The matrix Sylvester equation for congruence

Froilán M. Dopico

ICMAT and Departamento de Matemáticas, Universidad Carlos III de Madrid, Spain

joint work with Fernando De Terán, Nathan Guillery, Daniel Montealegre, and Nicolás Reyes

School of Mathematics, University of Edinburgh, Scotland
February 7, 2013

Thanks to Edinburgh Mathematical Society

Abstract (I)

The **matrix Sylvester equation**

$$AX - XB = C, \qquad A \in \mathbb{C}^{m\times m},\ B \in \mathbb{C}^{n\times n} \text{ are given}$$

is, probably, the most famous matrix equation. It arises

- as a step in algorithms for computing eigenvalues/vectors;
- in the perturbation theory of invariant subspaces of matrices;
- in the characterization of the matrices that commute with a given matrix $AX = XA$.
- Its particular case, the Lyapunov equation,

$$AX + XA^* = C$$

arises in control and linear system theory and in stability theory...

- Properties of Sylvester eq. are well-known and are presented in standard books on Matrix Analysis.
- Numerical methods for solution are also well-known.

Abstract (I)

The **matrix Sylvester equation**

$$AX - XB = C, \qquad A \in \mathbb{C}^{m\times m},\ B \in \mathbb{C}^{n\times n} \text{ are given}$$

is, probably, the most famous matrix equation. It arises

- as a step in algorithms for computing eigenvalues/vectors;
- in the perturbation theory of invariant subspaces of matrices;
- in the characterization of the matrices that commute with a given matrix $AX = XA$.
- Its particular case, the Lyapunov equation,

$$AX + XA^* = C$$

  arises in control and linear system theory and in stability theory...

- Properties of Sylvester eq. are well-known and are presented in standard books on Matrix Analysis.
- Numerical methods for solution are also well-known.

Abstract (I)

The **matrix Sylvester equation**

$$AX - XB = C, \qquad A \in \mathbb{C}^{m\times m},\ B \in \mathbb{C}^{n\times n} \text{ are given}$$

is, probably, the most famous matrix equation. It arises

- as a step in algorithms for computing eigenvalues/vectors;
- in the perturbation theory of invariant subspaces of matrices;
- in the characterization of the matrices that commute with a given matrix $AX = XA$.
- Its particular case, the Lyapunov equation,

$$AX + XA^* = C$$

arises in control and linear system theory and in stability theory...

- Properties of Sylvester eq. are well-known and are presented in standard books on Matrix Analysis.
- Numerical methods for solution are also well-known.

The **matrix Sylvester equation**

$$AX - XB = C, \qquad A \in \mathbb{C}^{m\times m},\ B \in \mathbb{C}^{n\times n} \text{ are given}$$

is, probably, the most famous matrix equation. It arises

- as a step in algorithms for computing eigenvalues/vectors;
- in the perturbation theory of invariant subspaces of matrices;
- in the characterization of the matrices that commute with a given matrix $AX = XA$.
- Its particular case, the Lyapunov equation,

$$AX + XA^* = C$$

arises in control and linear system theory and in stability theory...

- Properties of Sylvester eq. are well-known and are presented in standard books on Matrix Analysis.
- Numerical methods for solution are also well-known.

Abstract (I)

The **matrix Sylvester equation**

$$AX - XB = C, \qquad A \in \mathbb{C}^{m\times m},\ B \in \mathbb{C}^{n\times n} \text{ are given}$$

is, probably, the most famous matrix equation. It arises

- as a step in algorithms for computing eigenvalues/vectors;
- in the perturbation theory of invariant subspaces of matrices;
- in the characterization of the matrices that commute with a given matrix $AX = XA$.
- Its particular case, the Lyapunov equation,

  $$AX + XA^* = C$$

  arises in control and linear system theory and in stability theory...

▶ Properties of Sylvester eq. are well-known and are presented in standard books on Matrix Analysis.

▶ Numerical methods for solution are also well-known.

The **matrix Sylvester equation**

$$AX - XB = C, \qquad A \in \mathbb{C}^{m\times m},\ B \in \mathbb{C}^{n\times n} \text{ are given}$$

is, probably, the most famous matrix equation. It arises

- as a step in algorithms for computing eigenvalues/vectors;
- in the perturbation theory of invariant subspaces of matrices;
- in the characterization of the matrices that commute with a given matrix $AX = XA$.
- Its particular case, the Lyapunov equation,

$$AX + XA^* = C$$

  arises in control and linear system theory and in stability theory...

- Properties of Sylvester eq. are well-known and are presented in standard books on Matrix Analysis.
- Numerical methods for solution are also well-known.

Abstract (II)

Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

$$\boxed{AX + X^T B = C, \qquad A \in \mathbb{C}^{m\times n},\ B \in \mathbb{C}^{n\times m}}$$

has received considerable attention as a consequence of its relationship with palindromic eigenvalue problems

$$\boxed{Gx = -\lambda G^T x, \qquad G \in \mathbb{C}^{n\times n}.}$$

These problems arise in a number of applications:

- the mathematical modelling and numerical simulation of the behavior of periodic surface acoustic wave filters (2002, 2006);
- the analysis of rail track vibrations produced by high speed trains (2004, 2006, 2009);
- discrete-time optimal control problems (2008).

The spectrum of palindromic eigenproblems has the symmetry $(\lambda, 1/\lambda)$.

Abstract (II)

Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

$$AX + X^T B = C, \qquad A \in \mathbb{C}^{m\times n},\ B \in \mathbb{C}^{n\times m}$$

has received considerable attention as a consequence of its relationship with palindromic eigenvalue problems

$$Gx = -\lambda G^T x, \qquad G \in \mathbb{C}^{n\times n}.$$

These problems arise in a number of applications:

- the mathematical modelling and numerical simulation of the behavior of periodic surface acoustic wave filters (2002, 2006);
- the analysis of rail track vibrations produced by high speed trains (2004, 2006, 2009);
- discrete-time optimal control problems (2008).

The spectrum of palindromic eigenproblems has the symmetry $(\lambda, 1/\lambda)$.

Abstract (II)

Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

$$AX + X^T B = C, \qquad A \in \mathbb{C}^{m\times n},\ B \in \mathbb{C}^{n\times m}$$

has received considerable attention as a consequence of its relationship with palindromic eigenvalue problems

$$Gx = -\lambda G^T x, \qquad G \in \mathbb{C}^{n\times n}.$$

These problems arise in a number of applications:

- the mathematical modelling and numerical simulation of the behavior of periodic surface acoustic wave filters (2002, 2006);
- the analysis of rail track vibrations produced by high speed trains (2004, 2006, 2009);
- discrete-time optimal control problems (2008).

The spectrum of palindromic eigenproblems has the symmetry $(\lambda, 1/\lambda)$.

Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

$$AX + X^T B = C, \qquad A \in \mathbb{C}^{m\times n},\ B \in \mathbb{C}^{n\times m}$$

has received considerable attention as a consequence of its relationship with palindromic eigenvalue problems

$$Gx = -\lambda G^T x, \qquad G \in \mathbb{C}^{n\times n}.$$

These problems arise in a number of applications:

- the mathematical modelling and numerical simulation of the behavior of periodic surface acoustic wave filters (2002, 2006);
- the analysis of rail track vibrations produced by high speed trains (2004, 2006, 2009);
- discrete-time optimal control problems (2008).

The spectrum of palindromic eigenproblems has the symmetry $(\lambda, 1/\lambda)$.

Abstract (II)

Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

$$AX + X^T B = C, \qquad A \in \mathbb{C}^{m\times n},\ B \in \mathbb{C}^{n\times m}$$

has received considerable attention as a consequence of its relationship with palindromic eigenvalue problems

$$Gx = -\lambda G^T x, \qquad G \in \mathbb{C}^{n\times n}.$$

These problems arise in a number of applications:

- the mathematical modelling and numerical simulation of the behavior of periodic surface acoustic wave filters (2002, 2006);
- the analysis of rail track vibrations produced by high speed trains (2004, 2006, 2009);
- discrete-time optimal control problems (2008).

The spectrum of palindromic eigenproblems has the symmetry $(\lambda, 1/\lambda)$.

$$AX - XB = C \quad vs. \quad AX + X^T B = C$$

The "transposed second X" makes the study of both equations very different.
Not many? references available for T-Sylvester equation.

