An Arithmetic for Matrix Pencils: Theory and New Algorithms

Peter Benner*
Fakultät für Mathematik
TU Chemnitz
D-09107 Berlin
Germany
benner@mathematik.tu-chemnitz.de

Ralph Byers†
Department of Mathematics
University of Kansas
Lawrence, Kansas 66045
USA
byers@math.ukans.edu

April 2004

Abstract

This paper introduces arithmetic-like operations on matrix pencils. The pencil-arithmetic operations extend elementary formulas for sums and products of rational numbers and include the algebra of linear transformations as a special case. These operation give an unusual perspective on a variety of pencil related computations. We derive generalizations of monodromy matrices and the matrix exponential. A new algorithm for computing a pencil-arithmetic generalization of the matrix sign function does not use matrix inverses and gives an empirically forward numerically stable algorithm for extracting deflating subspaces.

1 Introduction

Corresponding to each matrix pencil $\lambda E - A$, E, $A \in \mathbb{C}^{m\times n}$, define the (left-handed) *matrix relation* on $\mathbb{C}^n$ to be

$$(E\backslash A) = \{(x,y) \in \mathbb{C}^n \times \mathbb{C}^n \mid Ey = Ax\}. \tag{1}$$

Matrix relations are vector subspaces of $\mathbb{C}^n \times \mathbb{C}^n$. In category theory, a matrix relation is called the *pullback* of E and A [37]. If E is a nonsingular n-by-n matrix, then $(E\backslash A)$ is the linear transformation with matrix representation $E^{-1}A$. If E does not have full column rank, then $(E\backslash A)$ might be described as a multi-valued linear transformation.

The linear descriptor difference equation $E_k x_{k+1} = A_k x_k$, $E_k \in \mathbb{C}^{m\times n}$ and $A_k \in \mathbb{C}^{m\times n}$ is equivalent to $(x_k, x_{k+1}) \in (E_k\backslash A_k)$. Similarly, the linear differential algebraic equation $E(t)\dot{x}(t) = A(t)x(t)$ is equivalent to $(x, \dot{x}) \in (E(t)\backslash A(t))$ [11].

We are especially interested in the case in which $m \neq n$ and/or E and/or A are rank deficient. However, for computational purposes, there are some advantages to representing $(E\backslash A)$ in terms of two matrices even when $m = n$ and E is nonsingular. If E is ill-conditioned with respect to inversion, then forming $E^{-1}A$ explicitly may introduce destructive levels of rounding error. Also, $(M\backslash I)$ is an inexpensive and rounding error free representation of M^{-1}. It is this observation that makes variations of the AB algorithm [32] and inverse-free spectral divide and conquer algorithms [7, 9, 38] free of inverses.

*Some of this work was completed at the University of Kansas. Partial support received by Deutsche Forschungsgemeinschaft, grant BE 2174/4-1.

†This material is based upon work partially supported by the DFG Research Center "Mathematics for Key Technologies" (FZT 86) in Berlin, the University of Kansas General Research Fund allocation 2301062-003 and by the National Science Foundation under awards 0098150, 0112375 and 9977352.

This paper reviews and extends the definitions and applications of sum-like and product-like operations on matrix relations that were introduced in [10, 11, 12] in Section 2. The introduced arithmetic-like operations allow to formally add and multiply matrix pencils. This gives a new perspective on several applications involving matrix pencils, or, in more general terms, matrix products and quotients. In Section 3 we consider monodromy relations for linear difference equations, characterize classical solutions of linear differential-algebraic equations with constant coefficients using exponential relations, and derive a generalization of the matrix sign function which is obtainable without matrix inversions. The new generalized sign function can be used to iteratively compute deflating subspaces of matrix pencils. It turns out to be a structure preserving algorithm for calculating deflating subspaces of Hamiltonian and skew-Hamiltonian pencils. The numerical tests in Section 4 report the performance of the new algorithm in comparison to the QZ algorithm [2, 40] and the classical generalized sign function iteration introduced in [24]. Empirically, the new, generalized matrix sign function gives a forward numerically stable method of extracting deflating subspaces.

1.1 Notation and Miscellaneous Definitions

We use the following notation.

- A superscript H indicates the Hermitian or complex conjugate transpose of a matrix, i.e., $E^H = \bar{E}^T$.
- The Moore-Penrose pseudo-inverse of a matrix $E \in \mathbb{C}^{m\times n}$ is denoted by $E^\dagger$. In particular, if $E \in \mathbb{C}^{m\times n}$ has full column rank, then $E^\dagger = (E^H E)^{-1} E^H$.
- The kernel or null space of $M \in \mathbb{C}^{m\times n}$ is $\mathrm{null}(M)$. The range or column space of M is $\mathrm{range}(M)$.
- The spectral norm of a matrix $M \in \mathbb{C}^{m\times n}$ is denoted by $\|M\|_2$. The Frobenius or Euclidean norm on $\mathbb{C}^{m\times n}$ is $\|M\|_F = \sqrt{\mathrm{trace}(M^H M)}$.
- An empty matrix with zero rows and n columns is denoted by $[-] \in \mathbb{C}^{0\times n}$. It is the matrix representation of the trivial linear transformation from $\mathbb{C}^n$ to the zero dimensional vector space $\{0\} = \mathbb{C}^0$.
- A matrix pencil $\lambda E - A$ is *regular* if E and A are square matrices and $\det(\lambda E - A) \neq 0$ for at least one $\lambda \in \mathbb{C}$. A pencil which is not regular is said to be *singular*.
- For a matrix pencil by $\lambda E - A$, a nonzero vector $x \in \mathbb{C}^n$ is an *eigenvector* if for some nonzero pair $(\varepsilon, \alpha) \in \mathbb{C} \setminus \{(0,0)\}$ $\varepsilon E x = \alpha A x$. If $\alpha = 0$, then x corresponds to an infinite eigenvalue. If $\alpha \neq 0$, then x corresponds to the finite eigenvalue $\lambda = \varepsilon/\alpha$.
- The columns of $X \in \mathbb{C}^{n\times k}$ span a a *right deflating subspace* of a regular matrix pencil $\lambda E - A$ if $\dim(\mathrm{range}(X)) = \dim(\mathrm{range}(EX) + \mathrm{range}(AX))$. Deflating subspaces are spanned by collections of eigenvectors and principal vectors. The deflating subspace is associated with the corresponding eigenvalues. If these eigenvalues are disjoint from the remaining eigenvalues of $\lambda E - A$, then the deflating subspace is uniquely determined by them.
- The right deflating subspace $\mathcal{V}_-(\lambda E - A)$ of the regular pencil $\lambda E - A$ corresponding to finite eigenvalues with negative real part is often called the *stable right deflating subspace*. The right deflating subspace corresponding to eigenvalues with positive real part, $\mathcal{V}_+(\lambda E - A)$, is the *unstable deflating subspace*. If $E = I$, we may write $\mathcal{V}_\pm(A)$ for $\mathcal{V}_\pm(\lambda I - A)$. Such deflating subspaces are required, e.g., by numerical algorithms for computing solutions to generalized algebraic Riccati equations and generalized Lyapunov equations [24, 34, 41] or more generally, for solving a variety of computational problems in systems and control [22, 39, 49].

2 Elementary Properties and Arithmetic-like Operations

This section reviews the definitions and elementary mathematical properties of the sum-like and product-like operations on matrix relations that were introduced in [10, 11].

For $x \in \mathbb{C}^n$, the x-section of $(E\backslash A)$ is the set $(E\backslash A)x \equiv \{y \in \mathbb{C}^n \mid (x, y) \in (E\backslash A)\}$. Note that depending on x, E and A, $(E\backslash A)x$ may or may not be empty. The domain of a matrix relation is its set of ordinates, i.e.,

$$\mathrm{Dom}(E\backslash A) = \{x \in \mathbb{C}^n \mid (E\backslash A)x \neq \emptyset\}.$$

The range of a matrix relation is its set of abscissas, i.e.,

$$\mathrm{Range}(E\backslash A) = \bigcup_{x\in\mathbb{C}^n} (E\backslash A)x.$$

Both $\mathrm{Dom}(E\backslash A)$ and $\mathrm{Range}(E\backslash A)$ are vector subspaces of $\mathbb{C}^n$. The matrix E has full column rank if and only if $(E\backslash A)$ is a linear transformation that maps $\mathrm{Dom}(E\backslash A)$ to $\mathrm{Range}(E\backslash A)$. The matrix relation $(E\backslash A)$ is a linear transformation that maps $\mathbb{C}^n$ to $\mathbb{C}^n$ if and only if E has full column rank and $A = EE^\dagger A$. In this case the matrix of the linear transformation is $E^\dagger A$ where $E^\dagger = (E^H E)^{-1} E^H$.

The representation of a matrix relation (1) in terms of matrices E and A is not unique.

Theorem 2.1 *For $M \in \mathbb{C}^{p\times m}$, and E, $A \in \mathbb{C}^{m\times n}$, $(E\backslash A) = (ME\backslash MA)$ if and only if* $\mathrm{null}(M) \cap \mathrm{range}([A,\ -E]) = \{0\}$.

Proof. If $\mathrm{null}(M)\cap\mathrm{range}([A,\ -E]) \neq \{0\}$, then there exists vectors x, $y \in \mathbb{C}^n$ such that $[A,\ -E]\begin{bmatrix} x \\ y \end{bmatrix} \neq 0$ but $M[A,\ -E]\begin{bmatrix} x \\ y \end{bmatrix} = 0$. Hence, $y \in (ME\backslash MA)x$ but $y \notin (E\backslash A)x$. Therefore, $\mathrm{null}(M) \cap \mathrm{range}([A,\ -E]) \neq \{0\}$ implies that $(E\backslash A) \neq (ME\backslash MA)$.

If $\mathrm{null}(M)\cap\mathrm{range}([A,\ -E]) = \{0\}$, then $y \in (ME\backslash MA)x$ implies that $[A,\ -E]\begin{bmatrix} x \\ y \end{bmatrix} \in \mathrm{null}(M)\cap \mathrm{range}([A,\ -E])$. Hence, $[A,\ -E]\begin{bmatrix} x \\ y \end{bmatrix} = 0$, i.e. $y \in (E\backslash A)x$ and $(ME\backslash MA) \subset (E\backslash A)$. If $y \in (E\backslash A)x$, i.e., $Ex = Ay$, then $MEx = MAy$, $y \in (ME\backslash MA)x$ and $(E\backslash A) \subset (ME\backslash MA)$. Therefore, $\mathrm{null}(M) \cap \mathrm{range}([A,\ -E]) = \{0\}$ implies $(E\backslash A) = (ME\backslash MA)$. □

An immediate corollary is the following.

Corollary 2.2 *If E, $A \in \mathbb{C}^{m\times n}$ and $\hat{E}$, $\hat{A} \in \mathbb{C}^{p\times n}$ satisfy $(E\backslash A) = (\hat{E}\backslash\hat{A})$, then there is a matrix $M \in \mathbb{C}^{p\times m}$ such that $\hat{E} = ME$, $\hat{A} = MA$ and* $\mathrm{null}(M) \cap \mathrm{range}([A,\ -E]) = \{0\}$.

The preceding theorem and corollary in particular show that matrix relations are invariant under left-sided linear transformations.

The universal relation $\mathbb{C}^n \times \mathbb{C}^n$ might be written as $([-]\backslash[-])$ where $[-] \in \mathbb{C}^{0\times n}$ is the empty matrix with zero rows and n columns. With this convention, each matrix relation has a representation in which $[E,\ A] \in \mathbb{C}^{m\times(2n)}$ has full row rank m.

2.1 Matrix relation Products

If E_1, $A_1 \in \mathbb{C}^{n\times n}$ and E_2, $A_2 \in \mathbb{C}^{p\times n}$, then the composite or product matrix relation of $(E_2\backslash A_2)$ with $(E_1\backslash A_1)$ is

$$(E_2\backslash A_2)(E_1\backslash A_1) = \left\{ (x, z) \in \mathbb{R}^n \times \mathbb{R}^n \;\middle|\; \begin{array}{l} \text{There exists } y \in \mathbb{C}^n \text{ such that} \\ y \in (E_1\backslash A_1)x \text{ and } z \in (E_2\backslash A_2)y \end{array} \right\} \tag{2}$$

$$= \left\{ (x, z) \in \mathbb{C}^n \times \mathbb{C}^n \;\middle|\; \begin{array}{l} \text{There exists } y \in \mathbb{R}^n \text{ such that} \\ \begin{bmatrix} A_1 & -E_1 & 0 \\ 0 & A_2 & -E_2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array} \right\}. \tag{3}$$

Note that the product relation may or may not have a matrix representation with the same number of rows as the factors. For example, although $([1]\backslash[0])$ and $([0]\backslash[1])$ are matrix relations on $\mathbb{C}^1$ which have representations in terms of 1-by-1 matrices. The product matrix relation, $([1]\backslash[0])([0]\backslash[1]) = \{(0,0)\}$, requires a matrix representation with at least two rows.

It is easy to verify that the product (2) is associative with multiplicative identity $(I\backslash I)$. Only matrix relations that are nonsingular linear transformations on $\mathbb{C}^n$ admit a multiplicative inverse.

The formula for the product of scalar fractions $(a_1/e_1)(a_2/e_2) = (a_1a_2)/(e_1e_2)$ has the following generalization to matrix relations.

Theorem 2.3 *Consider relations $(E_1\backslash A_1)$ and $(E_2\backslash A_2)$ where $E_1,\, A_1 \in \mathbb{C}^{m\times n}$ and $E_2,\, A_2 \in \mathbb{C}^{p\times n}$. If $\tilde{A}_2 \in \mathbb{R}^{q\times m}$ and $\tilde{E}_1 \in \mathbb{R}^{q\times p}$ satisfy*

$$\operatorname{null}[\tilde{A}_2,\, \tilde{E}_1] = \operatorname{range}\begin{bmatrix} -E_1 \\ A_2 \end{bmatrix}, \tag{4}$$

then

$$(E_2\backslash A_2)(E_1\backslash A_1) = ((\tilde{E}_1E_2)\backslash(\tilde{A}_2A_1)) = \left\{(x,z) \in \mathbb{C}^n \times \mathbb{C}^n \;\middle|\; \tilde{E}_1E_2z = \tilde{A}_2A_1x\right\} \tag{5}$$

Proof. See [11]. □

If $E_1 = I$ and $E_2 = I$, then $\tilde{E}_1 = I$ and $\tilde{A}_2 = A_2$ is a possibility in (4) and the theorem reduces to $(I\backslash A_2)(I\backslash A_1) = (I\backslash(A_2A_1))$ which is ordinary matrix multiplication. If $E_2 = I$ and $A_2 = \gamma I$ for some scalar $\gamma \in \mathbb{C}$, then for each pair $(\epsilon, \alpha) \in \mathbb{C}$ for which $\alpha/\epsilon = \gamma$, $\tilde{E}_1 = \epsilon I$ and $\tilde{A}_2 = \alpha I$ is a possibility in (4) and the theorem implies $(I\backslash(\gamma I))(E_1\backslash A_1) = (\epsilon E_1\backslash \alpha A_1)$ which is a special case of scalar multiplication. Thus, Theorem 2.3 is consistent with conventional matrix and scalar multiplication of linear transformations.

For convenience, we define scalar and matrix products with matrix relations as follows. If $M \in \mathbb{C}^{n\times n}$ is a matrix and $(E\backslash A)$ is a relation on $\mathbb{C}^n$, then $(E\backslash A)M = (E\backslash A)(I\backslash M)$ and $M(E\backslash A) = (I\backslash M)(E\backslash A)$. It is easy to show that $(E\backslash A)M = (E\backslash AM)$, and if M is nonsingular, then $M^{-1}(E\backslash A) = (EM\backslash A)$. If $\gamma \in \mathbb{C}$, then we define $\gamma(E\backslash A) = (I\backslash\gamma I)(E\backslash A)$. It is easy to show that $(I\backslash\gamma I)(E\backslash A) = (E\backslash A)(I\backslash\gamma I)$. If $\alpha,\, \beta \in \mathbb{C}$ and $\gamma = \alpha/\beta$, then $\gamma(E\backslash A) = ((\beta E)\backslash\alpha A)$. In particular $\gamma(E\backslash A) = (E\backslash\gamma A)$.

