1} \end{bmatrix} \end{bmatrix}$$

or, in extended form:

$$X_h^{(11)} = \left[\begin{array}{cc|cc} x_{11} & x_{12} & x_{13} & x_{14} \\ 0 & x_{11} & 0 & x_{13} \\ \hline 0 & x_{32} & 0 & x_{34} \end{array}\right], \quad X_h^{(22)} = \left[\begin{array}{ccc|c} 0 & x_{46} & x_{47} & x_{48} \\ 0 & 0 & x_{46} & 0 \end{array}\right]$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

for arbitrary complex $x_{11}, x_{12}, x_{13}, x_{14}, x_{32}, x_{34}, x_{46}, x_{47}$ and x_{48}. Therefore, we found all solutions of the equation in the case when it is solvable.

5. On the dimension of the space of solutions. Let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be given matrices. For the linear operator $\mathcal{T}(X) = AX - XB : \mathbb{C}^{m\times n} \to \mathbb{C}^{m\times n}$ we consider the following subspaces:

- $\mathcal{R}(\mathcal{T}) = \{C : AX - XB = C \text{ for some } X\}$,
- $\mathcal{N}(\mathcal{T}) = \{X : AX - XB = 0\}$.

The subspace $\mathcal{R}(\mathcal{T})$ is actually the set of all C such that the Sylvester equation is consistent, while $\mathcal{N}(\mathcal{T})$ consists precisely of all solutions of the homogeneous Sylvester equation. The general solution can be seen as the set of all homogeneous solutions translated by the particular solution, so the dimension of the space of solutions is the same as the dimension of the space of all solutions to homogeneous equation. By the basic theorem for the vector space homomorphism, we have:

$$\dim \mathcal{R}(\mathcal{T}) + \dim \mathcal{N}(\mathcal{T}) = \dim(\mathbb{C}^{m\times n}) = mn.$$

This fact helps us a lot in finding the dimension of $\mathcal{R}(\mathcal{T})$ when necessary, because it is significantly easier to deal with $\mathcal{N}(\mathcal{T})$, as follows.

If $A = J_m(0),\ B = J_n(0)$, then by Corollary 2.6, we have:

$$\mathcal{N}(\mathcal{T}) = \begin{cases} \begin{bmatrix} p_{n-1}(J_n(0)) \\ \hline 0_{(m-n)\times n} \end{bmatrix}, & m \geq n, \\ \begin{bmatrix} 0_{m\times(n-m)} & p_{m-1}(J_m(0)) \end{bmatrix}, & m \leq n, \end{cases}$$

and therefore, $\dim \mathcal{N}(\mathcal{T}) = \min\{m, n\}$ (so $\dim \mathcal{R}(\mathcal{T}) = mn - \min\{m, n\}$) and its basis is

$$\mathcal{B}_{\mathcal{N}(\mathcal{T})} = \left\{ \begin{bmatrix} I_n \\ 0 \end{bmatrix}, \begin{bmatrix} J_n(0) \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} J_n(0)^{n-1} \\ 0 \end{bmatrix} \right\},\ m \geq n,$$
$$\mathcal{B}_{\mathcal{N}(\mathcal{T})} = \left\{ \begin{bmatrix} 0 & I_m \end{bmatrix}, \begin{bmatrix} 0 & J_m(0) \end{bmatrix}, \ldots, \begin{bmatrix} 0 & J_m(0)^{m-1} \end{bmatrix} \right\},\ m \leq n.$$

For the case $A = J_{m_1}(0) \oplus \cdots \oplus J_{m_p}(0)$ and $B = J_{n_1}(0) \oplus \cdots \oplus J_{n_q}(0)$, by Theorem 3.5, we have $\mathcal{N}(\mathcal{T}) = \{X = [X_{ij}]_{p\times q} : J_{m_i}(0)X_{ij} = X_{ij}J_{n_j}(0),\ i = \overline{1,p},\ j = \overline{1,q}\}$, so

$$\dim \mathcal{N}(\mathcal{T}) = \sum_{(i,j)\in M} n_j + \sum_{(i,j)\in N} m_i.$$

In the most general case, described by Theorem 4.1, we have

$$\dim \mathcal{N}(\mathcal{T}) = \sum_{k=1}^{s} \left(\sum_{(i,j)\in M_k} q_{kj} + \sum_{(i,j)\in N_k} p_{ki} \right).$$

For the non-derogatory case, Corollary 4.3 gives

$$\dim \mathcal{N}(\mathcal{T}) = \sum_{i=1}^{s} \min\{p_i, q_i\},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

while in the diagonalizable case, given in Corollary 4.4, we have

$$\dim \mathcal{N}(\mathcal{T}) = \sum_{i=1}^{s} k_i \cdot \ell_i.$$

If we test this on Example 4.6, then we have $\dim \mathcal{N}(T) = (2+2+1) + (1+1+2) = 9$ which coincides with the number of independent parameters in the general solution.