In this talk, I will revise my research work on Sylvester equation for congruence that has been published in

- F. De Terán & D, *The solution of the equation* $XA + AX^T = 0$ *and its application to the theory of orbits*, Lin. Alg. Appl. 434 (2011) 44-67.
- F. De Terán & D, *The equation* $XA + AX^* = 0$ *and the dimension of *congruence orbits*, Electronic Journal of Linear Algebra, 22 (2011) 448-465.
- F. De Terán & D, *Consistency and efficient solution of the Sylvester equation for* $\star$*-congruence*, Elect. Journal of Linear Algebra 22 (2011) 849-863.
- F. De Terán & D & N. Guillery & D. Montealegre & N. Reyes, *The solution of the equation* $AX + X^\star B = 0$, published electronically in Lin. Alg. Appl., 2013.

This talk is my "personal journey" trough Sylvester eq. for congruence.

Abstract (III)

$$AX - XB = C \quad vs. \quad AX + X^T B = C$$

The "transposed second X" makes the study of both equations very different. Not many? references available for T-Sylvester equation.

In this talk, I will revise my research work on Sylvester equation for congruence that has been published in

- F. De Terán & D, *The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits*, Lin. Alg. Appl. 434 (2011) 44-67.
- F. De Terán & D, *The equation $XA + AX^* = 0$ and the dimension of *congruence orbits*, Electronic Journal of Linear Algebra, 22 (2011) 448-465.
- F. De Terán & D, *Consistency and efficient solution of the Sylvester equation for $\star$-congruence*, Elect. Journal of Linear Algebra 22 (2011) 849-863.
- F. De Terán & D & N. Guillery & D. Montealegre & N. Reyes, *The solution of the equation $AX + X^\star B = 0$*, published electronically in Lin. Alg. Appl., 2013.

This talk is my "personal journey" trough Sylvester eq. for congruence.

$$AX - XB = C \quad vs. \quad AX + X^T B = C$$

The “transposed second X” makes the study of both equations very different. Not many? references available for T-Sylvester equation.

In this talk, I will revise my research work on Sylvester equation for congruence that has been published in

- F. De Terán & D, *The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits*, Lin. Alg. Appl. 434 (2011) 44-67.
- F. De Terán & D, *The equation $XA + AX^* = 0$ and the dimension of *congruence orbits*, Electronic Journal of Linear Algebra, 22 (2011) 448-465.
- F. De Terán & D, *Consistency and efficient solution of the Sylvester equation for $\star$-congruence*, Elect. Journal of Linear Algebra 22 (2011) 849-863.
- F. De Terán & D & N. Guillery & D. Montealegre & N. Reyes, *The solution of the equation $AX + X^\star B = 0$*, published electronically in Lin. Alg. Appl., 2013.

This talk is my “personal journey” trough Sylvester eq. for congruence.

$$AX - XB = C \quad vs. \quad AX + X^T B = C$$

The "transposed second X" makes the study of both equations very different. Not many? references available for T-Sylvester equation.

In this talk, I will revise my research work on Sylvester equation for congruence that has been published in

- F. De Terán & D, *The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits*, Lin. Alg. Appl. 434 (2011) 44-67.
- F. De Terán & D, *The equation $XA + AX^* = 0$ and the dimension of *congruence orbits*, Electronic Journal of Linear Algebra, 22 (2011) 448-465.
- F. De Terán & D, *Consistency and efficient solution of the Sylvester equation for $\star$-congruence*, Elect. Journal of Linear Algebra 22 (2011) 849-863.
- F. De Terán & D & N. Guillery & D. Montealegre & N. Reyes, *The solution of the equation $AX + X^\star B = 0$*, published electronically in Lin. Alg. Appl., 2013.

This talk is my "personal journey" trough Sylvester eq. for congruence.

$$AX - XB = C \quad vs. \quad AX + X^* B = C$$

The "transposed second X" makes the study of both equations very different. Not many? references available for T-Sylvester equation.

In this talk, I will revise my work on Sylvester equation for congruence that has been presented in

- F. De Terán & D, *The solution of the equation* $XA + AX^T = 0$ *and its application to the theory of orbits*, Lin. Alg. Appl. 434 (2011) 44-67.
- F. De Terán & D, *The equation* $XA + AX^* = 0$ *and the dimension of *congruence orbits*, Electronic Journal of Linear Algebra, 22 (2011) 448-465.
- F. De Terán & D, *Consistency and efficient solution of the Sylvester equation for* $\star$*-congruence*, Elect. Journal of Linear Algebra, 22 (2011) 849-863.
- F. De Terán & D & N. Guillery & D. Montealegre & N. Reyes, *The solution of the equation* $AX + X^\star B = 0$, published electronically in Lin. Alg. Appl., 2013.

Both interesting, both related, but DIFFERENT!!!

$$AX - XB = C \quad vs. \quad AX + X^T B = C$$

The “transposed second X” makes the study of both equations very different. Not many? references available for T-Sylvester equation.

In this talk, I will revise my work on Sylvester equation for congruence that has been presented in

- F. De Terán & D, *The solution of the equation* $XA + AX^T = 0$ *and its application to the theory of orbits*, Lin. Alg. Appl. 434 (2011) 44-67.
- F. De Terán & D, *The equation* $XA + AX^* = 0$ *and the dimension of *congruence orbits*, Electronic Journal of Linear Algebra, 22 (2011) 448-465.
- F. De Terán & D, *Consistency and efficient solution of the Sylvester equation for* $\star$*-congruence*, Elect. Journal of Linear Algebra, 22 (2011) 849-863.
- F. De Terán & D & N. Guillery & D. Montealegre & N. Reyes, *The solution of the equation* $AX + X^\star B = 0$, published electronically in Lin. Alg. Appl., 2013.

In this talk for simplicity mostly T-case is considered,

$$AX - {\color{red}X}B = C \quad vs. \quad AX + {\color{red}X^\star}B = C$$

The "transposed second X" makes the study of both equations very different. Not many? references available for T-Sylvester.

In this talk, I will revise my work on Sylvester equation for congruence that has been presented in

- F. De Terán & D, *The solution of the equation* $XA + AX^T = 0$ *and its application to the theory of orbits*, Lin. Alg. Appl. 434 (2011) 44-67.
- F. De Terán & D, *The equation* $XA + AX^* = 0$ *and the dimension of *congruence orbits*, Electronic Journal of Linear Algebra, 22 (2011) 448-465.
- F. De Terán & D, *Consistency and efficient solution of the Sylvester equation for* $\star$*-congruence*, Elect. Journal of Linear Algebra, 22 (2011) 849-863.
- F. De Terán & D & N. Guillery & D. Montealegre & N. Reyes, *The solution of the equation* $AX + X^\star B = 0$, published electronically in Lin. Alg. Appl., 2013.

but sometimes both cases simultaneously: $\star = T$ **or** $*$

Outline

1. **Previous and related work**

2. **The equation $AX^T + XA = 0$**
   - Motivation: Orbits and the computation of canonical forms
   - Strategy for solving $AX^T + XA = 0$
   - The canonical form for congruence
   - The solution of $AX^T + XA = 0$
   - Generic canonical structure for congruence

3. **The general equation $AX + X^\star B = C$**
   - Motivation
   - Consistency of the Sylvester equation for $\star$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions

4. **General solution of $AX + X^\star B = 0$**

5. **Conclusions**

$$AX + X^\star A^\star = C \qquad (\star = T \text{ or } *)$$

arises in time-invariant Hamiltonian systems and R-matrix treatment of completely integrable mechanical systems.

- Hodges, Ann. Mat. Pura Appl., (1957): A nonsingular. Over finite fields.
- Taussky & Wielandt, Arch. Rational Mech. Anal., (1962): functions $G(X) = AX + X^\star A^\star$; its eigenvalues. Algebraically closed fields.
- Ballantine, Lin. Alg. Appl., (1969): $AX + X^*A = C$, with A Hermitian and positive definite. Neccesary and sufficient conditions for consistency.
- Lancaster & Rozsa, SIAM J. Alg. Disc. Meth., (1983): Necessary and sufficient conditions for consistency in terms of a rank factorization of A. Closed formula for general solution using submatrices of C and rank factorization of A, and dimension of the solution space.
- Braden, SIAM J. Matrix Anal. Appl., (1998): Similar results but in terms of projectors and generalized inverses $AGA = A$.
- Djordjević, J. Comput. Appl. Math., (2007): Extends Lancaster & Rozsa and Braden to A, C, X bounded linear operators on Hilbert spaces.

$$\boxed{AX + X^\star A^\star = C \qquad (\star = T \text{ or } *)}$$

arises in time-invariant Hamiltonian systems and R-matrix treatment of completely integrable mechanical systems.