The inverse relation is $(E\backslash A)^{-1} = (A\backslash E)$ and satisfies $y \in (E\backslash A)x$ if and only if $x \in (E\backslash A)^{-1}y$. In general, it is not the case that $(E\backslash A)^{-1}(E\backslash A)$ is the identity relation $(I\backslash I)$. For example, $([1]\backslash[0]) = \{(x,y) \in \mathbb{C}\times\mathbb{C} \mid y = 0\}$ and $([0]\backslash[1]) = \{(x,y) \in \mathbb{C}\times\mathbb{C} \mid x = 0\}$ are inverse relations, but their products are $([1]\backslash[0])([0]\backslash[1]) = \{(0,0)\}$ and $([0]\backslash[1])([1]\backslash[0]) = \mathbb{C}\times\mathbb{C}$. The next theorem summarizes the extent to which the inverse relation acts like a multiplicative inverse.

Lemma 2.4

1. *The inclusion $(I\backslash I) \subset (E\backslash A)^{-1}(E\backslash A)$ holds if and only if $\operatorname{range}(A) \subset \operatorname{range}(E)$ (i.e., Dom$(E\backslash A) = \mathbb{C}^n$).*
2. *The inclusion $(E\backslash A)^{-1}(E\backslash A) \subset (I\backslash I)$ holds if and only if A has full column rank n.*
3. *Both inclusions 1 and 2 hold and $(E\backslash A)^{-1}(E\backslash A) = (I\backslash I)$ if and only if $E^\dagger A$ is nonsingular, i.e., if and only if $(E\backslash A)$ is a nonsingular linear transformation on $\mathbb{C}^n$.*

Proof. To prove inclusion 1, suppose that $\operatorname{range}(A) \subset \operatorname{range}(E)$. For every $x \in \mathbb{C}^n$, there exists $y \in \mathbb{C}^n$ such that $Ey = Ax$ (i.e., $\operatorname{Dom}(E\backslash A) = \mathbb{C}^n$). Trivially, this implies that $x \in (E\backslash A)^{-1}y$. Hence, from (2), for all $x \in \mathbb{C}^n$, $x \in (E\backslash A)^{-1}(E\backslash A)x$ and $(I\backslash I) \subset (E\backslash A)^{-1}(E\backslash A)x$. Conversely, if $(I\backslash I) \subset (E\backslash A)^{-1}(E\backslash A)$, then, in particular, $\mathbb{C}^n = \operatorname{Dom}(E\backslash A)$ which implies $\operatorname{range} A \subset \operatorname{range} E$.

To prove inclusion 2, suppose that $(E\backslash A)^{-1}(E\backslash A) \subset (I\backslash I)$. The only solutions $x,\, y,\, z \in \mathbb{C}^n$ to $Ey = Ax$ and $Az = Ey$ have $x = z$. In particular, for the solution $x = y = z = 0$, the only solution to $Az = 0$ is $z = 0$. This implies that A has full column rank. Conversely, if A has full column rank, then solutions $x,\, y,\, z \in \mathbb{C}^n$ to $Ey = Ax$ and $Az = Ey$ satisfy $Az = Ax$ and, hence, $z = x$. It follows that $(E\backslash A)^{-1}(E\backslash A) \subset (I\backslash I)$.

The prove Statement 3 observe that the inclusion $(I\backslash I) \subset (E\backslash A)^{-1}(E\backslash A)$ implies that range$(A) \subset$ range(E), so Dom$(E\backslash A) = \mathbb{C}^n$. The inclusion $(E\backslash A)^{-1}(E\backslash A) \subset (I\backslash I)$ implies that A has full column rank n. Thus, range(A) has dimension n. This and range$(A) \subset$ range$(E) \subset \mathbb{C}^n$ implies that range$(A) =$ range(E) and E also has full column rank n. Hence, $(E\backslash A)$ is the linear transformation with matrix representation $E^\dagger A$ which is a rank n, n-by-n matrix. □

Some familiar properties of inverses do carry over to matrix relations including

$$((E_2\backslash A_2)(E_1\backslash A_1))^{-1} = (E_1\backslash A_1)^{-1}(E_2\backslash A_2)^{-1}$$

which follows directly from (2).

Products respect common eigenvectors in the following sense.

Theorem 2.5 *Suppose that E_1, E_2, A_1, $A_2 \in \mathbb{C}^{n\times n}$ and the matrix pencils $\lambda E_1 - A_1$ and $\lambda E_2 - A_2$ have a mutual eigenvector $x \neq 0$, i.e., suppose that there exist nonzero ordered pairs (ϵ_1, α_1), $(\epsilon_2, \alpha_2) \in \mathbb{C} \times \mathbb{C} \setminus \{(0,0)\}$ such that*

$$\begin{aligned} \epsilon_1 E_1 x &= \alpha_1 A_1 x && (6) \\ \epsilon_2 E_2 x &= \alpha_2 A_2 x. && (7) \end{aligned}$$

If $(E\backslash A) = (E_2\backslash A_2)(E_1\backslash A_1)$, then $(\epsilon_2\epsilon_1)Ex = (\alpha_2\alpha_1)Ax$. Moreover,

1. *If $(\epsilon_2\epsilon_1, \alpha_2\alpha_1) \neq (0,0)$, then x is an eigenvector of $\lambda E - A$.*
2. *If $\epsilon_1 = \alpha_2 = 0$, then $Ex = Ax = 0$ and the pencil $\lambda E - A$ is singular.*
3. *If $\epsilon_2 = \alpha_1 = 0$, and* rank $\begin{bmatrix} A_1 & -E_1 & 0 \\ 0 & A_2 & -E_2 \end{bmatrix} = 2n$, *then E and A are m-by-n matrices with $m > n$ and the pencil $\lambda E - A$ is singular.*

Proof. Multiply (6) by α_2 and (7) by ϵ_1 to get

$$\begin{aligned} E_1(\alpha_2\epsilon_1 x) &= A_1(\alpha_2\alpha_1 x) \\ E_2(\epsilon_2\epsilon_1 x) &= A_2(\epsilon_1\alpha_2 x) \end{aligned}$$

Hence, $(\epsilon_2\epsilon_1 x) \in (E_2\backslash A_2)(E_1\backslash A_1)(\alpha_2\alpha_1 x)$ and $(\epsilon_2\epsilon_1)Ex = (\alpha_2\alpha_1)Ax$. If $(\epsilon_2\epsilon_1, \alpha_2\alpha_1) \neq (0,0)$, then x is an eigenvector of $\lambda E - A$.

By hypothesis $(\epsilon_1, \alpha_1) \neq (0,0)$ and $(\epsilon_2, \alpha_2) \neq (0,0)$, so the condition $(\epsilon_1\epsilon_2, \alpha_1\alpha_2) = (0,0)$ implies that either $\epsilon_1 = \alpha_2 = 0$ or $\epsilon_2 = \alpha_1 = 0$. If $\epsilon_1 = \alpha_2 = 0$, then $0 \in (E_1\backslash A_1)x$. Since $0 \in (E_2\backslash A_2)0$, it follows that $0 \in (E_2\backslash A_2)(E_1\backslash A_1)x = (E\backslash A)x$, i.e., $Ax = 0$. Similarly, $x \in (E_2\backslash A_2)0$ and $0 \in (E_1\backslash A_1)0$, so $x \in (E_2\backslash A_2)(E_1\backslash A_1)0 = (E\backslash A)0$, i.e., $Ex = 0$. Because $x \neq 0$ is a mutual null vector of E and A, the matrix pencil $\lambda E - A$ is singular.

If $\epsilon_2 = \alpha_1 = 0$, then $E_1 x = 0$ and $A_2 x = 0$ and $[0, x^H, 0]^H \in \mathbb{C}^{3n}$ satisfies

$$\begin{bmatrix} A_1 & -E_1 & 0 \\ 0 & A_2 & -E_2 \end{bmatrix} \begin{bmatrix} 0 \\ x \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

By hypothesis, rank $\begin{bmatrix} A_1 & -E_1 & 0 \\ 0 & A_2 & -E_2 \end{bmatrix} = 2n$. Expand $[0, x^H, 0]^H$ to a basis of null $\begin{bmatrix} A_1 & -E_1 & 0 \\ 0 & A_2 & -E_2 \end{bmatrix}$, $\{[0, x^H, 0]$, $[u_2^H, v_2^H, w_2^H]^H$, $[u_3^H, v_3^H, w_3^H]^H$, $\ldots [u_n^H, v_n^H, w_n^H]^H$. Regarding $(E\backslash A)$ as the null space of $[-E, A] \in \mathbb{C}^{n\times 2n}$, we have from (3) that $\{[0, 0]$, $[u_2^H, w_2^H]^H$, $[u_3^H, w_3^H]^H$, $\ldots [u_n^H, w_n^H]^H$ is a spanning set of $(E\backslash A)$. Consequently, $(E\backslash A)$ has dimension less than n which implies that rank$[-E, A] > n$. Hence, E and A must be m-by-n matrices with $m > n$. Such rectangular pencils $\lambda E - A$ can not be regular. □

The rank hypothesis in Theorem 2.5 part 3 is relatively mild. It is satisfied, for example, whenever A_1 and E_2 are nonsingular. Some such hypothesis is needed in order to conclude that $\lambda E - A$ is singular. For example, $([1]\backslash[0])$ and $([0]\backslash[0])$ have a mutual eigenvector $x = [1]$, but $([1]\backslash[0])([0]\backslash[0]) = ([1]\backslash[0])$ is represented by the regular pencil $\lambda[1] - [0]$. We conjecture that the rank hypothesis in Statement 3 can be weakened to assuming that both $\lambda E_1 - A_1$ and $\lambda E_2 - A_2$ are regular.

Theorem 2.5 generalizes to deflating subspaces.

Theorem 2.6 *Let $(E_1\backslash A_1)$ and $(E_2\backslash A_2)$ be matrix relations on $\mathbb{C}^n$ and let $(E\backslash A)$ be the product $(E\backslash A) = (E_2\backslash A_2)(E_1\backslash A_1)$. Suppose that $X \in \mathbb{C}^{n\times k}$, and S_1, T_1, S_2, $T_2 \in \mathbb{C}^{k\times p}$ satisfy*

$$\begin{aligned} E_1XS_1 &= A_1XT_1 && (8)\\ E_2XS_2 &= A_2XT_2. && (9)\end{aligned}$$

If $\tilde{S}_1$, $\tilde{T}_2$ satisfy

$$\mathrm{null}[S_1,\, T_2] = \mathrm{range}\begin{bmatrix} -\tilde{T}_2 \\ \tilde{S}_1 \end{bmatrix}, \tag{10}$$

then

$$EX(S_2\tilde{S}_1) = AX(T_1\tilde{T}_2).$$

Moreover, if $\lambda(S_2\tilde{S}_1) - (T_1\tilde{T}_2)$ is regular, then range(X) *is a (right) deflating subspace of $\lambda E - A$.*

Proof. If $\tilde{E}_1$ and $\tilde{A}_2$ satisfy (4), then by Theorem 2.3 we may use $E = \tilde{E}_1E_2$ and $A = \tilde{A}_2A_1$ to represent the product $(E\backslash A) = (E_2\backslash A_2)(E_1\backslash A_1)$. Equations (8), (9) and (10) imply that

$$AXT_1\tilde{T}_2 = \tilde{A}_2A_1XT_1\tilde{T}_2 = \tilde{A}_2E_1XS_1\tilde{T}_2 = \tilde{E}_1A_2XT_2\tilde{S}_1 = \tilde{E}_1E_2XS_2\tilde{S}_1 = EXS_2\tilde{S}_1. \tag{11}$$

This proves the theorem for $E = \tilde{E}_1E_2$ and $A = \tilde{A}_2A_1$. It remains to show the theorem for the other pairs of matrices $\hat{E}$, $\hat{A} \in \mathbb{C}^{p\times n}$ such that $(\hat{E}\backslash\hat{A}) = (E\backslash A)$. If $(\hat{E}\backslash\hat{A}) = (E\backslash A)$, then, by Corollary 2.2, there is a matrix $M \in \mathbb{C}^{p\times m}$ such that $\hat{E} = ME$ and $\hat{A} = MA$ It follows from (11) that $\hat{E}$ and $\hat{A}$ also satisfy $\hat{E}X(S_2\tilde{S}_1) = \hat{A}X(T_1\tilde{T}_2)$. ◻

Remark 2.7 If both S_1 and S_2 are nonsingular or both T_1 and T_2 are nonsingular, then $\lambda(S_2\tilde{S}_1) - T_1\tilde{T}_2$ can be chosen to be regular. However, it is possible for $\lambda(S_2\tilde{S}_1) - (T_1\tilde{T}_2)$ to be a singular pencil even when both $\lambda S_1 - T_1$ and $\lambda S_2 - T_2$ are regular. A 1-by-1 example is $S_1 = [1]$, $T_1 = [0]$, $S_2 = [0]$, $T_2 = [1]$ in which case $S_2\tilde{S}_1 = [0]$ and $T_1\tilde{T}_2 = [0]$. Necessary and sufficient conditions for $\lambda S - T$ to be regular remain an open problem.

2.2 Matrix Relation Sums

The sum of $(E_1\backslash A_1)$ with $(E_2\backslash A_2)$ is the matrix relation

$$(E_1\backslash A_1) + (E_2\backslash A_2) = \left\{ (x,z) \in \mathbb{C}^n \times \mathbb{C}^n \;\middle|\; \begin{array}{l} \text{There exists } y_1,\, y_2 \in \mathbb{C}^n \text{, such that} \\ y_1 \in (E_1\backslash A_1)x,\; y_2 \in (E_2\backslash A_2)x \text{ and} \\ z = y_1 + y_2 \end{array} \right\} \tag{12}$$

or, equivalently,

$$(E_1\backslash A_1) + (E_2\backslash A_2) = \left\{ (x,z) \in \mathbb{C}^n \times \mathbb{C}^n \;\middle|\; \begin{array}{l} \text{There exists } y_1,\, y_2 \in \mathbb{C}^n \text{, such that} \\ \begin{bmatrix} A_1 & -E_1 & 0 & 0 \\ A_2 & 0 & -E_2 & 0 \\ 0 & I & I & -I \end{bmatrix} \begin{bmatrix} x \\ y_1 \\ y_2 \\ z \end{bmatrix} = 0. \end{array} \right\}$$

Here again, a representation of the sum relation may require matrices with a different number of rows than the matrix representations of the summands.

The matrix relation $(I\backslash 0)$ is an additive identity. A matrix relation has an additive inverse if and only if it is a linear transformation on $\mathbb{C}^n$.

The formula for the sum of scalar fractions $a_1/e_1 + a_2/e_2 = (e_2a_1 + e_1a_2)/(e_1e_2)$ has the following generalization to matrix relations.

Theorem 2.8 *Consider matrix relations $(E_1\backslash A_1)$ and $(E_2\backslash A_2)$ with E_1, $A_1 \in \mathbb{C}^{m\times n}$ and E_2, $A_2 \in \mathbb{C}^{p\times n}$. If $\tilde{E}_2 \in \mathbb{C}^{q\times m}$, and $\tilde{E}_1 \in \mathbb{C}^{q\times p}$ satisfy*

$$\mathrm{null}[\tilde{E}_2,\, \tilde{E}_1] = \mathrm{range}\begin{bmatrix} -E_1 \\ E_2 \end{bmatrix}, \tag{13}$$

then

$$(E_2\backslash A_2) + (E_1\backslash A_1) = ((\tilde{E}_1 E_2)\backslash(\tilde{E}_2 A_1 + \tilde{E}_1 A_2)) = \left\{(x,z) \in \mathbb{C}^n \times \mathbb{C}^n \;\middle|\; \tilde{E}_1 E_2 z = (\tilde{E}_2 A_1 + \tilde{E}_1 A_2)x\right\}. \quad (14)$$

Proof. See [11]. □

Observe that (13) implies that

$$\tilde{E}_1 E_2 = \tilde{E}_2 E_1, \quad (15)$$

so $\tilde{E}_1$, $\tilde{E}_2$, E_1 and E_2 are symmetrical in (14). If $E_1 = E_2 = I$, then $\tilde{E}_1 = \tilde{E}_2 = I$ is a possibility in (13). This choice gives conventional matrix addition $(I\backslash A_1) + (I\backslash A_2) = (I\backslash(A_1 + A_2))$.