6. The Schur decomposition approach. Finding the Jordan form of a given matrix having some multiple eigenvalues can be numerically unstable, so in this section, which is not closely related with previous considerations, we present some useful recursive methods for solving the Sylvester matrix equation $AX - XB = C$ in the case when A and B have common eigenvalue(s). We present the consistency conditions and the algorithm for finding the solutions, based on the Schur matrix decomposition.

For numerically solving the Sylvester equations, the most used methods are Bartels-Stewart [1] and Golub-Nash-van Loan [8], and their various improvements. Bartels-Stewart method uses QR-algorithm for the Schur decomposition of matrices A and B, while Golub-Nash-van Loan method uses the Hessenberg decomposition of matrices A and B. We emphasize that both methods assume disjointness of the spectra of A and B.

The main idea of the Bartels-Stewart algorithm is to apply the Schur decomposition to transform Sylvester equation into a triangular linear system which can be solved efficiently by forward or backward substitutions.

It is known that $A \in \mathbb{C}^{n\times n}$ can be expressed as $A = QUQ^{-1}$, where Q is unitary matrix (i.e., $Q^{-1} = Q^*$), and U is upper triangular matrix, which is called a Schur form for A. Since U is similar to A, it has the same multiset of eigenvalues, and since it is triangular, those eigenvalues are the diagonal entries of U. If A is real and $\sigma(A) \subset \mathbb{R}$ then U can be chosen to be real and orthogonal. For details on this topic, see e.g., [10, p. 79].

By using the idea of the constructive proof for Schur decomposition, suppose $\lambda \in \sigma(A) \cap \sigma(B)$ and there exist some unitary matrices S and T such that (a and b are the multiplicities of λ in $\sigma(A)$ and $\sigma(B)$, respectively; U_λ and V_λ are appropriate eigenspaces):

$$S^*AS = \begin{bmatrix} \lambda I_a & A_{12} \\ 0 & A_{22} \end{bmatrix} : U_\lambda \oplus U_\lambda^\perp \to U_\lambda \oplus U_\lambda^\perp,$$

$$T^*BT = \begin{bmatrix} \lambda I_b & B_{12} \\ 0 & B_{22} \end{bmatrix} : V_\lambda \oplus V_\lambda^\perp \to V_\lambda \oplus V_\lambda^\perp.$$

Note that $\lambda \notin \sigma(A_{22}) \cup \sigma(B_{22})$. Now the Sylvester equation becomes:

$$S \begin{bmatrix} \lambda I_a & A_{12} \\ 0 & A_{22} \end{bmatrix} S^*X - XT \begin{bmatrix} \lambda I_b & B_{12} \\ 0 & B_{22} \end{bmatrix} T^* = C$$

$$\Leftrightarrow \begin{bmatrix} \lambda I_a & A_{12} \\ 0 & A_{22} \end{bmatrix} S^*XT - S^*XT \begin{bmatrix} \lambda I_b & B_{12} \\ 0 & B_{22} \end{bmatrix} = S^*CT,$$

so the equation, with $S^*XT = Y = [Y_{ij}]$ and $S^*CT = D = [D_{ij}]$, becomes

$$\begin{bmatrix} \lambda I_a & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} - \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} \lambda I_b & B_{12} \\ 0 & B_{22} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

which is equivalent to the following system:

$$A_{12}Y_{21} = D_{11}, \tag{6.19}$$

$$Y_{12}(\lambda I - B_{22}) - Y_{11}B_{12} = D_{12} - A_{12}Y_{22}, \tag{6.20}$$

$$(A_{22} - \lambda I)Y_{21} = D_{21}, \tag{6.21}$$

$$A_{22}Y_{22} - Y_{22}B_{22} = D_{22} + Y_{21}B_{12}. \tag{6.22}$$

From (6.19) and (6.21) we conclude **the consistency condition**:

$$A_{12}(A_{22} - \lambda I)^{-1}D_{21} = D_{11}.$$

Also from (6.21) we have

$$Y_{21} = (A_{22} - \lambda I)^{-1}D_{21}.$$

Note that A_{22} and B_{22} (and consequently $A_{22} - \lambda I$ and $\lambda I - B_{22}$) are upper triangular matrices, so some known numerical method can be applied in computing their inverses.