- Hodges, Ann. Mat. Pura Appl., (1957): A nonsingular. Over finite fields.
- Taussky & Wielandt, Arch. Rational Mech. Anal., (1962): functions $G(X) = AX + X^\star A^\star$; its eigenvalues. Algebraically closed fields.
- Ballantine, Lin. Alg. Appl., (1969): $AX + X^* A = C$, with A Hermitian and positive definite. Neccesary and sufficient conditions for consistency.
- Lancaster & Rozsa, SIAM J. Alg. Disc. Meth., (1983): Necessary and sufficient conditions for consistency in terms of a rank factorization of A. Closed formula for general solution using submatrices of C and rank factorization of A, and dimension of the solution space.
- Braden, SIAM J. Matrix Anal. Appl., (1998): Similar results but in terms of projectors and generalized inverses $AGA = A$.
- Djordjević, J. Comput. Appl. Math., (2007): Extends Lancaster & Rozsa and Braden to A, C, X bounded linear operators on Hilbert spaces.

$$\boxed{AX + X^\star A^\star = C \qquad (\star = T \text{ or } *)}$$

arises in time-invariant Hamiltonian systems and R-matrix treatment of completely integrable mechanical systems.

- Hodges, Ann. Mat. Pura Appl., (1957): A nonsingular. Over finite fields.
- Taussky & Wielandt, Arch. Rational Mech. Anal., (1962): functions $G(X) = AX + X^\star A^\star$; its eigenvalues. Algebraically closed fields.
- Ballantine, Lin. Alg. Appl., (1969): $AX + X^* A = C$, with A Hermitian and positive definite. Neccesary and sufficient conditions for consistency.
- Lancaster & Rozsa, SIAM J. Alg. Disc. Meth., (1983): Necessary and sufficient conditions for consistency in terms of a rank factorization of A. Closed formula for general solution using submatrices of C and rank factorization of A, and dimension of the solution space.
- Braden, SIAM J. Matrix Anal. Appl., (1998): Similar results but in terms of projectors and generalized inverses $AGA = A$.
- Djordjević, J. Comput. Appl. Math., (2007): Extends Lancaster & Rozsa and Braden to A, C, X bounded linear operators on Hilbert spaces.

$$\boxed{AX + X^\star A^\star = C \qquad (\star = T \text{ or } *)}$$

arises in time-invariant Hamiltonian systems and R-matrix treatment of completely integrable mechanical systems.

- Hodges, Ann. Mat. Pura Appl., (1957): A nonsingular. Over finite fields.
- Taussky & Wielandt, Arch. Rational Mech. Anal., (1962): functions $G(X) = AX + X^\star A^\star$; its eigenvalues. Algebraically closed fields.
- Ballantine, Lin. Alg. Appl., (1969): $AX + X^* A = C$, with A Hermitian and positive definite. Neccesary and sufficient conditions for consistency.
- Lancaster & Rozsa, SIAM J. Alg. Disc. Meth., (1983): Necessary and sufficient conditions for consistency in terms of a rank factorization of A. Closed formula for general solution using submatrices of C and rank factorization of A, and dimension of the solution space.
- Braden, SIAM J. Matrix Anal. Appl., (1998): Similar results but in terms of projectors and generalized inverses $AGA = A$.
- Djordjević, J. Comput. Appl. Math., (2007): Extends Lancaster & Rozsa and Braden to A, C, X bounded linear operators on Hilbert spaces.

An important particular case of $AX + X^\star B = C$

$$AX + X^\star A^\star = C \qquad (\star = T \text{ or } *)$$

arises in time-invariant Hamiltonian systems and R-matrix treatment of completely integrable mechanical systems.

- Hodges, Ann. Mat. Pura Appl., (1957): A nonsingular. Over finite fields.
- Taussky & Wielandt, Arch. Rational Mech. Anal., (1962): functions $G(X) = AX + X^\star A^\star$; its eigenvalues. Algebraically closed fields.
- Ballantine, Lin. Alg. Appl., (1969): $AX + X^*A = C$, with A Hermitian and positive definite. Neccesary and sufficient conditions for consistency.
- Lancaster & Rozsa, SIAM J. Alg. Disc. Meth., (1983): Necessary and sufficient conditions for consistency in terms of a rank factorization of A. Closed formula for general solution using submatrices of C and rank factorization of A, and dimension of the solution space.
- Braden, SIAM J. Matrix Anal. Appl., (1998): Similar results but in terms of projectors and generalized inverses $AGA = A$.
- Djordjević, J. Comput. Appl. Math., (2007): Extends Lancaster & Rozsa and Braden to A, C, X bounded linear operators on Hilbert spaces.

$$\boxed{AX + X^\star A^\star = C \qquad (\star = T \text{ or } *)}$$

arises in time-invariant Hamiltonian systems and R-matrix treatment of completely integrable mechanical systems.

- Hodges, Ann. Mat. Pura Appl., (1957): A nonsingular. Over finite fields.
- Taussky & Wielandt, Arch. Rational Mech. Anal., (1962): functions $G(X) = AX + X^\star A^\star$; its eigenvalues. Algebraically closed fields.
- Ballantine, Lin. Alg. Appl., (1969): $AX + X^* A = C$, with A Hermitian and positive definite. Neccesary and sufficient conditions for consistency.
- Lancaster & Rozsa, SIAM J. Alg. Disc. Meth., (1983): Necessary and sufficient conditions for consistency in terms of a rank factorization of A. Closed formula for general solution using submatrices of C and rank factorization of A, and dimension of the solution space.
- Braden, SIAM J. Matrix Anal. Appl., (1998): Similar results but in terms of projectors and generalized inverses $AGA = A$.
- Djordjević, J. Comput. Appl. Math., (2007): Extends Lancaster & Rozsa and Braden to A, C, X bounded linear operators on Hilbert spaces.

$$AX + X^\star A^\star = C \qquad (\star = T \text{ or } *)$$

arises in time-invariant Hamiltonian systems and R-matrix treatment of completely integrable mechanical systems.

- Hodges, Ann. Mat. Pura Appl., (1957): A nonsingular. Over finite fields.
- Taussky & Wielandt, Arch. Rational Mech. Anal., (1962): functions $G(X) = AX + X^\star A^\star$; its eigenvalues. Algebraically closed fields.
- Ballantine, Lin. Alg. Appl., (1969): $AX + X^* A = C$, with A Hermitian and positive definite. Neccesary and sufficient conditions for consistency.
- Lancaster & Rozsa, SIAM J. Alg. Disc. Meth., (1983): Necessary and sufficient conditions for consistency in terms of a rank factorization of A. Closed formula for general solution using submatrices of C and rank factorization of A, and dimension of the solution space.
- Braden, SIAM J. Matrix Anal. Appl., (1998): Similar results but in terms of projectors and generalized inverses $AGA = A$.
- Djordjević, J. Comput. Appl. Math., (2007): Extends Lancaster & Rozsa and Braden to A, C, X bounded linear operators on Hilbert spaces.

$$AX + X^\star A = 0, \qquad A \in \mathbb{C}^{n\times n} \qquad (\star = T \text{ or } *)$$

- De Terán & D., Lin. Alg. Appl. and Elec. J. Lin. Alg., (2011):
  - General solution obtained in the spirit of classical methods of solution of standard Sylvester equation.
  - Related to the theory of orbits by the action of congruence.

References for general equation $AX + X^{\star}B = C$

- Wimmer, Lin. Alg. Appl., (1994): Necessary and sufficient conditions for **consistency**. Complex matrices
- Byers & Kressner, SIAM J. Matrix Anal. Appl., (2006): Necessary and sufficient conditions for **unique solution** for $\star = T$. Complex matrices.
- Kressner & Schröder & Watkins, Numer. Algor., (2009): Same for $\star = *$.
- De Terán & D., Elec. J. Lin. Alg., (2011): Efficient **algorithm** for computing the solution when it is unique.
- Chiang & Chu & Lin, Appl. Math. Comp., (2012): Repeat some of the results above in other language and considers more general eqs. $AXB + CX^{\star}D = E$.
- De Terán & D & Guillery & Montealegre & Reyes, Lin. Alg. Appl., (2013): General solution in the homogeneous case.
- Other papers Piao-Zhang-Wang, (2007) (involved, formula for solution gen. inverses, under certain assumptions); Cvetković-Ilić, (2008) (operators with certain restrictions); Ikramov (2009) (conditions for unique solvability); Vorontsov-Ikramov (2011) (algorithm).