Sum relations respect common eigenvectors in the following sense.

Theorem 2.9 *Suppose that $\lambda E_1 - A_1$ and $\lambda E_2 - A_2$ are matrix pencils that have a mutual eigenvector $x \neq 0$ and*

$$\epsilon_1 E_1 x = \alpha_1 A_1 x \quad (16)$$

$$\epsilon_2 E_2 x = \alpha_2 A_2 x \quad (17)$$

for some nonzero pairs (ϵ_1, α_1), $(\epsilon_2, \alpha_2) \in \mathbb{C} \times \mathbb{C} \setminus \{(0,0)\}$. Let $(E\backslash A) = ((E_2\backslash A_2)) + ((E_1\backslash A_1))$.

1. *If either $\alpha_1 \neq 0$ or $\alpha_2 \neq 0$, then $(\epsilon_2\alpha_1 + \epsilon_1\alpha_2, \alpha_1\alpha_2) \neq (0,0)$ and $(\alpha_1\epsilon_2 + \epsilon_1\alpha_2)Ex = \alpha_1\alpha_2 Ax$.*
2. *If $\alpha_1 = \alpha_2 = 0$, then $1(Ex) = 0(Ax)$.*

In any case, x is an eigenvector of $\lambda E - A$.

Proof.

If $\alpha_1 \neq 0$ or $\alpha_2 \neq 0$, then multiply (16) by α_2 and multiply (17) by α_1 to get

$$\epsilon_1\alpha_2 x = (E_1\backslash A_1)(\alpha_1\alpha_2 x)$$

$$\alpha_1\epsilon_2 x = (E_2\backslash A_2)(\alpha_1\alpha_2 x).$$

Hence, $(\epsilon_1\alpha_2 + \alpha_1\epsilon_2)x \in ((E_1\backslash A_1) + (E_2\backslash A_2))(\alpha_1\alpha_2 x)$. Under the assumption that $(\epsilon_1, \alpha_1) \neq (0,0)$ and $(\epsilon_2, \alpha_2) \neq (0,0)$ it is easy to show that if either $\alpha_1 \neq 0$ or $\alpha_2 \neq 0$, then $(\epsilon_2\alpha_1 + \epsilon_1\alpha_2, \alpha_1\alpha_2) \neq (0,0)$. So, $(\alpha_1\alpha_2)Ex = (\epsilon_1\alpha_2 + \alpha_1\epsilon_2)Ax$ and x is an eigenvector of $\lambda E - A$.

If $\alpha_1 = \alpha_2 = 0$, then $E_1 x = 0$ and $E_2 x = 0$. So, $x \in (E_1\backslash A_1)0$ and $x \in (E_2\backslash A_2)0$ which shows that $x + x \in ((E_1\backslash A_1) + (E_2\backslash A_2))\,0$. It follows that $Ex = 0$ and $1(Ex) = 0(Ax)$, so x is an eigenvector of $\lambda E - A$. □

Theorem 2.9 generalizes to deflating subspaces.

Theorem 2.10 *Let $(E_1\backslash A_1)$ and $(E_2\backslash A_2)$ be matrix relations on $\mathbb{C}^n$ and let $(E\backslash A)$ the be sum $(E\backslash A) = (E_1\backslash A_1) + (E_2\backslash A_2)$. Suppose that $X \in \mathbb{C}^{n\times k}$ and S_1, T_1, S_2, $T_2 \in \mathbb{C}^{k\times p}$ satisfy*

$$E_1 X S_1 = A_1 X T_1 \quad (18)$$

$$E_2 X S_2 = A_2 X T_2. \quad (19)$$

If $\tilde{T}_1$ and $\tilde{T}_2$ satisfy

$$\mathrm{null}[-T_1, T_2] = \mathrm{range}\begin{bmatrix} \tilde{T}_2 \\ \tilde{T}_1 \end{bmatrix}, \quad (20)$$

then

$$EX(S_1\tilde{T}_2 + S_2\tilde{T}_1) = AX(T_1\tilde{T}_2). \quad (21)$$

Moreover, if $\lambda(T_1\tilde{T}_2) - (S_1\tilde{T}_2 + S_2\tilde{T}_1)$ is regular, then range(X) *is a right deflating subspace of $\lambda E - A$.*

Proof. If $\tilde{E}_1$ and $\tilde{E}_2$ satisfy (13), i.e., if $\tilde{E}_2 E_1 = \tilde{E}_1 E_2$, then (18), (19) and (20) imply that

$$\begin{aligned} \tilde{E}_2 E_1 X S_1 \tilde{T}_2 &= \tilde{E}_2 A_1 X T_1 \tilde{T}_2 \\ \tilde{E}_1 E_2 X S_2 \tilde{T}_1 &= \tilde{E}_1 A_2 X T_2 \tilde{T}_1. \end{aligned}$$

Adding the two equations and using $\tilde{E}_1 E_2 = \tilde{E}_2 E_1$ and $T_1 \tilde{T}_2 = T_2 \tilde{T}_1$ gives

$$(\tilde{E}_1 E_2) X (S_1 \tilde{T}_2 + S_2 \tilde{T}_1) = (\tilde{E}_2 A_1 + \tilde{E}_1 A_2) X (T_1 \tilde{T}_2).$$

Equation (21) follows from Theorem 2.8. ◻

In the context of Theorem 2.9, if T_1 and T_2 are square and nonsingular, then $\tilde{T}_1$ and $\tilde{T}_2$ can be chosen to be square and nonsingular and, consequently the pencil $\lambda(T_1\tilde{T}_2) - (S_1\tilde{T}_2 + S_2\tilde{T}_1)$ is regular.

Remark 2.11 The classic proof of Ore's theorem [29, p. 170], characterizing rings having left quotient rings (or being a left order in a ring), uses expressions (5) and (14) in an abstract setting to define the multiplication and addition in the left quotient ring.

2.3 Distribution of Addition across Multiplication

In general, the distributive law of multiplication across addition does not hold. For example, if $\mathcal{R}_3 = ([0]\backslash[1])$, $\mathcal{R}_1 = ([1]\backslash[1])$ and $\mathcal{R}_2 = ([1]\backslash[-1])$,then $(\mathcal{R}_3\mathcal{R}_1) + (\mathcal{R}_3\mathcal{R}_2) = \mathcal{R}_3 = \{0\} \times \mathbb{C}$ but $\mathcal{R}_3(\mathcal{R}_1 + \mathcal{R}_2) = \mathbb{C} \times \mathbb{C}$. However, there is a partial distributive law.

Theorem 2.12 *For any three matrix relations on $\mathbb{C}^n$, $\mathcal{R}_1 = (E_1\backslash A_1)$, $\mathcal{R}_2 = (E_2\backslash A_2)$, and $\mathcal{R}_3 = (E_3\backslash A_3)$:*

$$(\mathcal{R}_3\mathcal{R}_1) + (\mathcal{R}_3\mathcal{R}_2) \subset \mathcal{R}_3(\mathcal{R}_1 + \mathcal{R}_2). \tag{22}$$

Proof. If $z \in ((\mathcal{R}_3\mathcal{R}_1) + (\mathcal{R}_3\mathcal{R}_1))\,x$, then there exist vectors $y_1 \in (E_3\backslash A_3)(E_1\backslash A_1)x$ and $y_2 \in (E_3\backslash A_3)(E_2\backslash A_2)x$ such that $z = y_1 + y_2$. This in turn implies that there exist vectors w_1, $w_2 \in \mathbb{C}^n$ such that

$$\begin{aligned} E_1 w_1 &= A_1 x \\ E_3 y_1 &= A_3 w_1 \\ E_2 w_2 &= A_2 x \\ E_3 y_2 &= A_3 w_2 \\ z &= y_1 + y_2. \end{aligned}$$

This implies that $E_3 z = E_3(y_1 + y_2) = A_3(w_1 + w_2)$ and $w_1 + w_2 \in ((E_1\backslash A_1) + (E_2\backslash A_2))\,x$. Hence,

$$z \in (E_3\backslash A_3)\,((E_1\backslash A_1) + (E_2\backslash A_3)) = \mathcal{R}_3(\mathcal{R}_1 + \mathcal{R}_2).$$

◻

The following theorem shows that the distributive law does hold in some special cases.

Theorem 2.13 *Let $\mathcal{R}_1 = (E_1\backslash A_1)$, $\mathcal{R}_2 = (E_2\backslash A_2)$, and $\mathcal{R}_3 = (E_3\backslash A_3)$ be three matrix relations. If* $\mathrm{range}(E_1\backslash A_1) \subset \mathit{Dom}(E_3\backslash A_3)$ *and* $\mathrm{range}(E_2\backslash A_2) \subset \mathit{Dom}(E_3\backslash A_3)$*, then*

$$\mathcal{R}_3(\mathcal{R}_1 + \mathcal{R}_2) \subset (\mathcal{R}_3\mathcal{R}_1) + (\mathcal{R}_3\mathcal{R}_2) \tag{23}$$

In particular, if $\mathcal{R}_3$ is a linear transformation on $\mathbb{C}^n$, then $\mathcal{R}_3(\mathcal{R}_1 + \mathcal{R}_2) = (\mathcal{R}_3\mathcal{R}_1) + (\mathcal{R}_3\mathcal{R}_2)$.

Proof. If $z \in (E_3\backslash A_3)\,((E_1\backslash A_1) + (E_2\backslash A_2))\,x$, then there exist vectors y_1, $y_2 \in \mathbb{C}^n$ such that

$$\begin{aligned} E_1 y_1 &= A_1 x \\ E_2 y_2 &= A_2 x \\ E_3 z &= A_3(y_1 + y_2) \end{aligned}$$

If range$(E_1\backslash A_1) \subset$ Dom$(E_3\backslash A_3)$ and range$(E_2\backslash A_2) \subset$ Dom$(E_3\backslash A_3)$, then there exist vectors z_1, $z_2 \in \mathbb{C}^n$ such that

$$\begin{aligned} E_3 z_1 &= A_3 y_1 \\ E_3 z_2 &= A_3 y_2. \end{aligned}$$

This implies that $E_3 z = E_3(z_1 + z_2) = A_3(y_1 + y_2)$. Set $w = z - (z_1 + z_2)$ and note that $E_3 w = 0$. Since $E_3(z_1 + w) = E_3 z_1 = A_3 y_1$, it follows that $z_1 + w \in (E_3\backslash A_3)(E_1\backslash A_1)x$ and

$$z = (z_1 + w + z_2) \in \left((E_3\backslash A_3)(E_1\backslash A_1) + (E_3\backslash A_3)(E_2\backslash A_2)\right)x.$$

Therefore, (23) holds. $\square$

Similar inclusions hold for distribution from the right.

Theorem 2.14 *For any three matrix relations on $\mathbb{C}^n$, $\mathcal{R}_1 = (E_1\backslash A_1)$, $\mathcal{R}_2 = (E_2\backslash A_2)$ and $\mathcal{R}_3 = (E_3\backslash A_3)$:*

$$(\mathcal{R}_1 + \mathcal{R}_2)\mathcal{R}_3 \subset (\mathcal{R}_1\mathcal{R}_3) + (\mathcal{R}_2\mathcal{R}_3). \tag{24}$$

If $\mathcal{R}_3$ is a linear transformation on $\mathbb{C}^n$, then $(\mathcal{R}_1\mathcal{R}_3) + (\mathcal{R}_2\mathcal{R}_3) = (\mathcal{R}_1 + \mathcal{R}_2)\mathcal{R}_3$.

Proof. Similar to the proofs of Theorems 2.12 and 2.13. $\square$

Comparing (22) and (24), we see that the inclusion flips between distribution from the left and distribution from the right.

2.4 Polynomials of Matrix Relations

If $p(x)$ is the polynomial $p(x) = \sum_{i=0}^{d} a_i x^i$, then for any matrix relation $(E\backslash A)$, we may define $p((E\backslash A))$ as $p\left((E\backslash A)\right) = \sum_{i=0}^{d} a_i (E\backslash A)^i$ where $(E\backslash A)^0$ is defined to be $(I\backslash I)$ and all other sums and products of relations are as defined above. The following lemmas enable occasional use of canonical forms to analyze matrix relations which will be necessary in Section 3 to analyze classical solution of linear differential-algebraic equations.

Lemma 2.15 *If $X, Y \in \mathbb{C}^{n\times n}$ are nonsingular and $(E\backslash A)$ is a matrix relation on $\mathbb{C}^n$, then $Y^{-1} p\left((E\backslash A)\right) Y = p\left(Y^{-1}(E\backslash A)Y\right) = p\left((XEY\backslash XAY)\right)$.*

Proof. By (2), $z \in Y^{-1}(E\backslash A)^i Y x$ if and only if there exist vectors x_0, x_1, …, x_i satisfying $x_0 = Yx$, $Ex_j = Ax_{j-1}$, for $j = 1, 2, \ldots, i$, and $Yz = x_i$. With $\tilde{x}_i = Y^{-1}x_i$ for $j = 0, 1, 2, \ldots, i$, this becomes $\tilde{x}_0 = x$, $EY\tilde{x}_j = AY\tilde{x}_{j-1}$ for $j = 1, 2, \ldots, i$, and $z = \tilde{x}_i$. which shows that $z \in (EY\backslash AY)^i x$. A similar argument shows that $z \in (EY\backslash AY)^i x$ implies that $z \in Y^{-1}(E\backslash A)^i Y x$. Hence, $Y^{-1}(E\backslash A)^i Y = (EY\backslash AY)^i$. Theorem 2.1 now implies $Y^{-1}(E\backslash A)^i Y = (XEY\backslash XAY)^i$.

Let $p(x) = \sum_{j=1}^{d} a_1 x^i$. Regarding Y and Y^{-1} as linear transformations on $\mathbb{C}^n$, Theorems 2.13 and 2.14 imply that the distributive law holds for Y and Y^{-1} and

$$Y^{-1}p((E\backslash A))Y = \sum_{j=1}^{d} a_1 Y^{-1}(E\backslash A)^i Y = \sum_{j=1}^{d} a_1 (EY\backslash AY)^i = \sum_{j=1}^{d} a_1 (XEY\backslash XAY)^i.$$

The last inequality follows from Theorem 2.1 and the fact that X is nonsingular. $\square$

Polynomials also respect block diagonal structure.

Lemma 2.16 *Suppose E_1, $A_1 \in \mathbb{C}^{m_1\times n_1}$, E_2, $A_2 \in \mathbb{C}^{m_2\times n_2}$ and $p(x)$ is a polynomial. If $(\hat{E}_1\backslash \hat{A}_1) = p(E_1\backslash A_1)$ and $(\hat{E}_2\backslash \hat{A}_2) = p(E_2\backslash A_2)$, then*

$$p(\operatorname{diag}(E_1, E_2)\backslash \operatorname{diag}(A_1, A_2)) = (\operatorname{diag}(\hat{E}_1, \hat{E}_2)\backslash \operatorname{diag}(\hat{A}_1, \hat{A}_2)).$$

3 Applications

3.1 Implicit Products

Building on [26, 38] the inverse free, spectral divide and conquer (IFSDC) algorithm [7] calculates an invariant subspace of a matrix A as range $\left(\lim_{k\to\infty}(I + A^k)^{-1}\right)$ or a right deflating subspace of a pencil $\lambda E - A$ as range $\left(\lim_{k\to\infty}(I + (E^{-1}A)^k)^{-1}\right)$. The calculation is carefully organized to avoid numerical instabilities. In particular, it represents the power A^{2^k} or $(E^{-1}A)^{2^k}$ in terms of a pair of matrices E_k, A_k as $(E_k^{-1}A_k) = A^{2^k}$ or $E_k^{-1}A_k = (E^{-1}A)^{2k}$. The matrices E_{k+1} and A_{k+1} are calculated from E_k and A_k using what is essentially Theorem 2.3 and (38). In the language of this paper, the IFSDC algorithm calculates $(E_k\backslash A_k) = (I\backslash A)^{2k}$ or $(E_k\backslash A_k) = (E\backslash A)^{2k}$ by successive matrix relation squaring $(E_{k+1}\backslash A_{k+1}) = (E_k\backslash A_k)^2$.