Case 1: If $\sigma(A)\cap\sigma(B) = \{\lambda\}$, i.e., if λ is the only common eigenvalue (and therefore, $\sigma(A_{22})\cap\sigma(B_{22}) = \emptyset$), because of (6.22), by Proposition 1.3, we have the unique Y_{22} (which can be computed by using some already known numerical method for solving the Sylvester equation), while from (6.20) we have Y_{12} via Y_{11}:

$$Y_{12} = (D_{12} - A_{12}Y_{22} + Y_{11}B_{12})(\lambda I - B_{22})^{-1}.$$

Therefore, we found the family of solutions.

Case 2: If there is some other common eigenvalue μ for A and B (which means $\mu \in \sigma(A_{22}) \cap \sigma(B_{22})$), the described method is applied to the equation (6.22), i.e.,

$$A_{22}Y_{22} - Y_{22}B_{22} = D_{22} + (A_{22} - \lambda I)^{-1}D_{21}B_{12}.$$

It remains to do backward substitution and the equation is fully solved.

Because any complex matrix from $\mathbb{C}^{n\times n}$ has at most n different eigenvalues, the algorithm terminates after at most n passes.

Acknowledgments. The author thanks to his colleague Miloš Cvetković for a valuable discussion on some topics presented in the paper. The author would like to express his gratitude to the two anonymous referees and the editor, Prof. Froilán M. Dopico, for their insightful comments and suggestions which significantly improved the paper.

REFERENCES

[1] R.H. Bartels and G.W. Stewart. A solution of the equation $AX + XB = C$. *Comm. ACM*, 15:820–826, 1972.

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

[3] Y. Chen and H. Xiao. The explicit solution of the matrix equation $AX - XB = C$. *Appl. Math. Mech. (English Ed.)*, 16:1133–1141, 1995.

[4] B.N. Datta and K. Datta. The matrix equation $XA = A^TX$ and an associated algorithm for solving the inertia and stability problems. *Linear Algebra Appl.*, 97:103–119, 1987.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 35, pp. 1-23, January 2019.

[5] F. De Terán, F.M. Dopico, N. Guillery, D. Montealegre, and N. Reyes. The solution of the equation $AX + X^{\star}B = 0$. *Linear Algebra Appl.*, 438:2817–2860, 2013.

[6] M.P. Drazin. On a result of J. J. Sylvester. *Linear Algebra Appl.*, 505:361–366, 2016.

[7] F.R. Gantmacher. *The Theory of Matrices*, Vol. 1. Chelsea Publish. Co., New York, 1959.

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

[9] E. Heinz. Beiträge zur Störungstheorie der Spektralzerlegung. *Math. Ann.*, 123:415–438, 1951.

[10] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[11] Q. Hu and D. Cheng. The polynomial solution to the Sylvester matrix equation. *Appl. Math. Lett.*, 19:859–864, 2006.

[12] Z.-Y. Li and B. Zhou. Spectral decomposition based solutions to the matrix equation $AX - XB = C$. *IET Control Theory Appl.*, 12:119–128, 2018.

[13] E.-C. Ma. A finite series solution of the matrix equation $AX - XB = C$. *SIAM J. Appl. Math.*, 14:490–495, 1966.

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

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

[16] W. Rudin. *Functional Analysis*, second edition. McGraw-Hill, Inc., New York, 1991.

[17] J.J. Sylvester. Sur l'equations en matrices $px = xq$. *C. R. Acad. Sci. Paris*, 99:67–71, 115–116, 1884.

[18] A.-G. Wu, G. Feng, G.-R. Duan, and W-J. Wu. Closed-form solutions to Sylvester-conjugate matrix equations. *Comput. Math. Appl.*, 60:95-111, 2010.