References for general equation $AX + X^\star B = C$

- Wimmer, Lin. Alg. Appl., (1994): Necessary and sufficient conditions for **consistency**. Complex matrices
- Byers & Kressner, SIAM J. Matrix Anal. Appl., (2006): Necessary and sufficient conditions for **unique solution** for $\star = T$. Complex matrices.
- Kressner & Schröder & Watkins, Numer. Algor., (2009): Same for $\star = *$.
- De Terán & D., Elec. J. Lin. Alg., (2011): Efficient **algorithm** for computing the solution when it is unique.
- Chiang & Chu & Lin, Appl. Math. Comp., (2012): Repeat some of the results above in other language and considers more general eqs. $AXB + CX^\star D = E$.
- De Terán & D & Guillery & Montealegre & Reyes, Lin. Alg. Appl., (2013): General solution in the homogeneous case.
- Other papers Piao-Zhang-Wang, (2007) (involved, formula for solution gen. inverses, under certain assumptions); Cvetković-Ilić, (2008) (operators with certain restrictions); Ikramov (2009) (conditions for unique solvability); Vorontsov-Ikramov (2011) (algorithm).

References for general equation $AX + X^\star B = C$

- Wimmer, Lin. Alg. Appl., (1994): Necessary and sufficient conditions for **consistency**. Complex matrices
- Byers & Kressner, SIAM J. Matrix Anal. Appl., (2006): Necessary and sufficient conditions for **unique solution** for $\star = T$. Complex matrices.
- Kressner & Schröder & Watkins, Numer. Algor., (2009): Same for $\star = *$.
- De Terán & D., Elec. J. Lin. Alg., (2011): Efficient **algorithm** for computing the solution when it is unique.
- Chiang & Chu & Lin, Appl. Math. Comp., (2012): Repeat some of the results above in other language and considers more general eqs. $AXB + CX^\star D = E$.
- De Terán & D & Guillery & Montealegre & Reyes, Lin. Alg. Appl., (2013): General solution in the homogeneous case.
- Other papers Piao-Zhang-Wang, (2007) (involved, formula for solution gen. inverses, under certain assumptions); Cvetković-Ilić, (2008) (operators with certain restrictions); Ikramov (2009) (conditions for unique solvability); Vorontsov-Ikramov (2011) (algorithm).

References for general equation $AX + X^\star B = C$

- Wimmer, Lin. Alg. Appl., (1994): Necessary and sufficient conditions for **consistency**. Complex matrices
- Byers & Kressner, SIAM J. Matrix Anal. Appl., (2006): Necessary and sufficient conditions for **unique solution** for $\star = T$. Complex matrices.
- Kressner & Schröder & Watkins, Numer. Algor., (2009): Same for $\star = *$.
- De Terán & D., Elec. J. Lin. Alg., (2011): Efficient **algorithm** for computing the solution when it is unique.
- Chiang & Chu & Lin, Appl. Math. Comp., (2012): Repeat some of the results above in other language and considers more general eqs. $AXB + CX^\star D = E$.
- De Terán & D & Guillery & Montealegre & Reyes, Lin. Alg. Appl., (2013): General solution in the homogeneous case.
- Other papers Piao-Zhang-Wang, (2007) (involved, formula for solution gen. inverses, under certain assumptions); Cvetković-Ilić, (2008) (operators with certain restrictions); Ikramov (2009) (conditions for unique solvability); Vorontsov-Ikramov (2011) (algorithm).

- Wimmer, Lin. Alg. Appl., (1994): Necessary and sufficient conditions for **consistency**. Complex matrices
- Byers & Kressner, SIAM J. Matrix Anal. Appl., (2006): Necessary and sufficient conditions for **unique solution** for $\star = T$. Complex matrices.
- Kressner & Schröder & Watkins, Numer. Algor., (2009): Same for $\star = *$.
- De Terán & D., Elec. J. Lin. Alg., (2011): Efficient **algorithm** for computing the solution when it is unique.
- Chiang & Chu & Lin, Appl. Math. Comp., (2012): Repeat some of the results above in other language and considers more general eqs. $AXB + CX^\star D = E$.
- De Terán & D & Guillery & Montealegre & Reyes, Lin. Alg. Appl., (2013): General solution in the homogeneous case.
- Other papers Piao-Zhang-Wang, (2007) (involved, formula for solution gen. inverses, under certain assumptions); Cvetković-Ilić, (2008) (operators with certain restrictions); Ikramov (2009) (conditions for unique solvability); Vorontsov-Ikramov (2011) (algorithm).

References for general equation $AX + X^\star B = C$

- Wimmer, Lin. Alg. Appl., (1994): Necessary and sufficient conditions for **consistency**. Complex matrices
- Byers & Kressner, SIAM J. Matrix Anal. Appl., (2006): Necessary and sufficient conditions for **unique solution** for $\star = T$. Complex matrices.
- Kressner & Schröder & Watkins, Numer. Algor., (2009): Same for $\star = *$.
- De Terán & D., Elec. J. Lin. Alg., (2011): Efficient **algorithm** for computing the solution when it is unique.
- Chiang & Chu & Lin, Appl. Math. Comp., (2012): Repeat some of the results above in other language and considers more general eqs. $AXB + CX^\star D = E$.
- De Terán & D & Guillery & Montealegre & Reyes, Lin. Alg. Appl., (2013): General solution in the homogeneous case.
- Other papers Piao-Zhang-Wang, (2007) (involved, formula for solution gen. inverses, under certain assumptions); Cvetković-Ilić, (2008) (operators with certain restrictions); Ikramov (2009) (conditions for unique solvability); Vorontsov-Ikramov (2011) (algorithm).

- Wimmer, Lin. Alg. Appl., (1994): Necessary and sufficient conditions for **consistency**. Complex matrices
- Byers & Kressner, SIAM J. Matrix Anal. Appl., (2006): Necessary and sufficient conditions for **unique solution** for $\star = T$. Complex matrices.
- Kressner & Schröder & Watkins, Numer. Algor., (2009): Same for $\star = *$.
- De Terán & D., Elec. J. Lin. Alg., (2011): Efficient **algorithm** for computing the solution when it is unique.
- Chiang & Chu & Lin, Appl. Math. Comp., (2012): Repeat some of the results above in other language and considers more general eqs. $AXB + CX^\star D = E$.
- De Terán & D & Guillery & Montealegre & Reyes, Lin. Alg. Appl., (2013): General solution in the homogeneous case.
- Other papers Piao-Zhang-Wang, (2007) (involved, formula for solution gen. inverses, under certain assumptions); Cvetković-Ilić, (2008) (operators with certain restrictions); Ikramov (2009) (conditions for unique solvability); Vorontsov-Ikramov (2011) (algorithm).

Outline

Due to roundoff errors, uncertainty in the data, ... , usually it **is not possible** to compute the **exact canonical forms of matrix eigenvalue problems**.

- $Ax = \lambda x$ (Jordan Canonical Form (JCF)).
- $Ax = \lambda Bx$ (Kronecker Canonical Form (KCF)).

Some related questions:

- Which are the **nearby** canonical structures (JCF, KCF) to a given one?
- Which is the **generic** canonical structure?

Same questions for matrices/matrix pencils in a particular subset (low-rank, palindromic, symmetric,...) and **structure preserving numerical methods**.

▶ The **theory of orbits** provides a theoretical framework for these questions.

Some interesting questions

Due to roundoff errors, uncertainty in the data, ... , usually it **is not possible** to compute the **exact canonical forms of matrix eigenvalue problems**.

- $Ax = \lambda x$ (Jordan Canonical Form (JCF)).
- $Ax = \lambda Bx$ (Kronecker Canonical Form (KCF)).

Some related questions:

- Which are the **nearby** canonical structures (JCF, KCF) to a given one?
- Which is the **generic** canonical structure?

Same questions for matrices/matrix pencils in a particular subset (low-rank, palindromic, symmetric,...) and **structure preserving numerical methods**.

▶ The **theory of orbits** provides a theoretical framework for these questions.

Due to roundoff errors, uncertainty in the data, ... , usually it **is not possible** to compute the **exact canonical forms of matrix eigenvalue problems**.

- $Ax = \lambda x$ (Jordan Canonical Form (JCF)).
- $Ax = \lambda Bx$ (Kronecker Canonical Form (KCF)).

Some related questions:

- Which are the **nearby** canonical structures (JCF, KCF) to a given one?
- Which is the **generic** canonical structure?