Consider the linear, discrete-time descriptor system

$$E_k x_{k+1} = A_k x_k \quad k = 1, 2, 3, \ldots \tag{25}$$

where E_k, $A_k \in \mathbb{C}^{m\times n}$, and $x_k \in \mathbb{C}^n$. If for all k, range$(A_k) \subset$ range(E_k) and E_k has full column rank n, then a sequence x_k satisfies (25) if and only if for all $k_1 > k_0$

$$x_{k_1+1} = \left(\prod_{k=k_0}^{k_1} E_k^\dagger A_k\right) x_{k_0}. \tag{26}$$

Moreover, for $k_1 > k_0$, a pair of vectors (x_{k_0}, x_{k_1+1}) are the k_0 and $k_1 + 1$st term in a sequence x_k of solutions of (25) if and only if (26) holds. The product $\prod_{k=k_0}^{k_1} E_k^\dagger A_k$ is sometimes called the (k_1, k_0) *monodromy matrix*.

A generalization of (26) to the case in which some or all of the E_k's fail to have full column rank is mentioned in [11, 12]. A sequence x_k satisfies (25) if and only if for all $k_1 > k_0$

$$x_{k_1+1} \in \left(\prod_{k=k_0}^{k_1} (E_k\backslash A_k)\right) x_{k_0}. \tag{27}$$

Moreover, for any pair of integers $k_1 > k_0$, a pair of vectors (x_{k_0}, x_{k_1+1}) are the k_0 and $k_1 + 1$st term in a sequence x_k of solutions of (25) if and only if (27) holds.

The matrix relation $(E_{k_1:k_0}\backslash A_{k_1:k_0}) = \prod_{k=k_0}^{k_1}(E_k\backslash A_k)$ might be called the (k_1, k_0) monodromy matrix relation. Theorem 2.3 suggests a way to explicitly compute E_{k_1,k_0} and A_{k_1,k_0} by computing a sequence of bases of null spaces and matrix products. For this purpose, algorithms proposed in [12] can be adapted. It is noted in [12] that if $E_{k_0}^\dagger A_{k_0}$ is nonsingular (i.e., if $(E_{k_0}\backslash A_{k_0})$ is a nonsingular linear transformation), then

$$\begin{aligned}
(E_{k_1:(k_0+1)}\backslash A_{k_1:(k_0+1)}) &= (E_{k_1:k_0}\backslash A_{k_1:k_0})(A_{k_0}\backslash E_{k_0}) \\
(E_{(k_1+1):(k_0+1)}\backslash A_{(k_1+1):(k_0+1)}) &= (E_{k_1+1}\backslash A_{k_1+1})(E_{k_1:k_0}\backslash A_{k_1:k_0})(A_{k_0}\backslash E_{k_0}).
\end{aligned}$$

In this case, some monodromy relations can be obtained from others at the cost of relatively few matrix relation products. Particularly, this approach allows to obtain deflating subspaces, eigenvalues and -vectors, as well as singular values and singular vectors of formal matrix products $\Pi_{k=1}^{\ell} A^{s_k}$ where $s_k = \pm 1$. In that way, the methods from [12, 21] generalize to formal matrix products involving rectangular factors.

3.2 Continuous-Time Descriptor Systems

The following generalization of the matrix exponential suitable for use with descriptor systems is mentioned in [11]. Consider the linear, time invariant, differential algebraic equation

$$E\dot{x} = Ax \tag{28}$$

where $E, A \in \mathbb{C}^{m\times n}$ and $x = x(t) : \mathbb{R} \to \mathbb{C}^n$ is a classical, smooth solution. (The notation $\dot{x}$ indicates the time derivative dx/dt.) This differential algebraic equation is well studied from both the theoretical and computational viewpoints. See, for example, [17, 33].

If $\mathrm{range}(A) \subset \mathrm{range}(E)$ and E has full column rank n, then classical solutions of (28) are characterized by the property that for all $t_0, t_1 \in \mathbb{R}^n$,

$$x(t_1) = \exp(E^\dagger A(t_1 - t_0))x(t_0). \tag{29}$$

Moreover, $x_1 = \exp(E^\dagger A(t_1 - t_0))x_0$ if and only if there is a classical solution $x(t)$ to (28) which interpolates x_1 and x_0 at t_1 and t_0.

If $\mathrm{range}(A) \not\subset \mathrm{range}(E)$ or E does not have full column rank, then the situation becomes more complicated. However, when expressed in terms of matrix relations, it is only a little more complicated. Define the exponential relation by

$$\exp(E\backslash(A(t_1 - t_0))) = \sum_{k=0}^{\infty} \frac{(t_1 - t_0)^k}{k!}((E\backslash A))^k \tag{30}$$

where the terms in the sum are interpreted as in Subsection 2.4. As defined above, the infinite sum is a limit of matrix relations, i.e., a limit of subspaces of $\mathbb{C}^n \times \mathbb{C}^n$ in the usual largest-canonical-angle/gap metric topology [28], [44, Ch.II§4].

Theorem 3.1 *For all $E, A \in \mathbb{C}^{m\times n}$, the exponential relation (30) is well defined and converges.*

Proof. See Appendix A for a more detailed statement of the theorem and a proof. □

In many ways, the matrix relation exponential characterizes solutions to (28).

Theorem 3.2 *If $\lambda E - A$ is a regular pencil on $\mathbb{C}^{n\times n}$, then $x(t)$ is a classical solution of (28) if and only if for all $t_0, t_1 \in \mathbb{R}$,*

$$x(t_1) \in \exp(E\backslash(A(t_1 - t_0)))x(t_0).$$

Proof. By hypothesis, $\lambda E - A$ is regular, so it has Weierstraß canonical form

$$X(\lambda E - A)Y = \lambda \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} - \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \tag{31}$$

where $X, Y \in \mathbb{C}^{n\times n}$ are nonsingular, $J \in \mathbb{C}^{k\times k}$ is in Jordan Form, and $N \in \mathbb{C}^{(n-k)\times(n-k)}$ is a nilpotent matrix also in Jordan form [23, Vol.II,§2].

Partition $z(t) = Y^{-1}x(t)$ conformally with (31) as $z(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$ with $z_1(t) \in \mathbb{C}^k$ and $z_2(t) \in \mathbb{C}^{n-k}$. Then $x(t)$ is a classical solution of (28) if and only if for any $t_0, t_1 \in \mathbb{R}$, $z_1(t_1) = e^{J(t_1-t_0)}z_1(t_0)$ and $z_2(t) \equiv 0$.

It is easy to verify (see Appendix A) that

$$x(t_1) \in \exp\left(E\backslash(A(t_1 - t_0))\right)x(t_0)$$

if and only if

$$\begin{aligned} z(t_1) &\in \exp\left(\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} \backslash \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix}(t_1 - t_0)\right) z(t_0) \\ &= \left(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \backslash \begin{bmatrix} e^{J(t_1-t_0)} & 0 \\ 0 & I \end{bmatrix}\right) z(t_0). \end{aligned}$$

The last equality is a tedious but straight forward application of (30), see Appendix A. So,

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1(t_1) \\ z_2(t_1) \end{bmatrix} = \begin{bmatrix} e^{J(t_1-t_0)} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} z_1(t_0) \\ z_2(t_0) \end{bmatrix}.$$

Hence, for all t_0, $t_1 \in \mathbb{R}$, $z_1(t_1) = e^{J(t_1-t_0)} z_1(t_0)$ and $z_2(t_0) = 0$. (Note that $z_2 \equiv 0$ because t_0 varies throughout $\mathbb{R}$.) $\square$

As emphasized at the end of the proof, Theorem 3.2 characterizes solutions to (28) as both t_0 and t_1 vary through $\mathbb{R}$. In contrast, (29) still characterizes solutions if t_0 is fixed *a priori* and only t_1 varies.

A corollary to the proof of Theorem 3.2 is useful for numerical computation.

Corollary 3.3 *Suppose that $\lambda E - A$ is a regular pencil on $\mathbb{C}^{n\times n}$ and x_0, $x_1 \in \mathbb{C}^n$. There exists a classical solution $x(t)$ of (28) such that $x(t_0) = x_0$ and $x(t_1) = x_1$ if and only if*

$$\begin{aligned} x(t_1) &\in \exp(E\backslash(A(t_1-t_0)))x(t_0) \\ x(t_0) &\in \exp(E\backslash(A(t_0-t_1)))x(t_1). \end{aligned}$$

Theorem 3.2 has an extension to singular pencils. Let $\lambda E - A$ have Kronecker canonical form (Theorem A.1 in Appendix A) [23],

$$X(\lambda E - A)Y = \mathrm{diag}(\lambda E_0 - A_0,\, L_1,\, L_2,\, \ldots,\, L_p,\, L_{p+1}^T,\, L_{p+2}^T,\, \ldots,\, L_{p+q}^T) \tag{32}$$

where X and Y are nonsingular, $\lambda E_0 - A_0$ is regular and the L_j's are ϵ_j-by-(ϵ_j+1) matrices of the form

$$L_j = \lambda[I_{\epsilon_j},\ 0_{\epsilon_j,1}] - [0_{\epsilon_j},\ I_{\epsilon_j,1}].$$

Here I_{ϵ_j} is the ϵ_j-by-ϵ_j identity matrix and $0_{\epsilon_j,1}$ is the ϵ_j-by-1 zero matrix. Let $x(t)$ be a classical solution of (28) and let $z(t) = Y^{-1}x(t)$. Partition $z(t)$ conformally with (32) as $z^T = [z_0^T, z_1^T, \ldots, z_{p+q}^T]^T$. It is easy to show that $x(t)$ is a classical solution of (28) if and only if $E_0\dot{z}_0(t) = A_0 z_0(t)$, $z_j(t) \equiv 0$ for $j = p+1,\, p+2, \ldots p+q$, and for $j = 1,\, 2,\, 3, \ldots p$, $z_j(t)$ satisfies the under determined differential equation

$$[I_{\epsilon_j},\, 0_{\epsilon_j,1}]\dot{z}_j(t) = [0_{\epsilon_j,1},\, I_{\epsilon_j}]z(t). \tag{33}$$

Using the explicit expression for the matrix relation exponential in Appendix A, an elementary but tedious calculation shows that $x(t_1) = \exp(E\backslash A(t_1-t_0))\, x(t_0)$ if and only if $z_0(t_1) = \exp(E_0\backslash A_0(t_1-t_0))\, z_0(t_0)$, and $z_j = 0$ for $j = p+1,\, p+2, \ldots p+q$. The exponential matrix relation $x(t_1) = \exp(E\backslash A(t_1-t_0))\, x(t_0)$ does not capture (33).

Nevertheless, although the exponential matrix relation puts no restriction on $z_j(t)$ in (33), the conclusion of Corollary 3.3 still holds. If $t_0 \neq t_1$, then for any choice of y_0, $y_1 \in \mathbb{C}^{\epsilon_j+1}$, there is a solution of (33) that interpolates y_0 and y_1 at t_0 and t_1. The solutions of (33) take the form $z_2(t) = \dot{z}_1(t)$, $z_3(t) = \dot{z}_2(t) = \ddot{z}_1(t)$, $\ldots$, $z_{\epsilon_j+1}(t) = \dot{z}_{\epsilon_j}(t) = z^{(\epsilon_j)}(t)$. A solution of (33) that interpolates y_0 and y_1 at t_0 and t_1 is obtained by choosing $z_1(t)$ to be the polynomial of degree $2\epsilon_j+1$ satisfying the osculatory interpolation conditions $z_1(t_0) = y_{10}$, $\dot{z}_1(t_0) = y_{20}$, $\ldots z_1^{(\epsilon_j)}(t_0) = y_{\epsilon_j+1,0}$ and $z_1(t_1) = y_{11}$, $\dot{z}_1(t_1) = y_{21}$, $\ldots z_1^{(\epsilon_j)}(t_1) = y_{\epsilon_j+1,1}$.

It follows that every choice of t_0, $t_1 \in \mathbb{R}$ and boundary conditions x_0, $x_1 \in \mathbb{C}^n$ there is a solution $x(t)$ of (28) such that $x(t_0) = x_0$ and $x(t_1) = x_1$ if and only if $x_1 \in \exp(E\backslash A(t_1-t_0))\, x_0$ and $x_0 \in \exp(E\backslash A(t_0-t_1))\, x_1$. This can be extended to an arbitrary number of boundary conditions at distinct values of t.

3.3 An Inverse-Free Sign Function Iteration

The matrix sign function [22, 31, 41] gives rise to an unusual family of algorithms for finding deflating subspaces, solving (generalized) algebraic Riccati equations and (generalized) Lyapunov equations. Because they are rich in matrix-matrix operations, matrix sign function algorithms are well suited to computers with advanced architectures [4, 6, 15, 24, 25]. Matrix sign function algorithms have attracted much attention through the last three decades; the survey [31] lists over 100 references. The rounding error analysis and perturbation theory are becoming understood [5, 18, 19, 20, 46]. Its presence in the SLICOT library [14, SB02OD], the PLiC library [15, PMD05RD] and as a prototype in the ScaLAPACK library [16] is an indication of its maturity and acceptance.

For $z \in \mathbb{C}$, define $\operatorname{sign}(z)$ by

$$\operatorname{sign}(z) = \begin{cases} 1 & \text{if the real part of } z \text{ is positive} \\ -1 & \text{if the real part of } z \text{ is negative.} \end{cases}$$

If z has zero real part, then we leave $\operatorname{sign}(z)$ undefined. If $A \in \mathbb{C}^{n\times n}$ has no eigenvalue with zero real part and has Jordan canonical form $A = M(\Lambda + N)M^{-1}$ where $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, N is nilpotent and $N\Lambda = \Lambda N$, then (see [41])

$$\operatorname{sign}(A) = M \operatorname{diag}(\operatorname{sign}(\lambda_1), \operatorname{sign}(\lambda_2), \operatorname{sign}(\lambda_3), \ldots, \operatorname{sign}(\lambda_n))M^{-1}.$$

Note that $\operatorname{sign}(A)^2 = I$, i.e., $\operatorname{sign}(A)$ is a square root of I, and $\operatorname{null}(\operatorname{sign}(A) \pm I)$ is $\mathcal{V}_{\mp}(A)$ the invariant subspaces of A corresponding to eigenvalues in the open left- and right-half plane, respectively.

Gardiner and Laub [24] proposed a generalization of the matrix sign to matrix pencils $\lambda E - A$ in which both A and E are nonsingular. They defined the sign of A with respect to E as the matrix $\operatorname{sign}(A, E) = E \operatorname{sign}(E^{-1}A) = \operatorname{sign}(AE^{-1})E$. The *pencil sign function* is the pencil $\operatorname{sign}(\lambda E - A) = \lambda E - \operatorname{sign}(A, E)$. If both A and E are nonsingular, then $\lambda E - A$ has Weierstraß canonical form $XEY = I$, $XAY = J$ where X, Y and J are nonsingular and J is in Jordan canonical form. The pencil sign function $\lambda E - \operatorname{sign}(A, E)$ has Weierstraß canonical form $X\hat{E}Y = I$, $X\hat{A}Y = \operatorname{sign}(J)$. Note that $\operatorname{null}(\operatorname{sign}(A, E) \pm E)$ is $\mathcal{V}_{\mp}(\lambda E - A)$ the right deflating subspaces corresponding to eigenvalues in the open left- and right-half plane respectively. Such deflating subspaces are the key computation in some numerical algorithms in computational control [22, 24, 34, 39, 41, 49].

A *right-handed sign pencil* is any pencil in the form

$$\lambda\tilde{E} - \tilde{A} = \lambda(\tilde{X}E) - (\tilde{X} \operatorname{sign}(A, E))$$

for some nonsingular matrix $\tilde{X}$. If $\lambda\tilde{E} - \tilde{A}$ is a right-handed sign pencil, then

$$\operatorname{null}(A \pm E) = \mathcal{V}_{\mp}(\lambda E - A).$$

A *left-handed sign pencil* is any pencil in the form

$$\lambda\tilde{E} - \tilde{A} = \lambda(E\tilde{Y}) - (\operatorname{sign}(A, E)\tilde{Y})$$

for some nonsingular matrix $\tilde{Y}$. The pencil sign function $\lambda E - \operatorname{sign}(A, E)$ is ambidextrous.