Same questions for matrices/matrix pencils in a particular subset (low-rank, palindromic, symmetric,...) and **structure preserving numerical methods**.

▶ The **theory of orbits** provides a theoretical framework for these questions.

Some interesting questions

Due to roundoff errors, uncertainty in the data, ... , usually it **is not possible** to compute the **exact canonical forms of matrix eigenvalue problems**.

- $Ax = \lambda x$ (Jordan Canonical Form (JCF)).
- $Ax = \lambda Bx$ (Kronecker Canonical Form (KCF)).

Some related questions:

- Which are the **nearby** canonical structures (JCF, KCF) to a given one?
- Which is the **generic** canonical structure?

Same questions for matrices/matrix pencils in a particular subset (low-rank, palindromic, symmetric,...) and **structure preserving numerical methods**.

▶ The **theory of orbits** provides a theoretical framework for these questions.

Given $A,\ B \in \mathbb{C}^{n\times n}$

$$\begin{aligned} \mathcal{O}(A) &= \left\{PAP^T \ : \ P \text{ nonsingular}\right\} && \text{Congruence orbit of } A \\ \mathcal{O}_s(A) &= \left\{PAP^{-1} \ : \ P \text{ nonsingular}\right\} && \text{Similarity orbit of } A \\ \mathcal{O}_e(A-\lambda B) &= \{P(A-\lambda B)Q \ : \ P,Q \text{ nonsing.}\} && \text{Equivalency orbit of } A-\lambda B \end{aligned}$$

Similarity/equivalency orbits

- have been widely studied: Arnold (1971), Demmel-Edelman (1995), Edelman-Elmroth-Kågström (1997, 1999), Johansson (2006), ...
- correspond to matrices with the same Jordan Canonical Form (JCF) / Pencils with the same Kronecker Canonical Form (KCF).
- The **dimension** of these orbits gives us an idea of their “size”.
- The description of the hierarchy of inclusions between closures of different orbits allows us to identify nearby Jordan/Kronecker structures and is useful in the design of algorithms to compute the JCF/KCF.

Congruence orbits? Important in structure preserving methods for palindromic eigenproblems.

Given $A,\ B \in \mathbb{C}^{n\times n}$

$$\begin{array}{rl}\mathcal{O}(A) = \left\{PAP^T \ : \ P \text{ nonsingular}\right\} & \text{Congruence orbit of } A \\ \mathcal{O}_s(A) = \left\{PAP^{-1} \ : \ P \text{ nonsingular}\right\} & \text{Similarity orbit of } A \\ \mathcal{O}_e(A-\lambda B) = \{P(A-\lambda B)Q \ : \ P,Q \text{ nonsing.}\} & \text{Equivalency orbit of } A-\lambda B\end{array}$$

Similarity/equivalency orbits

- have been widely studied: Arnold (1971), Demmel-Edelman (1995), Edelman-Elmroth-Kågström (1997, 1999), Johansson (2006), ...
- correspond to matrices with the same Jordan Canonical Form (JCF) / Pencils with the same Kronecker Canonical Form (KCF).
- The **dimension** of these orbits gives us an idea of their "size".
- The description of the hierarchy of inclusions between closures of different orbits allows us to identify nearby Jordan/Kronecker structures and is useful in the design of algorithms to compute the JCF/KCF.

Congruence orbits? Important in structure preserving methods for palindromic eigenproblems.

Given $A, B \in \mathbb{C}^{n\times n}$

$$\mathcal{O}(A) = \{PAP^T \ : \ P \text{ nonsingular}\}$$ Congruence orbit of A

$$\mathcal{O}_s(A) = \{PAP^{-1} \ : \ P \text{ nonsingular}\}$$ Similarity orbit of A

$$\mathcal{O}_e(A-\lambda B) = \{P(A-\lambda B)Q \ : \ P, Q \text{ nonsing.}\}$$ Equivalency orbit of $A - \lambda B$

Similarity/equivalency orbits

- have been widely studied: Arnold (1971), Demmel-Edelman (1995), Edelman-Elmroth-Kågström (1997, 1999), Johansson (2006), ...
- correspond to matrices with the same Jordan Canonical Form (JCF) / Pencils with the same Kronecker Canonical Form (KCF).
- The **dimension** of these orbits gives us an idea of their “size".
- The description of the hierarchy of inclusions between closures of different orbits allows us to identify nearby Jordan/Kronecker structures and is useful in the design of algorithms to compute the JCF/KCF.

Congruence orbits? Important in structure preserving methods for palindromic eigenproblems.

Given $A, B \in \mathbb{C}^{n\times n}$

$$\begin{array}{rll} \mathcal{O}(A) = \left\{PAP^T \ : \ P \text{ nonsingular}\right\} & & \text{Congruence orbit of } A \\ \mathcal{O}_s(A) = \left\{PAP^{-1} \ : \ P \text{ nonsingular}\right\} & & \text{Similarity orbit of } A \\ \mathcal{O}_e(A-\lambda B) = \{P(A-\lambda B)Q \ : \ P, Q \text{ nonsing.}\} & & \text{Equivalency orbit of } A-\lambda B \end{array}$$

Similarity/equivalency orbits

- have been widely studied: Arnold (1971), Demmel-Edelman (1995), Edelman-Elmroth-Kågström (1997, 1999), Johansson (2006), ...
- correspond to matrices with the same Jordan Canonical Form (JCF) / Pencils with the same Kronecker Canonical Form (KCF).
- The **dimension** of these orbits gives us an idea of their “size".
- The description of the hierarchy of inclusions between closures of different orbits allows us to identify nearby Jordan/Kronecker structures and is useful in the design of algorithms to compute the JCF/KCF.

Congruence orbits? Important in structure preserving methods for palindromic eigenproblems.

Given $A,\ B \in \mathbb{C}^{n\times n}$

$$\mathcal{O}(A) = \{PAP^T \ : \ P \text{ nonsingular}\} \qquad \text{Congruence orbit of } A$$
$$\mathcal{O}_s(A) = \{PAP^{-1} \ : \ P \text{ nonsingular}\} \qquad \text{Similarity orbit of } A$$
$$\mathcal{O}_e(A-\lambda B) = \{P(A-\lambda B)Q \ : \ P, Q \text{ nonsing.}\} \qquad \text{Equivalency orbit of } A-\lambda B$$

Similarity/equivalency orbits

- have been widely studied: Arnold (1971), Demmel-Edelman (1995), Edelman-Elmroth-Kågström (1997, 1999), Johansson (2006), ...
- correspond to matrices with the same Jordan Canonical Form (JCF) / Pencils with the same Kronecker Canonical Form (KCF).
- The **dimension** of these orbits gives us an idea of their “size".
- The description of the hierarchy of inclusions between closures of different orbits allows us to identify nearby Jordan/Kronecker structures and is useful in the design of algorithms to compute the JCF/KCF.

Congruence orbits? Important in structure preserving methods for palindromic eigenproblems.

Given $A, B \in \mathbb{C}^{n\times n}$

$$\begin{array}{rl} \mathcal{O}(A) = \left\{PAP^T \ : \ P \text{ nonsingular}\right\} & \text{Congruence orbit of } A \\ \mathcal{O}_s(A) = \left\{PAP^{-1} \ : \ P \text{ nonsingular}\right\} & \text{Similarity orbit of } A \\ \mathcal{O}_e(A-\lambda B) = \{P(A-\lambda B)Q \ : \ P, Q \text{ nonsing.}\} & \text{Equivalency orbit of } A-\lambda B \end{array}$$

Similarity/equivalency orbits

- have been widely studied: Arnold (1971), Demmel-Edelman (1995), Edelman-Elmroth-Kågström (1997, 1999), Johansson (2006), ...
- correspond to matrices with the same Jordan Canonical Form (JCF) / Pencils with the same Kronecker Canonical Form (KCF).
- The **dimension** of these orbits gives us an idea of their “size".
- The description of the hierarchy of inclusions between closures of different orbits allows us to identify nearby Jordan/Kronecker structures and is useful in the design of algorithms to compute the JCF/KCF.

Congruence orbits? Important in structure preserving methods for palindromic eigenproblems.