One of the first numerical iterations proposed to compute the matrix sign function is

$$A_0 = A, \quad A_{j+1} = (A_j + A_j^{-1})/2, \quad j = 0, 1, 2, \ldots \tag{34}$$

If A has no eigenvalue with zero real part, then $\lim_{j\to\infty} A_j = \operatorname{sign}(A)$ [41]. This is Newton's method applied to the nonlinear equation $X^2 - I = 0$. Thus, (34) has local quadratic convergence rate.

Iteration (34) extends to matrix pencils $\lambda E - A$ [24] as follows.

$$\hat{A}_0 = A, \quad \hat{A}_j = \frac{1}{2}\left(\hat{A}_{j-1} + E\hat{A}_{j-1}^{-1}E\right), \quad j = 0, 1, 2, \ldots \tag{35}$$

If both E and A are nonsingular, then $\hat{A}_j$ converges to $\operatorname{sign}(A, E)$.

The Gardiner-Laub algorithm [24] to calculate $\operatorname{sign}(A, E)$ is essentially an explicit implementation of (35). It avoids explicitly forming the product $E^{-1}\hat{A}_j$. However, it does explicitly form $E\hat{A}_j^{-1}E$. The condition number of $\hat{A}_j$ for inversion is not closely linked to the conditioning of the deflating subspaces of $\lambda E - A$. However, inverting ill-conditioned $\hat{A}_j$ in (35) may introduce significant rounding errors that change the deflating subspaces. The following example demonstrates this.

	Backward Errors		
p	(35)	QZ	Alg. 1
1	10^{-16}	10^{-16}	10^{-16}
2	10^{-15}	10^{-15}	10^{-15}
3	10^{-12}	10^{-15}	10^{-15}
4	10^{-10}	10^{-15}	10^{-13}
5	10^{-9}	10^{-16}	10^{-13}
6	10^{-7}	10^{-16}	10^{-12}
7	10^{-6}*	10^{-16}	10^{-12}
8	10^{-5}*	10^{-16}	10^{-11}
9	10^{-4}*	10^{-16}	10^{-10}
10	10^{-3}*	10^{-16}	10^{-11}

	Forward Errors			
p	(35)	QZ	Alg. 1	ε/dif
1	10^{-15}	10^{-16}	10^{-16}	10^{-16}
2	10^{-15}	10^{-15}	10^{-15}	10^{-15}
3	10^{-12}	10^{-15}	10^{-14}	10^{-14}
4	10^{-9}	10^{-13}	10^{-13}	10^{-13}
5	10^{-8}	10^{-13}	10^{-12}	10^{-12}
6	10^{-6}	10^{-14}	10^{-12}	10^{-11}
7	10^{-5}*	10^{-12}	10^{-11}	10^{-10}
8	10^{-4}*	10^{-11}	10^{-10}	10^{-10}
9	10^{-3}*	10^{-12}	10^{-11}	10^{-10}
10	10^{-3}*	10^{-11}	10^{-10}	10^{-9}

Table 1: Rounding error induced forward and backward errors in the computed eigenvector of with eigenvalue $-1/p$ of the pencil (36). The tables compare (35), the QZ algorithm [2, 40] and Algorithm 1. Asterisks indicate examples in which (35) failed to satisfy its convergence tolerance $\|\hat{A}_{j+1} - \hat{A}_j\|_F \leq 10^{-10}\|\hat{A}_{j+1}\|_F$ in 50 iterations. For $p = 10$, (35) encounters many highly ill-conditioned matrix inverses. In the right-most-column, ε/dif is is an estimate largest forward error that can be caused by a backward error of roughly the unit round. See [5, 45] for details.

Example 3.4 Let p be a nonzero scalar, E be the 10-by-10 matrix all of whose entries are one, U be the 10-by-10 elementary reflector $U = I - 0.2 \cdot E$, H_p be the 10-by-10 Jordan block with eigenvalue $1/p$, K_p be the 10-by-10 diagonal matrix with $(1,1)$ entry equal to -1 and all remaining diagonal entries equal to 1. Now let

$$\lambda E_p - A_p = \lambda(UH_pU) - (UK_pU). \tag{36}$$

The pencils $\lambda E_p - A_p$ have one eigenvalue equal to $-1/p$ and a multiplicity 9 eigenvalue equal to $1/p$. The eigenvalue $1/p$ corresponds to a single 9-by-9 Jordan block. The eigenvalue $-1/p$ is simple, and the conditioning of the corresponding one dimensional right deflating grows only moderately as p increases from 1 to 10.

We calculated a normalized basis of the one dimensional stable right deflating subspace corresponding to the eigenvalue $-1/p$ as an eigenvector obtained from the QZ algorithm [2, 40] and also as the null space of $E + \operatorname{sign}(A, E)$ using (35) to calculate $\operatorname{sign}(A, E)$. (The computations were run under MATLAB version 6 on a Dell Precision workstation with unit roundoff approximately 2.22×10^{-16}.) This produced rounding error corrupted, normalized, approximate eigenvectors $v_{p,\mathrm{qz}}$ from the QZ algorithm, $v_{p,\mathrm{gl}}$ from (35) (and $v_{p,\mathrm{ps}}$ from Algorithm 1 presented in Section 3.3 below). Comparing these to the first column of U_p, u_p, gives the forward or absolute errors $\|v_{p,\mathrm{qz}} - u_p\|_2$ and $\|v_{p,\mathrm{gl}} - u_p\|_2$. The backward errors are the smallest singular values of $[E_p v_{p,\mathrm{qz}},\ A_p v_{p,\mathrm{qz}}]$ and $[E_p v_{p,\mathrm{gl}},\ A_p v_{p,\mathrm{gl}}]$. Table 1 displays the rounding error induced forward and backward errors for $p = 1, 2, 3, \ldots, 10$. It also lists an estimate of the largest forward error that could be caused by perturbing the entries of E_p and A_p by quantities as large as the unit roundoff. In the table, the estimate is denoted by ε/dif. See [5, 45] for details. It demonstrates how ill-conditioned E and/or $\hat{A}_j$ in (35) can adversely affect both forward and backward errors. In contrast, the backward numerically stable QZ algorithm is unaffected.

As p varies from $p = 1$ to $p = 10$, the condition number of E_p, $\|E_p^{-1}\|\,\|E_p\|$, varies from 6 to 10^{10}. For $p > 7$ the iterates $\hat{A}_k$ in (35) are so ill-conditioned that our program failed to meet its stopping criterion $\|\hat{A}_{j+1} - \hat{A}_j\|_F \leq 10^{-10}\|\hat{A}_{j+1}\|_F$ where ε is the machine precision 2.22×10^{-16}. In that case, we terminated the program after 50 iterations. For $p = 10$, many iterates had condition numbers larger than 10^{14}.

3.3.1 An Inverse Free Pencil Sign Function Iteration

This subsection describes in detail a modification and extension of (35) that was briefly introduced in [10, 13]. Using matrix relations, the modified iteration avoids explicit matrix inversions.

Multiplying (35) from the left by E_j^{-1} yields (34) applied to $E^{-1}A$. Iteration (35) defines the sequence of pencils $\lambda E - \hat{A}_j$. Regarded as a sequence of matrix relations $(E\backslash\hat{A}_j)$, (35) becomes

$$(E_{j+1}\backslash A_{j+1}) = (\beta_j I\backslash\alpha_j I)\left[(E_j\backslash A_j) + ((E_j\backslash A_j))^{-1}\right]. \tag{37}$$

where $\alpha_j,\ \beta_j \in \mathbb{C}$ satisfy $\alpha_j/\beta_j = 1/2$. If E_j is nonsingular, then one choice of $[\tilde{E}_2,\ \tilde{E}_1]$ in Theorem 2.8 is $[\tilde{E}_2,\ \tilde{E}_1] = [I,\ E_j A_j^{-1}]$. With this choice and $\alpha_j \equiv 1$, $\beta_j \equiv 2$ (37) reduces to (35).

Many other choices of $[\tilde{E}_2,\ \tilde{E}_1]$ are possible. For example, if

$$\begin{bmatrix} -E_j \\ A_j \end{bmatrix} = \begin{bmatrix} Q_{j,11} & Q_{j,12} \\ Q_{j,21} & Q_{j,22} \end{bmatrix} \begin{bmatrix} R_j \\ 0 \end{bmatrix} \tag{38}$$

is a QR factorization, then $[\tilde{E}_2,\ \tilde{E}_1] = [Q_{j,12}^H,\ Q_{j,22}^H]$ is a possibility. With this choice, (37) becomes

$$\begin{array}{rcl} A_{j+1} & = & \alpha_j\left(Q_{j,12}^H A_j + Q_{j,22}^H E_j\right) \\ E_{j+1} & = & \beta_j Q_{j,12}^H E_j \end{array} \tag{39}$$

where $E_0 = E$, $A_0 = A$ and $\alpha_j,\ \beta_j \in \mathbb{C}$ are numbers for which $\alpha_j/\beta_j = 1/2$. Equivalently, using $Q_{j,12}^H E_j = Q_{j,22} A_j$ from (15), (39) can also be expressed as

$$\begin{bmatrix} E_{j+1} \\ A_{j+1} \end{bmatrix} = \frac{\beta}{2} \begin{bmatrix} Q_{j,12} & Q_{j,22} \\ Q_{j,22} & Q_{j,12} \end{bmatrix}^H \begin{bmatrix} E_j \\ A_j \end{bmatrix}.$$

For simplicity, throughout this paper we will choose α_j and β_j to be real and independent of j. Consequently, we drop the subscript j. We show below that $\alpha = 1/\sqrt{2}$ and $\beta = \sqrt{2}$ is necessary for convergence of the sequences E_j and A_j. An explicit implementation of (38), (39) is an inverse free pencil sign iteration. (See Algorithm 1 below.) The resulting algorithm is matrix-matrix multiplication rich and well suited to computers with advanced architectures. Note that although (35) and (39) generate different sequences of matrices, they define the same sequence of matrix relations, i.e., $(E_j\backslash A_j) = (E\backslash\hat{A}_j)$, for all $j \in \mathbb{N}_0$.

Theorem 3.5 *If both A and E are nonsingular, (i.e., if $\lambda E - A$ has neither an infinite eigenvalue nor an eigenvalue with zero real part), then the sequences of matrices $E_j \in \mathbb{C}^{n\times n}$ and $A_j \in \mathbb{C}^{n\times n}$ generated by (38) and (39) have the following properties for all $j = 0, 1, 2, \ldots$.*

1. *Both A_j and E_j are nonsingular, i.e., the pencil $\lambda E_j - A_j$ has neither an infinite eigenvalue nor an eigenvalue with zero real part eigenvalues.*

2. *If $\lambda E_j - A_j$ has an eigenvalue $\lambda \in \mathbb{C}$ with corresponding right eigenvector $x \in \mathbb{C}^n$, then $\lambda E_{j+1} - A_{j+1}$ has an eigenvalue $(\lambda + \lambda^{-1})/2$ with the same eigenvector x.*

3. *If $\mathcal{V}$ is a right deflating subspace of $\lambda E - A$ then $\mathcal{V}$ is a right deflating subspace of $\lambda E_j - A_j$.*

4. *The sequence of relations $(E_j\backslash A_j)$ generated by (38) and (39) converges quadratically (as a sequence of subspaces of $\mathbb{C}^n \times \mathbb{C}^n$ in the usual largest-canonical-angle/gap-metric topology [28] [44, Ch.II§4]) to a limiting relation $(E_\infty\backslash A_\infty)$ and $\lambda E_\infty - A_\infty$ is a right-handed sign pencil.*

5. $\lim_{j\to\infty} E_j^{-1} A_j = \mathrm{sign}(E^{-1}A)$.

Proof. Recall that (35) and (39) generate the same sequence of matrix relations, i.e., in the notation of (35) and (39), for all j, $(E_j\backslash A_j) = (E\backslash\hat{A}_j)$. The stated properties follow from the corresponding properties of $\hat{A}_j$ in (35) [24]. ▫

Theorem 3.5 shows that right eigenvectors and right deflating subspaces are invariant throughout (39). In particular, (39) preserves special structure that the right deflating subspaces may have. Linear quadratic and H_∞ optimal control problems [24, 34, 39, 35] along with quadratic eigenvalue linear damping models [27] lead to invariant subspace problems whose right deflating subspaces have symplectic structure. The symplectic structure is preserved by (39).

Recall that for any nonsingular matrix M_j, $((M_jE_j)\backslash(M_jA_j)) = (E_j\backslash A_j)$. Consequently, convergence of $(E_j\backslash A_j)$ does not imply convergence of the individual sequences of matrices E_j and A_j. For example, if $E = A = I$, $\alpha = 1$ and $\beta = 2$ in (39), then a direct calculation shows that one may choose $Q_{j,12} = Q_{j,22} = I/\sqrt{2}$ in (38). With these choices, $E_j = (\sqrt{2})^j I$ and $A_j = (\sqrt{2})^j I$ and $\lim_{j\to\infty} E_j = \lim_{j\to\infty} A_j = \infty$. If $\alpha = 1/2$ and $\beta = 1$, then $\lim_{j\to\infty} E_j = \lim_{j\to\infty} A_j = 0$. Converging to zero is at least as problematic as not converging. Note that in this example, $\operatorname{sign}(A, E) = I = A$ and for all j, $(E\backslash A) = (E_j\backslash A_j) = (I\backslash I)$. The sequence of matrix relations is stationary. Using $\alpha = 1/\sqrt{2}$, $\beta = \sqrt{2}$ in (39) yields a stationary sequence of matrices $E_j = I$, $A_j = I$. The following theorem shows that this choice of $\alpha = 1/\sqrt{2}$ and $\beta = \sqrt{2}$ is necessary for convergence of the individual matrices.

Theorem 3.6 *If E_j and A_j obey (38) and (39) with nonnegative diagonal entries in R, and if both $E_\infty = \lim_{j\to\infty} E_j$ and $A_\infty = \lim_{j\to\infty} A_j$ exist and are nonsingular, then $\alpha = \pm 1/\sqrt{2}$, $\beta = \pm\sqrt{2}$ and there are unitary matrices $W, Y \in \mathbb{C}^{n\times n}$ with $W = W^H$ for which Q_∞ has CS decomposition*

$$Q_\infty = \lim_{j\to\infty} Q_j = \frac{\pm 1}{\sqrt{2}}\begin{bmatrix} Y & I \\ -WY & W \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & W \end{bmatrix}\left(\frac{\pm 1}{\sqrt{2}}\begin{bmatrix} I & I \\ -I & I \end{bmatrix}\right)\begin{bmatrix} Y & 0 \\ 0 & I \end{bmatrix}. \tag{40}$$

Proof. Taking limits in (39) and noting that both E_∞ and A_∞ are nonsingular shows that $Q_{\infty,12} = \lim_{j\to\infty} Q_{j,12}$, $Q_{\infty,22} = \lim_{j\to\infty} Q_{j,22}$ exist and

$$\begin{aligned} Q_{\infty,12} &= \beta^{-1} I \\ A_\infty &= \tfrac{\alpha}{\beta} A_\infty + \alpha Q_{\infty,22}^H E_\infty. \end{aligned} \tag{41}$$

Solving for $Q_{\infty,22}$ and using $\alpha/\beta = 1/2$ gives $(2\alpha)^{-1} A_\infty E_\infty^{-1} = Q_{22}^H$ or, equivalently,

$$(2\alpha)^{-1} E_\infty \left(E_\infty^{-1} A_\infty\right) E_\infty^{-1} = Q_{22}^H. \tag{42}$$

Recall that $E_\infty^{-1} A_\infty = \operatorname{sign}(E^{-1}A)$ which is a square root of I. Squaring both sides of (42) gives $(2\alpha)^{-2} I = \left(Q_{\infty,22}^H\right)^2$ or, equivalently,

$$I = (2\alpha)^2 Q_{\infty,22}^2. \tag{43}$$

The columns of $\begin{bmatrix} Q_{\infty,12} \\ Q_{\infty,22} \end{bmatrix} = \begin{bmatrix} \beta^{-1} I \\ Q_{\infty,22} \end{bmatrix}$ are orthonormal, because they are the limit of the corresponding columns of the orthonormal matrices $\begin{bmatrix} Q_{j,12} \\ Q_{j,22} \end{bmatrix}$. Hence,

$$\begin{bmatrix} \beta^{-1} I \\ Q_{\infty,22} \end{bmatrix}^H \begin{bmatrix} \beta^{-1} I \\ Q_{\infty,22} \end{bmatrix} = \beta^{-2} I + Q_{\infty,22}^H Q_{\infty,22} = I \tag{44}$$

which implies that $Q_{\infty,22} = \pm\sqrt{1-\beta^{-2}}\, W$ for some unitary matrix $W \in \mathbb{R}^{n\times n}$. It follows from (43) and the scalar-times-unitary structure of $Q_{\infty,22}$ that $(2\alpha)^{-2} = 1 - \beta^{-2}$. This and the constraint $\alpha/\beta = 1/2$ establish $\alpha = \pm 1/\sqrt{2}$, $\beta = \pm\sqrt{2}$. The $(1,2)$ block of (40) now follows from (41). The $(2,2)$ block of (40) follows from (44). Equation (43) now simplifies to $W^2 = I$. Multiplying this by W^H gives $W = W^H$.