Given $A,\ B \in \mathbb{C}^{n\times n}$

$$\begin{aligned} \mathcal{O}(A) &= \left\{PAP^T \ :\ P \text{ nonsingular}\right\} && \text{Congruence orbit of } A \\ \mathcal{O}_s(A) &= \left\{PAP^{-1} \ :\ P \text{ nonsingular}\right\} && \text{Similarity orbit of } A \\ \mathcal{O}_e(A-\lambda B) &= \{P(A-\lambda B)Q \ :\ P, Q \text{ nonsing.}\} && \text{Equivalency orbit of } A-\lambda B \end{aligned}$$

Similarity/equivalency orbits

- have been widely studied: Arnold (1971), Demmel-Edelman (1995), Edelman-Elmroth-Kågström (1997, 1999), Johansson (2006), ...
- correspond to matrices with the same Jordan Canonical Form (JCF) / Pencils with the same Kronecker Canonical Form (KCF).
- The **dimension** of these orbits gives us an idea of their “size".
- The description of the hierarchy of inclusions between closures of different orbits allows us to identify nearby Jordan/Kronecker structures and is useful in the design of algorithms to compute the JCF/KCF.

Congruence orbits? Important in structure preserving methods for palindromic eigenproblems.

$$T_{\mathcal{O}(A)}(A) = \{XA + AX^T \;:\; X \in \mathbb{C}^{n\times n}\}$$ Tangent space of $\mathcal{O}(A)$ at A

$$T_{\mathcal{O}_s(A)}(A) = \{XA - AX \;:\; X \in \mathbb{C}^{n\times n}\}$$ Tangent space of $\mathcal{O}_s(A)$ at A

Then:

(a) $\operatorname{codim} \mathcal{O}(A) = \operatorname{codim} T_{\mathcal{O}(A)}(A) = \dim($solution space of $XA + AX^T = 0)$

(b) $\operatorname{codim} \mathcal{O}_s(A) = \operatorname{codim} T_{\mathcal{O}_s(A)}(A) = \dim($solution space of $XA - AX = 0)$

General solution of $XA - AX = 0$: known since the 1950's (Gantmacher) and probably before. Depends on the JCF of A.

Our goal: Solve $XA + AX^T = 0$

(In this talk are mainly interested in the **dimension** of the solution space, but we are able also to give the solution!)

$$T_{\mathcal{O}(A)}(A) = \{XA + AX^T \ : \ X \in \mathbb{C}^{n\times n}\}$$ Tangent space of $\mathcal{O}(A)$ at A

$$T_{\mathcal{O}_s(A)}(A) = \{XA - AX \ : \ X \in \mathbb{C}^{n\times n}\}$$ Tangent space of $\mathcal{O}_s(A)$ at A

Then:

(a) $\operatorname{codim} \mathcal{O}(A) = \operatorname{codim} T_{\mathcal{O}(A)}(A) = \dim($solution space of $XA + AX^T = 0)$

(b) $\operatorname{codim} \mathcal{O}_s(A) = \operatorname{codim} T_{\mathcal{O}_s(A)}(A) = \dim($solution space of $XA - AX = 0)$

General solution of $XA - AX = 0$: known since the 1950's (Gantmacher) and probably before. Depends on the JCF of A.

Our goal: Solve $XA + AX^T = 0$

(In this talk are mainly interested in the **dimension** of the solution space, but we are able also to give the solution!)

$$T_{\mathcal{O}(A)}(A) = \left\{XA + AX^T \ : \ X \in \mathbb{C}^{n\times n}\right\} \quad \text{Tangent space of } \mathcal{O}(A) \text{ at } A$$
$$T_{\mathcal{O}_s(A)}(A) = \{XA - AX \ : \ X \in \mathbb{C}^{n\times n}\} \quad \text{Tangent space of } \mathcal{O}_s(A) \text{ at } A$$

Then:

(a) $\operatorname{codim} \mathcal{O}(A) = \operatorname{codim} T_{\mathcal{O}(A)}(A) = \dim($solution space of $XA + AX^T = 0)$

(b) $\operatorname{codim} \mathcal{O}_s(A) = \operatorname{codim} T_{\mathcal{O}_s(A)}(A) = \dim($solution space of $XA - AX = 0)$

General solution of $XA - AX = 0$: known since the 1950's (Gantmacher) and probably before. Depends on the JCF of A.

Our goal: Solve $XA + AX^T = 0$

(In this talk are mainly interested in the **dimension** of the solution space, but we are able also to give the solution!)

$$T_{\mathcal{O}(A)}(A) = \left\{XA + AX^T \ : \ X \in \mathbb{C}^{n\times n}\right\}$$ Tangent space of $\mathcal{O}(A)$ at A

$$T_{\mathcal{O}_s(A)}(A) = \{XA - AX \ : \ X \in \mathbb{C}^{n\times n}\}$$ Tangent space of $\mathcal{O}_s(A)$ at A

Then:

(a) $\operatorname{codim} \mathcal{O}(A) = \operatorname{codim} T_{\mathcal{O}(A)}(A) = \dim($solution space of $XA + AX^T = 0)$

(b) $\operatorname{codim} \mathcal{O}_s(A) = \operatorname{codim} T_{\mathcal{O}_s(A)}(A) = \dim($solution space of $XA - AX = 0)$

General solution of $XA - AX = 0$: known since the 1950's (Gantmacher) and probably before. Depends on the JCF of A.

Our goal: Solve $XA + AX^T = 0$

(In this talk are mainly interested in the **dimension** of the solution space, but we are able also to give the solution!)

$$T_{\mathcal{O}(A)}(A) = \{XA + AX^T \ : \ X \in \mathbb{C}^{n\times n}\}$$ Tangent space of $\mathcal{O}(A)$ at A

$$T_{\mathcal{O}_s(A)}(A) = \{XA - AX \ : \ X \in \mathbb{C}^{n\times n}\}$$ Tangent space of $\mathcal{O}_s(A)$ at A

Then:

(a) $\operatorname{codim} \mathcal{O}(A) = \operatorname{codim} T_{\mathcal{O}(A)}(A) = \dim($solution space of $XA + AX^T = 0)$

(b) $\operatorname{codim} \mathcal{O}_s(A) = \operatorname{codim} T_{\mathcal{O}_s(A)}(A) = \dim($solution space of $XA - AX = 0)$

General solution of $XA - AX = 0$: known since the 1950's (Gantmacher) and probably before. Depends on the JCF of A.

Our goal: Solve $XA + AX^T = 0$

(In this talk are mainly interested in the **dimension** of the solution space, but we are able also to give the solution!)

Outline

1. Previous and related work
2. The equation $AX^T + XA = 0$
   - Motivation: Orbits and the computation of canonical forms
   - Strategy for solving $AX^T + XA = 0$
   - The canonical form for congruence
   - The solution of $AX^T + XA = 0$
   - Generic canonical structure for congruence
3. The general equation $AX + X^\star B = C$
   - Motivation
   - Consistency of the Sylvester equation for $\star$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions
4. General solution of $AX + X^\star B = 0$
5. Conclusions

Notation: $\mathcal{S}_A = \left\{ X \in \mathbb{C}^{n\times n} : \; AX^T + XA = 0 \right\}$ (solution space)

Consider $B := PAP^T$ (P nonsingular) then

$$B\left(PXP^{-1}\right)^T + \left(PXP^{-1}\right)B = 0$$

and $\mathcal{S}_A = P^{-1}\mathcal{S}_B P$.

In particular: $\boxed{\dim \mathcal{S}_A = \dim \mathcal{S}_B}$

Procedure to solve $AX^T + XA = 0$:

1. Set $C_A = PAP^T$, the canonical form of A for congruence !?.
2. Solve $C_A Y^T + Y C_A = 0$.
3. Undo the change: $X = P^{-1} Y P$.

Notation: $\mathcal{S}_A = \left\{ X \in \mathbb{C}^{n\times n} : \; AX^T + XA = 0 \right\}$ (solution space)

Consider $B := PAP^T$ (P nonsingular) then

$$B \left(PXP^{-1}\right)^T + \left(PXP^{-1}\right) B = 0$$

and $\mathcal{S}_A = P^{-1}\mathcal{S}_B P$.

In particular: $\boxed{\dim \mathcal{S}_A = \dim \mathcal{S}_B}$

Procedure to solve $AX^T + XA = 0$:

1. Set $C_A = PAP^T$, the canonical form of A for congruence !?.
2. Solve $C_A Y^T + Y C_A = 0$.
3. Undo the change: $X = P^{-1} Y P$.

Getting a simpler equation via congruence

Notation: $\mathcal{S}_A = \left\{ X \in \mathbb{C}^{n\times n} : \; AX^T + XA = 0 \right\}$ (solution space)

Consider $B := PAP^T$ (P nonsingular) then

$$B\left(PXP^{-1}\right)^T + \left(PXP^{-1}\right)B = 0$$

and $\mathcal{S}_A = P^{-1}\mathcal{S}_B P$.