Continuity of the QR factorization of full column rank matrices together with the nonnegative diagonal entries of R implies that $\lim_{j\to\infty} Q_{j,11} = Q_{\infty,11}$ and $\lim_{j\to\infty} Q_{j,21} = Q_{\infty,21}$ exist. The fact that Q_∞ is unitary and the previously established form of $Q_{\infty,12} = (\pm 1/\sqrt{2}) I$ and $Q_{\infty,22} = (\pm 1/\sqrt{2}) W$ implies that

$$0 = Q_{\infty,11}^H Q_{\infty,12} + Q_{\infty,21}^H Q_{\infty,22} = \frac{\pm 1}{\sqrt{2}}\left(Q_{\infty,11}^H + Q_{\infty,12}^H W\right).$$

So, $Q_{\infty,11} = -W^H Q_{\infty,12} = -WQ_{\infty,12}$. This gives

$$I = Q^H_{\infty,11} Q_{\infty,11} + Q^H_{\infty,12} Q_{\infty,12} = Q^H_{\infty,11} Q_{\infty,11} + Q^H_{\infty,11} Q_{\infty,11}$$

which establishes the $(1,1)$ and $(2,1)$ block of (40). $\square$

Submatrices $Q_{j,12}$ and $Q_{j,22}$ are determined by (38) only up to left multiplication by an arbitrary unitary factor. Even with $\alpha = 1/\sqrt{2}$ and $\beta = \sqrt{2}$, the sequences E_j and A_j in (38) and (39) may or may not converge depending on how this nonuniqueness is resolved. The following particular choice of Q in (38) generates empirically convergent sequences E_j and A_j in (39) and admits a particularly efficient implementation.

Assume that both A_j and E_j are nonsingular. Let $E_j = U_{E,j} T_{E,j}$ and $A_j = U_{A,j} T_{A,j}$ be QR factorizations of E_j and A_j, respectively where $U_{E,j}$, $U_{A,j} \in \mathbb{R}^{n\times n}$ are unitary and $T_{E,j}$, $T_{A,j} \in \mathbb{R}^{n\times n}$ are upper triangular with non-negative diagonal entries. Let

$$\begin{aligned} \begin{bmatrix} -T_{E,j} \\ T_{A,j} \end{bmatrix} &= \begin{bmatrix} V_{j,11} & V_{j,12} \\ V_{j,21} & V_{j,22} \end{bmatrix} \begin{bmatrix} R_j \\ 0 \end{bmatrix} \\ \begin{bmatrix} \text{◹} \\ \text{◹} \end{bmatrix} &= \begin{bmatrix} \text{◹} & \text{◺} \\ \text{◹} & \text{◺} \end{bmatrix} \begin{bmatrix} \text{◹} \\ 0 \end{bmatrix} \end{aligned} \tag{45}$$

be the QR factorization with orthogonal factor V in which, as indicated schematically, $V_{j,11}$ and $V_{j,21}$ are upper triangular and $V_{j,12}$ and $V_{j,22}$ are lower triangular. To promote convergence, select the diagonal entries of $V_{j,12}$ to be non-negative. (We give an explicit algorithm to compute this factorization below.) One particular choice of the orthogonal factor Q_j in (38) is

$$Q_j = \begin{bmatrix} U_{E,j} & 0 \\ 0 & U_{A,j} \end{bmatrix} \begin{bmatrix} V_{j,11} & V_{j,12} \\ V_{j,21} & V_{j,22} \end{bmatrix} = \begin{bmatrix} \square & 0 \\ 0 & \square \end{bmatrix} \begin{bmatrix} \text{◹} & \text{◺} \\ \text{◹} & \text{◺} \end{bmatrix}. \tag{46}$$

If both E_j and A_j are nonsingular, then the matrix Q_j is uniquely determined by the nonsingular matrices E_j and A_j and the choice of nonnegative diagonal entries in the triangular factors $T_{E,j}$, $T_{A,j}$, R_j and $V_{j,12}$.

In the notation of (45) and (46) with $\alpha = 1/\sqrt{2}$, $\beta = \sqrt{2}$, the iteration (39) becomes

$$\begin{aligned} A_{j+1} &= \frac{1}{\sqrt{2}} \left(V^H_{j,12} U^H_{E,j} A_j + V^H_{j,22} U^H_{A,j} E_j \right) \\ E_{j+1} &= \sqrt{2} V^H_{j,12} U^H_{E,j} E_j = \sqrt{2} V^H_{j,22} U^H_{A,j} A_j. \end{aligned} \tag{47}$$

Note that E_{j+1} is upper triangular with non-negative diagonal entries, because it is the product of the upper triangular matrices $V^H_{j,12}$ and $U^H_{E,j} E_j = T_{E,j}$ both of which were chosen to have nonnegative diagonal entries. Consequently, $U_{E,j+1} = I$.

Scaling

If the pencil $\lambda E - A = \lambda E_0 - A_0$ has eigenvalues far away from their limiting value of ± 1, then initially convergence may be slow. Initially slow convergence can sometimes be avoided by scaling, i.e., at the beginning of each iteration, select a positive scalar $\gamma_j > 0$ and replace A_j (say) by $\gamma_j A_j$ before calculating the factorization (38) [19, 4, 31]. It is easy to show that if $\gamma > 0$, then $\operatorname{sign}(\lambda E - A) = \operatorname{sign}(\lambda E - \gamma A)$, so scaling does not change the pencil sign function. With this modification, (38) becomes

$$\begin{bmatrix} -E_j \\ \gamma_j A_j \end{bmatrix} = \begin{bmatrix} Q_{j,11} & Q_{j,12} \\ Q_{j,21} & Q_{j,22} \end{bmatrix} \begin{bmatrix} R_j \\ 0 \end{bmatrix}$$

The variation of determinantal scaling suggested in [24] applies here:

$$\gamma_j = \frac{|\det(E_j)|^{-1/n}}{|\det(A_j)|^{-1/n}}.$$

This choice makes the eigenvalues of the scaled pencil, $\lambda E_j - \gamma_j A_j$, have geometric mean equal to one. In the context of Algorithm 1, E_j is triangular and the magnitude of the determinant of A_j can be obtained from the triangular QR factor T_A which must be computed anyway. Hence, the scale factor γ_j may be computed with negligible extra work relative to the rest of Algorithm 1.

3.3.2 An Inverse Free Pencil Sign Function Algorithm

In this section we describe a procedure for computing a right-handed pencil sign function using (45), (46), and (47). The QR factorizations $E_j = U_{E,\,j}T_{E,\,j}$ and $A_j = U_{A,\,j}T_{A,\,j}$ may be obtained from a well known procedure like Householder's algorithm [28]. The orthogonal matrix V with triangular blocks in (45) may be obtained as a product of rotations. The process is well illustrated by an $n = 3$ example. Initially $\left[\begin{smallmatrix} -T_{E,\,j} \\ T_{A,\,j} \end{smallmatrix}\right]$ has the zero structure

$$\begin{bmatrix} -T_{E,\,j} \\ T_{A,\,j} \end{bmatrix} = \begin{bmatrix} \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \\ \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \end{bmatrix}$$

where $\mathtt{x}$'s represent entries that may be nonzero and blanks represent entries that must be zero. Select plane rotations $R_{j,\,14}$, $R_{j,\,25}$ and $R_{j,\,36}$ in the $(1,4)$, $(2,5)$ and $(3,6)$ planes, respectively, to zero the $(4,1)$, $(5,2)$ and $(6,3)$ elements of $\left[\begin{smallmatrix} -T^{(1)}_{E,\,j} \\ T^{(1)}_{A,\,j} \end{smallmatrix}\right] := R_{j,\,14}R_{j,\,25}R_{j,\,36}\left[\begin{smallmatrix} -T_{E,\,j} \\ T_{A,\,j} \end{smallmatrix}\right]$. Schematically, this becomes

$$(R_{j,\,14}R_{j,\,25}R_{j,\,36})\begin{bmatrix} -T_{E,\,j} \\ T_{A,\,j} \end{bmatrix} = \left[\begin{array}{ccc|ccc} \mathtt{x} & 0 & 0 & \mathtt{x} & 0 & 0 \\ 0 & \mathtt{x} & 0 & 0 & \mathtt{x} & 0 \\ 0 & 0 & \mathtt{x} & 0 & 0 & \mathtt{x} \\ \mathtt{x} & 0 & 0 & \mathtt{x} & 0 & 0 \\ 0 & \mathtt{x} & 0 & 0 & \mathtt{x} & 0 \\ 0 & 0 & \mathtt{x} & 0 & 0 & \mathtt{x} \end{array}\right] \begin{bmatrix} \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \\ \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \end{bmatrix} = \begin{bmatrix} \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \\ & & \end{bmatrix}.$$

Select plane rotations $R_{j,\,24}$ and $R_{j,\,35}$ in the $(2,4)$ and $(3,5)$ planes, respectively, to zero the $(4,2)$ and $(5,3)$ elements

$$(R_{j,\,24}R_{j,\,35})\begin{bmatrix} -T^{(1)}_{E,\,j} \\ T^{(1)}_{A,\,j} \end{bmatrix} = \left[\begin{array}{cccccc} \mathtt{x} & 0 & 0 & \mathtt{x} & 0 & 0 \\ \mathtt{x} & \mathtt{x} & 0 & \mathtt{x} & \mathtt{x} & 0 \\ 0 & \mathtt{x} & \mathtt{x} & 0 & \mathtt{x} & \mathtt{x} \\ \mathtt{x} & \mathtt{x} & 0 & \mathtt{x} & \mathtt{x} & 0 \\ 0 & \mathtt{x} & \mathtt{x} & 0 & \mathtt{x} & \mathtt{x} \\ 0 & 0 & \mathtt{x} & 0 & 0 & \mathtt{x} \end{array}\right] \begin{bmatrix} \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \\ \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \end{bmatrix} = \begin{bmatrix} \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \\ & & \mathtt{x} \\ & & \\ & & \end{bmatrix}.$$

Call the result $\left[\begin{smallmatrix} -T^{(2)}_{E,\,j} \\ T^{(2)}_{A,\,j} \end{smallmatrix}\right]$. Finally select a plane rotation $R_{j,\,34}$ in the $(3,4)$ plane to zero the $(4,3)$ element

$$R_{j,\,34}\begin{bmatrix} -T^{(2)}_{E,\,j} \\ T^{(2)}_{A,\,j} \end{bmatrix} = \left[\begin{array}{cccccc} \mathtt{x} & 0 & 0 & \mathtt{x} & 0 & 0 \\ \mathtt{x} & \mathtt{x} & 0 & \mathtt{x} & \mathtt{x} & 0 \\ \mathtt{x} & \mathtt{x} & \mathtt{x} & \mathtt{x} & \mathtt{x} & \mathtt{x} \\ \mathtt{x} & \mathtt{x} & \mathtt{x} & \mathtt{x} & \mathtt{x} & \mathtt{x} \\ 0 & \mathtt{x} & \mathtt{x} & 0 & \mathtt{x} & \mathtt{x} \\ 0 & 0 & \mathtt{x} & 0 & 0 & \mathtt{x} \end{array}\right] \begin{bmatrix} \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \\ \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \end{bmatrix} = \begin{bmatrix} \mathtt{x} & \mathtt{x} & \mathtt{x} \\ & \mathtt{x} & \mathtt{x} \\ & & \mathtt{x} \\ & & \\ & & \\ & & \end{bmatrix} =: R_j$$

INPUT: A pencil $\lambda E - A$ with neither an infinite eigenvalue nor an eigenvalue with zero real part. A convergence tolerance $\tau > 0$.

OUTPUT: $\lambda E - A$ is overwritten by a (an approximate) right-handed pencil sign function.

$E \to U_E T_E$ $\{QR$ factorization$\}$
$E \leftarrow T_E$; $\quad A \leftarrow U_E^H A$
$R_- \leftarrow 2\tau I$; $\quad R \leftarrow 0$
while $\|R_- - R\|_F > \tau \|R\|_F$ **do**
  $A \to U_A T_A$ $\{QR$ factorization$\}$
  $\gamma \to \det(E)^{-1/n} / \det(T_A)^{-1/n}$
  {Use $\gamma \to 1$ for unscaled algorithm. Note that E and T_A are upper triangular.}
  $R_- \leftarrow R$; $R \leftarrow \begin{bmatrix} -E \\ \gamma T_A \end{bmatrix}$; $V \leftarrow I_{2n}$
  for $k = 1, 2, 3, \ldots n$ **do**
    for $i = 1, 2, 3, \ldots n - k + 1$ **do**
      $j \leftarrow i + k - 1$
      Calculate a rotation $\left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right]$ such that $\left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right] \left[\begin{smallmatrix} r_{jj} \\ r_{ij} \end{smallmatrix}\right] = \left[\begin{smallmatrix} \sqrt{r_{jj}^2 + r_{ij}^2} \\ 0 \end{smallmatrix}\right]$.
      $R_{[j,n+i],:} \leftarrow \left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right] R_{[j,n+i],:}$
      $V_{:,[j,n+i]} \leftarrow V_{:,[j,n+i]} \left[\begin{smallmatrix} c & -s \\ s & c \end{smallmatrix}\right]^H$
    end for
  end for
  $V_{:,(n+1):(2n)} \leftarrow V_{:,(n+1):(2n)} \operatorname{diag}(\operatorname{sign}(v_{1,n+1}), \operatorname{sign}(v_{2,n+2}), \ldots, \operatorname{sign}(v_{n,2n}))$
  $A \leftarrow \left(\sqrt{2} V_{1:n,(n+1):(2n)}^H (\gamma A) + V_{(n+1):(2n),(n+1):(2n)}^H U_A^H E\right) / \sqrt{2}$
  $E \leftarrow \sqrt{2} V_{(1:n),(n+1:2n)}^H E$
end while

Algorithm 1: Inverse Free Matrix Sign Function

In this notation, V_j is the product

$$V_j = (R_{j,34} R_{j,24} R_{j,35} R_{j,14} R_{j,25} R_{j,36})^H = \begin{bmatrix} x & x & x & x & & \\ & x & x & x & x & \\ & & x & x & x & x \\ & & & & & \\ x & x & x & x & & \\ & x & x & x & x & \\ & & x & x & x & x \end{bmatrix}. \qquad (48)$$

Algorithm 1 computes a right-handed pencil sign function employing the efficient implementation of (38) described above and making use of the resulting triangular matrix structures. It uses $28n^3/3$ floating point operations to compute the QR factorization using Householder's method [28] and accumulate the two n-by-n blocks Q_{12} and Q_{22}. Taking advantage of the triangular structure in E, $V_{(1:n),(n+1:2n)}$ and $V_{1:n,(n+1):(2n)}$ the three matrix-matrix multiplies use roughly $5n^3/3$ floating point operations. So, each iteration uses a total of roughly $11n^3$ floating point operations. We observe empirically that 6–10 iterations is typical for well scaled pencils.

Singularities

The unscaled inverse free iteration (38) and (39) remains defined even in the case that A or E or both are singular. (This corresponds to the unscaled, $\gamma = 1$, version of Algorithm 1.) It is natural to ask what becomes of the sequence of iterates E_j, A_j in this case. Sun and Quintana–Ortí [48] studied (35) in the case that E is singular but A is nonsingular. Since, (35) and (39) are related by $(E \backslash \hat{A}_j) = (E_j \backslash A_j)$, the results in [48] readily generalize to (39).

Theorem 3.7

Suppose that $\lambda E - A$ is a regular pencil with Weierstraß canonical form

$$X(\lambda E - A)Y = \lambda \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & N \end{bmatrix} - \begin{bmatrix} J & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & I \end{bmatrix} \tag{49}$$

where J is nonsingular and in Jordan canonical form, M is nilpotent and N is nilpotent. (Any of J, M or N may be void, "0-by-0" matrices.) If $(E_j \backslash A_j)$ is the sequence of matrix relations generated by (38) and (39), then there is a sequence of nonsingular matrices $\tilde{X}_j$ for which for $j = 1, 2, 3, \ldots$

$$\tilde{X}_j E_j Y = \begin{bmatrix} I & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & N \end{bmatrix}, \qquad \tilde{X}_j A_j Y = \begin{bmatrix} K_j & 0 & 0 \\ 0 & H_j & 0 \\ 0 & 0 & G_j \end{bmatrix}$$

where $K_0 = J$, $G_0 = I$, $H_0 = I$, $\ldots K_j = (K_{j-1} + K_{j-1}^{-1})/2$ and for $j = 1, 2, 3, \ldots$

$$\begin{aligned} G_j &= G_{j-1} + N G_{j-1}^{-1} N = 2^{-j} I + \left(\frac{4^j - 1}{3(2^{j-1})} \right) N^2 + \text{ higher powers of } N^2 \\ H_j &= H_{j-1} + M H_{j-1}^{-1} M = 2^{-j} I + \left(\frac{4^j - 1}{3(2^{j-1})} \right) M^2 + \text{ higher powers of } M^2 \ . \end{aligned}$$

Proof. It suffices to analyze each diagonal block of (49) separately. In terms of relations, each iteration of (39) transforms $(I \backslash K_{j-1})$ to $(I \backslash (K_{j-1} + K_{j-1}^{-1})/2)$. The two nilpotent blocks $(I \backslash M)$ and $(N \backslash I)$ initially transform to

$$\begin{aligned} ((I \backslash M) + (M \backslash I))/2 &= (M \backslash I + M^2) \\ ((N \backslash I) + (I \backslash N))/2 &= (N \backslash I + N^2). \end{aligned}$$

respectively. The form of G_j and H_j follows by induction. $\square$

Remark 3.8 Although Theorem 3.7 does not show it, when $N^2 \neq 0$, then infinite eigenvalue structure grows relative to the finite eigenvalue structure and may be lost.

Remark 3.9 In the context of Theorem 3.7, if $N = 0$, then the matrix relation corresponding to the infinite eigenvalue structure of $\lambda E_j - A_j$, $(0 \backslash 2^{-j} I) = (0 \backslash I)$, is invariant throughout (39).

4 Numerical Examples

This section presents a few numerical examples demonstrating the effectiveness and empirical numerical stability of Algorithm 1.

Remark 4.1 A rounding error analysis of a single matrix relation product appears in [12]. There it is shown that if the factors are appropriately scaled and $\tilde{A}_2$ and $\tilde{E}_1$ in (4) are selected to have small componentwise condition number [8, 42], then the computation of a single matrix relation product is numerically backward stable. A rounding error analysis of a matrix relation sum and combinations of matrix relation sums and matrix relation products remains an open problem.

In addition to the examples below, we applied Algorithm 1 to several generalized algebraic Riccati equations derived from control systems in generalized state-space form [1]. The results are given in [13]. They show that Algorithm 1 yields accuracy comparable to and sometimes better than the generalized Schur vector method [3, 36].

Example 4.2 Consider again the pencil $\lambda E_p - A_p$ from (36) in Example 3.4. Using Algorithm 1 to obtain a right-handed sign pencil $\lambda \check{E}_p - \check{A}_p$, we calculated the one dimensional deflating subspace corresponding to the eigenvalue $-1/p$ by calculating a basis of null$(\check{E}_p + \check{A}_p)$. Forward and backward errors were computed as described in Example 3.4 and reported in Table 1.

In this example, backward and forward errors in the deflating subspace calculated from Algorithm 1 are five or more orders of magnitude smaller than in the deflating subspace calculated from (35). Algorithm 1's forward errors are comparable to and nearly as small as the forward errors of the backward stable QZ algorithm and comparable to the theoretically worst case forward errors of a backward stable algorithm. So, at least in this example, Algorithm 1 exhibits forward stability in the sense of [30, page 10], but (35) does not. (The backward stable QZ algorithm is *a fortiori* forward stable.)

Table 1 also shows gradual growth in the backward error. Our implementation of Algorithm 1 is not quite as numerically stable as the QZ algorithm.

Algorithm 1 takes between six and ten iterations to meet its stopping criterion $\tau = 10^{-10}$.

Example 4.3 This is Example 4.2 from [48] and Example 2 from [47]. Let

$$\hat{A} = Q^T \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{11}^T \end{bmatrix} Q$$

where A_{12} is a 10-by-10 matrix whose entries are random numbers drawn uniformly from $[0, 1]$ and

$$A_{11} = \begin{bmatrix} 1-\alpha & 0 & 0 & \cdots & \alpha \\ \alpha & 1-\alpha & 0 & \cdots & 0 \\ 0 & \alpha & 1-\alpha & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \\ 0 & 0 & \cdots & \alpha & 1-\alpha \end{bmatrix}.$$

Find the QR (orthogonal-triangular) factorization $\hat{A} = QR$, and set $A = R$ and $E = Q^T$. In this example, we find the stable right deflating subspace, i.e., the right deflating subspace of $\lambda E - A$ corresponding to eigenvalues with negative real part.

In the $\alpha = (1-10^{-7})/2$ example, (35) failed to satisfy its convergence tolerance $\|\hat{A}_{j+1} - \hat{A}_j\|_F \leq 10^{-10}\|\hat{A}_{j+1}\|_F$.

Table 2 reports the backward errors in the rounding error perturbed stable invariant subspace as calculated by (35), the QZ algorithm, and Algorithm 1. A rounding-error-free expression of the stable invariant subspace is unavailable, so the forward errors are calculated relative to the stable invariant subspace computed from the QZ algorithm. The right most column of Table 2 lists ε/dif, an estimate of the largest forward error that could be caused by perturbing the entries of E and A by quantities as large as the unit roundoff of the finite precision arithmetic.

In this example, the backward errors from Algorithm 1 are comparable to the backwards errors from the QZ algorithm. As α approaches 0.5, the backward errors from (35) grow several orders of magnitude larger. Algorithm 1's forward errors are consistent with forward stability, but forward stability of (35) is doubtful.

Algorithm 1 and (35) take between seven and sixteen iterations to meet their stopping criteria, except for $\alpha = (1 - 10^{-7})/2$, (35) failed to meet its stopping criterion $\|\hat{A}_{j+1} - \hat{A}_j\|_F \leq 10^{-10}\|\hat{A}_{j+1}\|_F$ and was arbitrarily halted after 50 iterations.

Example 4.4 This is Example 4.3 from [48]. (Example 3 from [47] is slightly different.) Let Q be a random orthogonal matrix generated as described in [43]. Using random numbers uniformly distributed over $[0, 1]$, let

$$\hat{A} = Q^T \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22}^T \end{bmatrix} Q$$

where A_{12} is a 5-by-5 random matrix A_{11} is an upper triangular random matrix with positive diagonal entries scaled by a real number $\beta < 0$, and A_{22} is an upper triangular random matrix

Backward Errors			
α	(35)	QZ	Alg. 1
$(1-10^{-1})/2$	10^{-15}	10^{-15}	10^{-15}
$(1-10^{-3})/2$	10^{-12}	10^{-15}	10^{-15}
$(1-10^{-5})/2$	10^{-9}	10^{-15}	10^{-14}
$(1-10^{-7})/2$	$10^{-5}*$	10^{-15}	10^{-14}

Forward Errors Relative to QZ			
α	(35)	Alg. 1	ε/dif
$(1-10^{-1})/2$	10^{-15}	10^{-15}	10^{-15}
$(1-10^{-3})/2$	10^{-12}	10^{-14}	10^{-13}
$(1-10^{-5})/2$	10^{-9}	10^{-12}	10^{-11}
$(1-10^{-7})/2$	$10^{-4}*$	10^{-10}	10^{-9}

Table 2: Example 4.3: Backward errors corresponding to the rounding error perturbed stable invariant subspace computed by (35), the QZ Algorithm [2, 40], and Algorithm 1. Asterisks indicate examples in which (35) failed to satisfy its convergence tolerance $\|\hat{A}_{j+1} - \hat{A}_j\|_F \leq 10^{-10}\|\hat{A}_{j+1}\|_F$. The forward errors are relative to the invariant subspace obtained from the backward stable QZ algorithm. The right-hand column is an estimate of the largest forward error that could be caused by perturbing the entries of E and A by quantities as large as the unit roundoff of the finite precision arithmetic.

Backward Errors			
β	(35)	QZ	Alg. 1
1.0	10^{-15}	10^{-15}	10^{-15}
0.5	10^{-11}	10^{-15}	10^{-15}
0.3	$10^{-9}*$	10^{-15}	10^{-14}
0.2	$10^{-7}*$	10^{-16}	10^{-12}
0.1	$10^{-2}*$	10^{-15}	10^{-10}

Forward Errors				
β	(35)	QZ	Alg. 1	ε/dif
1.0	10^{-15}	10^{-15}	10^{-15}	10^{-13}
0.5	10^{-9}	10^{-12}	10^{-12}	10^{-12}
0.3	$10^{-7}*$	10^{-10}	10^{-11}	10^{-12}
0.2	$10^{-6}*$	10^{-10}	10^{-9}	10^{-10}
0.1	$10^{-4}*$	10^{-8}	10^{-8}	10^{-8}

Table 3: Example 4.4: Backward errors corresponding to the rounding error perturbed stable invariant subspace computed by (35), the QZ Algorithm [2, 40], and Algorithm 1. Asterisks indicate examples in which (35) failed to satisfy its convergence tolerance $\|\hat{A}_{j+1} - \hat{A}_j\|_F \leq 10^{-10}\|\hat{A}_{j+1}\|_F$. In the $\beta = 0.1$ example, (35) inverted many highly ill-conditioned matrices. The right-hand column is as in Table 2.

with positive diagonal entries scaled by the positive real number $-\beta$. Find the QR (orthogonal-triangular) factorization $\hat{A} = QR$, and set $A = R$ and $E = Q^T$. In this example, we find the stable right deflating subspace, i.e., the right deflating subspace of $\lambda E - A$ corresponding to eigenvalues with negative real part.

Table 3 reports the backward errors in the rounding error perturbed stable invariant subspace as calculated by (35), the QZ algorithm, and Algorithm 1. For $\beta \leq 0.3$, (35) failed to satisfy its convergence criterion $\|\hat{A}_{j+1} - \hat{A}_j\|_F \leq 10^{-10}\|\hat{A}_{j+1}\|_F$ and was arbitrarily halted after 50 iterations. In the $\beta = 0.1$ example, (35) inverted many highly ill-conditioned matrices.

Algorithm 1 exhibits forward stability in this example. The $\beta = 0.5$ example is not consistent with forward stability of (35).

As in Example 3.4, Table 3 lists ε/dif, an estimate of the largest forward error that could be caused by perturbing the entries of E and A by quantities as large as the unit roundoff of the finite precision arithmetic.

These numerical experiments show that neither (35) nor Algorithm 1 are numerically backward stable. They also show that (35) is not, in general, forward stable. However, Algorithm 1 appears to be numerically forward stable. A proof or counter example to forward stability remains an open question.

5 Conclusions

This paper introduces arithmetic-like operations on matrix pencils. The pencil-arithmetic operations extend elementary formulas for sums and products of rational numbers and include the algebra of linear transformations as a special case. The set of m-by-n matrix relations has an identity element for multiplication, $(I\backslash I)$, and an identity element for addition $(I\backslash 0)$ Multiplication is associative and addition is commutative. Some matrix relations lack an additive or multiplicative inverse. The matrix relation $([0]\backslash[1])$ lacks both. There is only a partial distributive law for multiplication across addition.

Matrix relations lead to generalizations of the monodromy matrix, the matrix exponential and the matrix sign function that give a simplified and more unified understanding of these functions in the case of pencils with zero and infinite eigenvalues and pencils which are singular.

The rounding error analysis of matrix relation algorithms is encouraging but incomplete. The inverse free matrix sign function algorithm, Algorithm 1, is not backward stable, but the empirical evidence suggests that it is forward stable in the sense of [30, page 10].

References

[1] J. Abels and P. Benner. CAREX – a collection of benchmark examples for continuous-time algebraic Riccati equations (version 2.0). SLICOT Working Note 1999-14, Nov. 1999. Available from `http://www.win.tue.nl/niconet/NIC2/reports.html`.

[2] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. SIAM, Philadelphia, PA, third edition, 1999.

[3] W. Arnold, III and A. Laub. Generalized eigenproblem algorithms and software for algebraic Riccati equations. *Proc. IEEE*, 72:1746–1754, 1984.

[4] Z. Bai and J. Demmel. Design of a parallel nonsymmetric eigenroutine toolbox, Part I. In R. S. et al., editor, *Proceedings of the Sixth SIAM Conference on Parallel Processing for Scientific Computing*, pages 391–398. SIAM, Philadelphia, PA, 1993. *See also:* Tech. Report CSD-92-718, Computer Science Division, University of California, Berkeley, CA 94720.

[5] Z. Bai and J. Demmel. Using the matrix sign function to compute invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 19(1):205–225, 1998.

[6] Z. Bai, J. Demmel, J. Dongarra, A. Petitet, H. Robinson, and K. Stanley. The spectral decomposition of nonsymmetric matrices on distributed memory parallel computers. *SIAM J. Sci. Comput.*, 18:1446–1461, 1997.

[7] Z. Bai, J. Demmel, and M. Gu. An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems. *Numer. Math.*, 76(3):279–308, 1997.

[8] F. L. Bauer. Genauigkeitsfragen bei der Lösung linearer Gleichunssysteme. *Z. Angew. Math. Mech.*, 46:409–421, 1966.

[9] P. Benner. *Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems.* Logos–Verlag, Berlin, Germany, 1997. *Also:* Dissertation, Fakultät für Mathematik, TU Chemnitz–Zwickau, 1997.

[10] P. Benner and R. Byers. An arithmetic for matrix pencils. In A. Beghi, L. Finesso, and G. Picci, editors, *Mathematical Theory of Networks and Systems*, pages 573–576, Il Poligrafo, Padova, Italy, 1998.

[11] P. Benner and R. Byers. An arithmetic for rectangular matrix pencils. In O. Gonzalez, editor, *Proc. 1999 IEEE Intl. Symp. CACSD, Kohala Coast-Island of Hawai'i, Hawai'i, USA, August 22–27, 1999* (CD-Rom), pages 75–80, 1999.

[12] P. Benner and R. Byers. Evaluating products of matrix pencils and collapsing matrix products for parallel computation. *Numer. Linear Algebra Appl.*, 8:357–380, 2001.

[13] P. Benner and R. Byers. A structure-preserving method for generalized algebraic Riccati equations based on pencil arithmetic. In *Proc. European Control Conference ECC 2003, Cambridge, UK*, 2003.

[14] P. Benner, V. Mehrmann, V. Sima, S. V. Huffel, and A. Varga. SLICOT - a subroutine library in systems and control theory. In B. Datta, editor, *Applied and Computational Control, Signals, and Circuits*, volume 1, chapter 10, pages 499–539. Birkhäuser, Boston, MA, 1999.

[15] P. Benner, E. Quintana-Ortí, and G. Quintana-Ortí. A portable subroutine library for solving linear control problems on distributed memory computers. In G. Cooperman, E. Jessen, and G. Michler, editors, *Workshop on Wide Area Networks and High Performance Computing, Essen (Germany), September 1998*, Lecture Notes in Control and Information, pages 61–88. Springer-Verlag, Berlin/Heidelberg, Germany, 1999.

[16] S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley. *ScaLAPACK Users' Guide*. SIAM, Philadelphia, PA, 1997. *See also* `http://www.netlib.org/scalapack` and `http://www.netlib.org/scalapack/prototype`.

[17] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical solution of initial-value problems in differential-algebraic equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996. Revised and corrected reprint of the 1989 original.