In particular: $\boxed{\dim \mathcal{S}_A = \dim \mathcal{S}_B}$

Procedure to solve $AX^T + XA = 0$:

1. Set $C_A = PAP^T$, the canonical form of A for congruence !?.
2. Solve $C_A Y^T + Y C_A = 0$.
3. Undo the change: $X = P^{-1} Y P$.

Getting a simpler equation via congruence

Notation: $\mathcal{S}_A = \left\{ X \in \mathbb{C}^{n\times n} : \ AX^T + XA = 0 \right\}$ (solution space)

Consider $B := PAP^T$ (P nonsingular) then

$$B \left(PXP^{-1}\right)^T + \left(PXP^{-1}\right) B = 0$$

and $\mathcal{S}_A = P^{-1} \mathcal{S}_B P$.

In particular: $\boxed{\dim \mathcal{S}_A = \dim \mathcal{S}_B}$

Procedure to solve $AX^T + XA = 0$:

1. Set $C_A = PAP^T$, the canonical form of A for congruence !?.
2. Solve $C_A Y^T + Y C_A = 0$.
3. Undo the change: $X = P^{-1} Y P$.

Notation: $\mathcal{S}_A = \left\{X \in \mathbb{C}^{n\times n} : \ AX^T + XA = 0\right\}$ (solution space)

Consider $B := PAP^T$ (P nonsingular) then

$$B\left(PXP^{-1}\right)^T + \left(PXP^{-1}\right)B = 0$$

and $\mathcal{S}_A = P^{-1}\mathcal{S}_B P$.

In particular: $\boxed{\dim \mathcal{S}_A = \dim \mathcal{S}_B}$

Procedure to solve $AX^T + XA = 0$:

1. Set $C_A = PAP^T$, the canonical form of A for congruence !?.
2. Solve $C_A Y^T + YC_A = 0$.
3. Undo the change: $X = P^{-1}YP$.

Outline

1. Previous and related work
2. The equation $AX^T + XA = 0$
   - Motivation: Orbits and the computation of canonical forms
   - Strategy for solving $AX^T + XA = 0$
   - The canonical form for congruence
   - The solution of $AX^T + XA = 0$
   - Generic canonical structure for congruence
3. The general equation $AX + X^{\star}B = C$
   - Motivation
   - Consistency of the Sylvester equation for $\star$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions
4. General solution of $AX + X^{\star}B = 0$
5. Conclusions

Theorem (Canonical form for congruence (Horn & Sergeichuk, 2006))

Each matrix $A \in \mathbb{C}^{n\times n}$ is **congruent** *to a direct sum, uniquely determined up to permutation of summands, of blocks of types 0, I and II.*

$$(Type\ 0) \qquad J_k(0) = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}_{k\times k}$$

$$(Type\ I) \qquad \Gamma_k = \begin{bmatrix} 0 & & & & & (-1)^{k+1} \\ & & & & \iddots & (-1)^k \\ & & & -1 & \iddots & \\ & & 1 & 1 & & \\ & -1 & -1 & & & \\ 1 & 1 & & & & 0 \end{bmatrix}_{k\times k}, \qquad \Gamma_1 = [1]$$

$$(Type\ II) \quad H_{2k}(\mu) = \begin{bmatrix} 0 & I_k \\ J_k(\mu) & 0 \end{bmatrix}_{2k\times 2k}, \quad H_2(\mu) = \begin{bmatrix} 0 & 1 \\ \mu & 0 \end{bmatrix}, \quad (0 \neq \mu \neq (-1)^{k+1})$$

The canonical form for congruence: a brief history (I)

- Turnbull (U. St. Andrews, Scotland) & Aitken (U. Edinburgh, Scotland), *An Introduction to the Theory of Canonical Matrices*, 1932.

  For complex matrices. Six types of blocks.

- Gabriel, J. Algebra (1974), studied equivalence of bilinear forms in fields with characteristic $\neq 2$.
- Riehm, J. Algebra (1974), reduced the problem of equivalence of bilinear forms to equivalence of Hermitian forms.
- Sergeichuk, Math. USSR Izvestiya (1988) complete study via quivers and Hermitian forms in fields with characteristic $\neq 2$.
- Thompson, Linear Algebra and its Applications (1991). Complex and real matrices: Symmetric/Skew-Symmetric pencils.
- Lee and Weinberg, Linear Algebra and its Applications (1996). Complex and real matrices based on Thompson and $A = S + K$, with $S = S^T$ and $K = -K^T$. Six blocks for complex (Turnbull and Aitken). Eight blocks for real matrices.

The canonical form for congruence: a brief history (I)

- Turnbull (U. St. Andrews, Scotland) & Aitken (U. Edinburgh, Scotland), *An Introduction to the Theory of Canonical Matrices*, 1932.

  For complex matrices. Six types of blocks.

- Gabriel, J. Algebra (1974), studied equivalence of bilinear forms in fields with characteristic $\neq 2$.

- Riehm, J. Algebra (1974), reduced the problem of equivalence of bilinear forms to equivalence of Hermitian forms.

- Sergeichuk, Math. USSR Izvestiya (1988) complete study via quivers and Hermitian forms in fields with characteristic $\neq 2$.

- Thompson, Linear Algebra and its Applications (1991). Complex and real matrices: Symmetric/Skew-Symmetric pencils.

- Lee and Weinberg, Linear Algebra and its Applications (1996). Complex and real matrices based on Thompson and $A = S + K$, with $S = S^T$ and $K = -K^T$. Six blocks for complex (Turnbull and Aitken). Eight blocks for real matrices.

The canonical form for congruence: a brief history (I)

- Turnbull (U. St. Andrews, Scotland) & Aitken (U. Edinburgh, Scotland), *An Introduction to the Theory of Canonical Matrices*, 1932.

  For complex matrices. Six types of blocks.
- Gabriel, J. Algebra (1974), studied equivalence of bilinear forms in fields with characteristic $\neq 2$.
- Riehm, J. Algebra (1974), reduced the problem of equivalence of bilinear forms to equivalence of Hermitian forms.
- Sergeichuk, Math. USSR Izvestiya (1988) complete study via quivers and Hermitian forms in fields with characteristic $\neq 2$.
- Thompson, Linear Algebra and its Applications (1991). Complex and real matrices: Symmetric/Skew-Symmetric pencils.
- Lee and Weinberg, Linear Algebra and its Applications (1996). Complex and real matrices based on Thompson and $A = S + K$, with $S = S^T$ and $K = -K^T$. Six blocks for complex (Turnbull and Aitken). Eight blocks for real matrices.

- Turnbull (U. St. Andrews, Scotland) & Aitken (U. Edinburgh, Scotland), *An Introduction to the Theory of Canonical Matrices*, 1932.

  For complex matrices. Six types of blocks.
- Gabriel, J. Algebra (1974), studied equivalence of bilinear forms in fields with characteristic $\neq 2$.
- Riehm, J. Algebra (1974), reduced the problem of equivalence of bilinear forms to equivalence of Hermitian forms.
- Sergeichuk, Math. USSR Izvestiya (1988) complete study via quivers and Hermitian forms in fields with characteristic $\neq 2$.
- Thompson, Linear Algebra and its Applications (1991). Complex and real matrices: Symmetric/Skew-Symmetric pencils.
- Lee and Weinberg, Linear Algebra and its Applications (1996). Complex and real matrices based on Thompson and $A = S + K$, with $S = S^T$ and $K = -K^T$. Six blocks for complex (Turnbull and Aitken). Eight blocks for real matrices.

- Turnbull (U. St. Andrews, Scotland) & Aitken (U. Edinburgh, Scotland), *An Introduction to the Theory of Canonical Matrices*, 1932.

  For complex matrices. Six types of blocks.

- Gabriel, J. Algebra (1974), studied equivalence of bilinear forms in fields with characteristic $\neq 2$.
- Riehm, J. Algebra (1974), reduced the problem of equivalence of bilinear forms to equivalence of Hermitian forms.
- Sergeichuk, Math. USSR Izvestiya (1988) complete study via quivers and Hermitian forms in fields with characteristic $\neq 2$.
- Thompson, Linear Algebra and its Applications (1991). Complex and real matrices: Symmetric/Skew-Symmetric pencils.
- Lee and Weinberg, Linear Algebra and its Applications (1996). Complex and real matrices based on Thompson and $A = S + K$, with $S = S^T$ and $K = -K^T$. Six blocks for complex (Turnbull and Aitken). Eight blocks for real matrices.