[18] R. Byers. Numerical stability and instability in matrix sign function based algorithms. In C. Byrnes and A. Lindquist, editors, *Computational and Combinatorial Methods in Systems Theory*, pages 185–200. Elsevier (North-Holland), New York, 1986.

[19] R. Byers. Solving the algebraic Riccati equation with the matrix sign function. *Linear Algebra Appl.*, 85:267–279, 1987.

[20] R. Byers, C. He, and V. Mehrmann. The matrix sign function method and the computation of invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 18(3):615–632, 1997.

[21] D. Chu, L. De Lathauwer, and B. De Moor. A QR-type reduction for computing the SVD of a general matrix product/quotient. *Numer. Math.*, 95:101–121, 2003.

[22] E. Denman and A. Beavers. The matrix sign function and computations in systems. *Appl. Math. Comput.*, 2:63–94, 1976.

[23] F. R. Gantmacher. *The theory of matrices. Vols. 1, 2.* Chelsea Publishing Co., New York, 1959.

[24] J. Gardiner and A. Laub. A generalization of the matrix-sign-function solution for algebraic Riccati equations. *Internat. J. Control*, 44:823–832, 1986.

[25] J. Gardiner and A. Laub. Parallel algorithms for algebraic Riccati equations. *Internat. J. Control*, 54(6):1317–1333, 1991.

[26] S. Godunov. Problem of the dichotomy of the spectrum of a matrix. *Siberian Math. J.*, 27(5):649–660, 1986.

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

[28] G. Golub and C. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, third edition, 1996.

[29] I. N. Herstein. *Noncommutative Rings*. Mathematical Association of America, Washington, DC, 1994.

[30] N. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM Publications, Philadelphia, PA, 1996.

[31] C. Kenney and A. Laub. The matrix sign function. *IEEE Trans. Automat. Control*, 40(8):1330–1348, 1995.

[32] V. Kublanovskaya. AB-algorithm and its modifications for the spectral problem of linear pencils of matrices. *Numer. Math.*, 43:329–342, 1984.

[33] P. Kunkel and V. Mehrmann. A new class of discretization methods for the solution of linear differential-algebraic equations with variable coefficients. *SIAM J. Numer. Anal.*, 33:1941–1961, 1996.

[34] P. Lancaster and L. Rodman. *The Algebraic Riccati Equation*. Oxford University Press, Oxford, 1995.

[35] A. Laub. A Schur method for solving algebraic Riccati equations. *IEEE Trans. Automat. Control*, AC-24:913–921, 1979.

[36] A. Laub and P. Gahinet. Numerical improvements for solving Riccati equations. *IEEE Trans. Automat. Control*, 42(9):1303–1308, 1997.

[37] S. MacLane. *Categories for the Working Mathematician*. Springer-Verlag, Berlin/Heidelberg, Germany, 2nd edition, 1998.

[38] A. Malyshev. Parallel algorithm for solving some spectral problems of linear algebra. *Linear Algebra Appl.*, 188/189:489–520, 1993.

[39] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*. Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[40] C. B. Moler and G. W. Stewart. An algorithm for generalized matrix eigenvalue problems. *SIAM J. Numer. Anal.*, 10:241–256, 1973.

[41] J. Roberts. Linear model reduction and solution of the algebraic Riccati equation by use of the sign function. *Internat. J. Control*, 32:677–687, 1980. (Reprint of Technical Report No. TR-13, CUED/B-Control, Cambridge University, Engineering Department, 1971).

[42] R. Skeel. Scaling for numerical stability in gaussian elimination. *J. Assoc. Comput. Mach.*, 26:494–526, 1979.

[43] G. W. Stewart. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM J. Numer. Anal.*, 17:403–409, 1980.

[44] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, New York, 1990.

[45] J.-G. Sun. Perturbation analysis of singular subspaces of deflating subspaces. *Numer. Math.*, 73:235–263, 1996.

[46] J.-G. Sun. Perturbation analysis of the matrix sign function. *Linear Algebra Appl.*, 250:177–206, 1997.

[47] X. Sun and E. Quintana-Ortí. Spectral division methods for block generalized Schur decompositions. PRISM Working Note #32, 1996. Available from `http://www-c.mcs.anl.gov/Projects/PRISM`.

[48] X. Sun and E. Quintana-Ortí. The generalized Newton iteration for the matrix sign function. *SIAM J. Sci. Statist. Comput.*, 24:669–683, 2002.

[49] P. Van Dooren. The generalized eigenstructure problem in linear system theory. *IEEE Trans. Automat. Control*, AC-26:111–129, 1981.

A Proof of Theorem 3.1

Recall the Kronecker canonical form.

Theorem A.1 *[23] For each pair A, $E \in \mathbb{C}^{m\times n}$, there exist nonsingular matrices $X \in \mathbb{C}^{m\times m}$ and $Y \in \mathbb{C}^{n\times n}$ such that*

$$X(\lambda E - A)Y = \operatorname{diag}(0_{\gamma\delta},\, L_{\epsilon_1},\, L_{\epsilon_2} \ldots, L_{\eta_1}^T, L_{\eta_2}^T, \ldots, \lambda\hat{E} - \hat{A}) \tag{50}$$

where $\lambda\hat{E} - \hat{A}$ is regular and L_ϵ is the $\epsilon \times (\epsilon+1)$ matrix $L_\epsilon = \lambda[I_\epsilon,\, 0_{\epsilon,1}] - [0_{\epsilon,1},\, I_\epsilon]$. Here I_ϵ is the ϵ-by-ϵ identity matrix and $0_{\gamma\delta}$ is the γ-by-δ zero matrix. The regular part of the pencil $\hat{A} - \lambda\hat{E}$ simultaneously takes Weierstraß canonical form [23, Vol.II,§2].

$$\lambda\hat{E} - \hat{A} = \lambda \begin{bmatrix} I_\theta & 0 \\ 0 & N \end{bmatrix} - \begin{bmatrix} J & 0 \\ 0 & I_\psi \end{bmatrix} \tag{51}$$

where $J \in \mathbb{C}^{\theta\times\theta}$ is in Jordan canonical form and $N \in \mathbb{R}^{\psi\times\psi}$ is nilpotent.

We will prove the following slightly stronger version of Theorem 3.1.

Theorem A.2 *If E, $A \in \mathbb{C}^{m\times n}$ have Kronecker canonical form (50)–(51) and $t \neq 0$, then the matrix relation exponential converges in the usual largest-canonical-angle metric,*

$$\exp(E\backslash(At)) = \sum_{k=0}^{\infty} \frac{t^k}{k!}((E\backslash A))^k = (X\mathcal{E}Y\backslash X\mathcal{A}Y),$$

$$\mathcal{E} = \operatorname{diag}([-_\delta],\, [-_{\epsilon_1+1}],\, [-_{\epsilon_2+1}],\, \ldots,\, \begin{bmatrix} I_{\eta_1} \\ 0_{\eta_1,\eta_1} \end{bmatrix},\, \begin{bmatrix} I_{\eta_2} \\ 0_{\eta_2,\eta_2} \end{bmatrix},\, \ldots I_\theta,\, 0_{\psi,\psi}),$$

and

$$\mathcal{A} = \operatorname{diag}([-_\delta],\, [-_{\epsilon_1+1}],\, [-_{\epsilon_2+1}],\, \ldots,\, \begin{bmatrix} 0_{\eta_1,\eta_1} \\ I_{\eta_1} \end{bmatrix},\, \begin{bmatrix} 0_{\eta_2\eta_2} \\ I_{\eta_2} \end{bmatrix},\, \ldots \exp(tJ),\, I_\psi).$$

Here, $[-_p]$ is the 0-by-p empty matrix.

Proof. Let E, $A \in \mathbb{C}^{m\times n}$ have Kronecker canonical form (50)–(51). Consider the partial sum

$$\sum_{k=0}^{K} \frac{t^k}{k!}((E\backslash A))^k \tag{52}$$

for some real numbers $t_0 \neq t_1$. (By convention, $(E\backslash At)^0 = (I\backslash I)$.) By Lemmas 2.15 and 2.16 (52) is block diagonal with diagonal blocks

$$\begin{array}{ll}
\sum_{k=0}^{K} \frac{t^k}{k!} & (0_{\gamma\delta}\backslash 0_{\gamma\delta})^k \\
\sum_{k=0}^{K} \frac{t^k}{k!} & ([I_{\epsilon_j},\, 0_{\epsilon_j,1}]\backslash[0_{\epsilon_j,1},\, I_{\epsilon_j}])^k \\
\sum_{k=0}^{K} \frac{t^k}{k!} & (\begin{bmatrix} I_{\eta_j} \\ 0_{1,\eta_j} \end{bmatrix}\backslash\begin{bmatrix} 0_{1,\eta_j} \\ I_{\eta_j} \end{bmatrix})^k \\
\sum_{k=0}^{K} \frac{t^k}{k!} & (I_\theta\backslash J)^k \\
\sum_{k=0}^{K} \frac{t^k}{k!} & (N\backslash I_\psi)^k.
\end{array}$$

It suffices to consider each diagonal block separately. A direct application of the definitions (2) and (12) show that for $K > 0$,

$$\sum_{k=0}^{K} \frac{t^k}{k!}(0_{\gamma\delta}\backslash 0_{\gamma\delta})^k = (I\backslash I) + \mathbb{C}^\delta \times \mathbb{C}^\delta + \mathbb{C}^\delta \times \mathbb{C}^\delta \cdots = \mathbb{C}^\delta \times \mathbb{C}^\delta = ([-_\delta]\backslash[-_\delta]).$$

is independent of K. Hence, the matrix relation sum converges trivially and

$$\sum_{k=0}^{\infty} \frac{t^k}{k!}(0_{\gamma\delta}\backslash 0_{\gamma\delta})^k = (0_{\gamma\delta}\backslash 0_{\gamma\delta}) = ([-_\delta]\backslash[-_\delta])$$

From (2), $\mathrm{Dom}([I_{\epsilon_j},\, 0_{\epsilon_j,1}]\backslash[0_{\epsilon_j,1},\, I_{\epsilon_j}])^\ell = \mathbb{C}^{\epsilon_j+1}$, and $z \in ([I_{\epsilon_j},\, 0_{\epsilon_j,1}]\backslash[0_{\epsilon_j,1},\, I_{\epsilon_j}])^k x$ if and only if there exist vectors $w_\ell \in \mathbb{C}^{\epsilon_j+1}$ for $\ell = 0, 1, \ldots k$, satisfying $w_0 = x$, $[I_{\epsilon_j},\, 0_{\epsilon_j,1}]w_\ell = [0_{\epsilon_j,1},\, I_{\epsilon_j}]w_{\ell-1}$, and $z = w_k$. Note that for $p = 1, 2, \ldots \epsilon_j$, $w_{p\ell} = w_{p+1,\ell-1}$ and $w_{\epsilon_j+1,\ell}$ may be chosen arbitrarily. So, for $k > \epsilon_j$, every entry of $z = w_k$ may be chosen arbitrarily. Hence, for $k > \epsilon_j$, $([I_{\epsilon_j},\, 0_{\epsilon_j,1}]\backslash[0_{\epsilon_j,1},\, I_{\epsilon_j}])^k = \mathbb{C}^{\epsilon_j+1} \times \mathbb{C}^{\epsilon_j+1} = (0_{1,\epsilon_j+1}\backslash 0_{1,\epsilon_j+1}). = ([-_{\epsilon_j+1}]\backslash[-_{\epsilon_j+1}])$. It follows from (12) that for $K > \epsilon_j$, $\sum_{k=0}^{K} \frac{t^k}{k!}([I_{\epsilon_j},\, 0_{\epsilon_j,1}]\backslash[0_{\epsilon_j,1},\, I_{\epsilon_j}])^k = \mathbb{C}^{\epsilon_j+1} \times \mathbb{C}^{\epsilon_j+1} = ([-_{\epsilon_j+1}]\backslash[-_{\epsilon_j+1}])$ is constant. Hence, the matrix relation sum converges trivially and

$$\sum_{k=0}^{\infty} \frac{t^k}{k!}([I_{\epsilon_j},\, 0_{\epsilon_j,1}]\backslash[0_{\epsilon_j,1},\, I_{\epsilon_j}])^k = \mathbb{C}^{\epsilon_j+1} \times \mathbb{C}^{\epsilon_j+1} = ([-_{\epsilon_j+1}]\backslash[-_{\epsilon_j+1}]).$$

Using (2) again, $z \in (\begin{bmatrix} I_{\eta_j} \\ 0_{1,\eta_j} \end{bmatrix}\backslash\begin{bmatrix} 0_{1,\eta_j} \\ I_{\eta_j} \end{bmatrix})^k x$ if and only if there exist vectors w_ℓ, for $\ell = 0, 2, 3, \ldots, k$ satisfying $w_0 = x$, $\begin{bmatrix} I_{\eta_j} \\ 0_{1,\eta_j} \end{bmatrix} w_\ell = \begin{bmatrix} 0_{1,\eta_j} \\ I_{\eta_j} \end{bmatrix} w_{\ell-1}$, and $z = w_k$. Note that for each ℓ, $1 \le \ell \le k$, $w_{1\ell} = 0$, $0 = w_{\eta_j,\ell-1}$ and for $p = 2, 3, \ldots, \eta_j$, $w_{p\ell} = w_{p-1,\ell-1}$. For $k \ge \eta_j$, the only solution is $0 = x = w_0 = w_1 = \cdots = w_k = z$, i.e., $(\begin{bmatrix} I_{\eta_j} \\ 0_{1,\eta_j} \end{bmatrix}\backslash\begin{bmatrix} 0_{1,\eta_j} \\ I_{\eta_j} \end{bmatrix})^k = \{0, 0\} = (\begin{bmatrix} I_{\eta_j} \\ 0_{\eta_j,\eta_j} \end{bmatrix}\backslash\begin{bmatrix} 0_{\eta_j\eta_j} \\ I_{\eta_j} \end{bmatrix})$. It follows that $\sum_{k=0}^{K} \frac{t^k}{k!}(\begin{bmatrix} I_{\eta_j} \\ 0_{1,\eta_j} \end{bmatrix}\backslash\begin{bmatrix} 0_{1,\eta_j} \\ I_{\eta_j} \end{bmatrix})^k = \{0, 0\}$ independent of $K \ge \eta_j$. Hence, the matrix relation sum converges trivially and

$$\sum_{k=0}^{\infty} \frac{t^k}{k!}(\begin{bmatrix} I_{\eta_j} \\ 0_{1,\eta_j} \end{bmatrix}\backslash\begin{bmatrix} 0_{1,\eta_j} \\ I_{\eta_j} \end{bmatrix})^k = \{0, 0\} = (\begin{bmatrix} I_{\eta_j} \\ 0_{\eta_j\eta_j} \end{bmatrix}\backslash\begin{bmatrix} 0_{\eta_j\eta_j} \\ I_{\eta_j} \end{bmatrix}).$$

Because $(I_\theta \backslash J)$ is a linear transformation on $\mathbb{C}^n$, it follows that

$$\sum_{k=0}^{K} \frac{t^k}{k!}(I_\theta \backslash J)^k = (I_\theta \backslash \sum_{k=0}^{K} \frac{t^k}{k!} J^k) = \text{null}\left[-I_\theta, \sum_{k=0}^{K} \frac{t^k}{k!} J^k\right]$$

The right-hand equality follows from (1). Taking limits as K tends to infinity, we have

$$\sum_{k=0}^{\infty} \frac{t^k}{k!}(I_\theta \backslash J)^k = \text{null}\left[-I_\theta, \exp(Jt)\right] = (I_\theta \backslash \exp(Jt)).$$

An application of (2) or Theorem 2.3 shows that $(N \backslash I_\psi)^k = (N^k \backslash I_\psi)$. Because N is nilpotent, for $k \geq \psi$, $(N \backslash I_\psi)^k = (0_{\psi,\psi} \backslash I_\psi) = \mathbb{C}^\psi \times \{0\}$. Now, (12) implies that for K large enough, $\sum_{k=0}^{K} \frac{t^k}{k!}(N \backslash I_\psi)^k = (0_\psi \backslash I_\psi)$. So, the matrix relation sum converges trivially and

$$\sum_{k=0}^{\infty} \frac{t^k}{k!}(N \backslash I_\psi)^k = \mathbb{C}^\psi \times \{0\} = (0_\psi \backslash I_\psi).$$

□