- Turnbull (U. St. Andrews, Scotland) & Aitken (U. Edinburgh, Scotland), *An Introduction to the Theory of Canonical Matrices*, 1932.

  For complex matrices. Six types of blocks.
- Gabriel, J. Algebra (1974), studied equivalence of bilinear forms in fields with characteristic $\neq 2$.
- Riehm, J. Algebra (1974), reduced the problem of equivalence of bilinear forms to equivalence of Hermitian forms.
- Sergeichuk, Math. USSR Izvestiya (1988) complete study via quivers and Hermitian forms in fields with characteristic $\neq 2$.
- Thompson, Linear Algebra and its Applications (1991). Complex and real matrices: Symmetric/Skew-Symmetric pencils.
- Lee and Weinberg, Linear Algebra and its Applications (1996). Complex and real matrices based on Thompson and $A = S + K$, with $S = S^T$ and $K = -K^T$. Six blocks for complex (Turnbull and Aitken). Eight blocks for real matrices.

- Corbas and Williams, J. Pure Appl. Algebra (2001), canonical forms over algebraically closed fields with characteristic not 2.
- Lanscaster and Rodman, Linear Algebra and its Applications (2005) and SIAM Review (2005). Same approach as Thompson and Lee-Weinberg.
- Simplest form for complex matrices with only 3 types of blocks: Horn and Sergeichuk, Linear Algebra and its Applications (2004, 2006). Proofs: first based on quivers and second constructive and based only on basic Matrix Analysis.

► F. De Terán, *"Canonical forms for congruence of matrices: a tribute to H. W. Turnbull and A. C. Aitken"*, Actas del II congreso de la red ALAMA, Valencia, 2-4 june, 2010.

The canonical form for congruence: a brief history (II)

- Corbas and Williams, J. Pure Appl. Algebra (2001), canonical forms over algebraically closed fields with characteristic not 2.
- Lanscaster and Rodman, Linear Algebra and its Applications (2005) and SIAM Review (2005). Same approach as Thompson and Lee-Weinberg.
- Simplest form for complex matrices with only 3 types of blocks: Horn and Sergeichuk, Linear Algebra and its Applications (2004, 2006). Proofs: first based on quivers and second constructive and based only on basic Matrix Analysis.

► F. De Terán, *"Canonical forms for congruence of matrices: a tribute to H. W. Turnbull and A. C. Aitken"*, Actas del II congreso de la red ALAMA, Valencia, 2-4 june, 2010.

The canonical form for congruence: a brief history (II)

- Corbas and Williams, J. Pure Appl. Algebra (2001), canonical forms over algebraically closed fields with characteristic not 2.
- Lanscaster and Rodman, Linear Algebra and its Applications (2005) and SIAM Review (2005). Same approach as Thompson and Lee-Weinberg.
- Simplest form for complex matrices with only 3 types of blocks: Horn and Sergeichuk, Linear Algebra and its Applications (2004, 2006). Proofs: first based on quivers and second constructive and based only on basic Matrix Analysis.

▶ F. De Terán, *"Canonical forms for congruence of matrices: a tribute to H. W. Turnbull and A. C. Aitken"*, Actas del II congreso de la red ALAMA, Valencia, 2-4 june, 2010.

- Corbas and Williams, J. Pure Appl. Algebra (2001), canonical forms over algebraically closed fields with characteristic not 2.
- Lanscaster and Rodman, Linear Algebra and its Applications (2005) and SIAM Review (2005). Same approach as Thompson and Lee-Weinberg.
- Simplest form for complex matrices with only 3 types of blocks: Horn and Sergeichuk, Linear Algebra and its Applications (2004, 2006). Proofs: first based on quivers and second constructive and based only on basic Matrix Analysis.

▶ F. De Terán, *"Canonical forms for congruence of matrices: a tribute to H. W. Turnbull and A. C. Aitken"* , Actas del II congreso de la red ALAMA, Valencia, 2-4 june, 2010.

Procedure to solve $AX^T + XA = 0$:

1. Set $C_A = PAP^T$, the canonical form of A for congruence.
2. Solve $C_A Y^T + Y C_A = 0$.
3. Undo the change: $X = P^{-1} Y P$.

Partition into blocks to solve $XC_A + C_A X^T = 0$ (I)

Set $C_A = D_1 \oplus \cdots \oplus D_s$, $D_i = J_k(0),\ \Gamma_k,$ or $H_{2k}(\mu)$ (Canonical form of A)

Partition $X = \begin{bmatrix} X_{11} & \dots & X_{1s} \\ \vdots & & \vdots \\ X_{s1} & \dots & X_{ss} \end{bmatrix}$ conformally with C_A.

Equating the (i,j) and (j,i) blocks of $XC_A + C_A X^T = 0$, we get:

- $i = j$: $X_{ii} D_i + D_i X_{ii}^T = 0$ $\rightarrow$ codim D_i (codimension)
- $i \neq j$: $\begin{array}{ll} (i,j) & X_{ij} D_j + D_i X_{ji}^T = 0 \\ (j,i) & X_{ji} D_i + D_j X_{ij}^T = 0 \end{array}$ $\rightarrow$ inter (D_i, D_j) (interaction)

Then:

$\dim \mathcal{S}_A =$ codim $\mathcal{O}(A) = \sum_i$ codim $D_i + \sum_{i \neq j}$ inter (D_i, D_j)

Partition into blocks to solve $XC_A + C_A X^T = 0$ (I)

Set $C_A = D_1 \oplus \cdots \oplus D_s$, $D_i = J_k(0),\ \Gamma_k,$ or $H_{2k}(\mu)$ (Canonical form of A)

Partition $X = \begin{bmatrix} X_{11} & \dots & X_{1s} \\ \vdots & & \vdots \\ X_{s1} & \dots & X_{ss} \end{bmatrix}$ conformally with C_A.

Equating the (i,j) and (j,i) blocks of $XC_A + C_A X^T = 0$, we get:

- $i = j$: $X_{ii} D_i + D_i X_{ii}^T = 0$ $\rightarrow$ codim D_i (codimension)
- $i \neq j$: $\begin{array}{ll} (i,j) & X_{ij} D_j + D_i X_{ji}^T = 0 \\ (j,i) & X_{ji} D_i + D_j X_{ij}^T = 0 \end{array}$ $\rightarrow$ inter (D_i, D_j) (interaction)

Then:

$\dim \mathcal{S}_A =$ codim $\mathcal{O}(A) = \sum_i$ codim $D_i + \sum_{i \neq j}$ inter (D_i, D_j)

Partition into blocks to solve $XC_A + C_A X^T = 0$ (I)

Set $C_A = D_1 \oplus \cdots \oplus D_s$, $D_i = J_k(0),\ \Gamma_k,$ or $H_{2k}(\mu)$ (Canonical form of A)

Partition $X = \begin{bmatrix} X_{11} & \dots & X_{1s} \\ \vdots & & \vdots \\ X_{s1} & \dots & X_{ss} \end{bmatrix}$ conformally with C_A.

Equating the (i,j) and (j,i) blocks of $XC_A + C_A X^T = 0$, we get:

- $i = j$: $X_{ii} D_i + D_i X_{ii}^T = 0 \qquad \rightarrow$ codim D_i (codimension)
- $i \neq j$: $\begin{array}{ll} (i,j) & X_{ij} D_j + D_i X_{ji}^T = 0 \\ (j,i) & X_{ji} D_i + D_j X_{ij}^T = 0 \end{array} \quad \rightarrow$ inter (D_i, D_j) (interaction)

Then:

$$\dim \mathcal{S}_A = \text{codim}\, \mathcal{O}(A) = \sum_i \text{codim}\, D_i + \sum_{i \neq j} \text{inter}\, (D_i, D_j)$$

The problem reduces to solve matrix equations of the types:

(a) $XD + DX^T = 0$ (easier Sylvester equation for congruence)

with $D = J_k(0)$ (type 0), Γ_k (type I), or $H_{2k}(\mu)$ (type II)
(3 different types of eqs.)

(b) $$\begin{aligned} XD_1 + D_2 Y^T &= 0 \\ YD_2 + D_1 X^T &= 0 \end{aligned}$$ (system of two matrix equations)

with $D_1, D_2 = J_k(0)$ (type 0), Γ_ℓ (type I), or $H_{2m}(\mu)$ (type II)
(6 different types of eqs.)

The codimension formula

Theorem (De Terán & D, Lin. Alg. Appl., 2011)

Let $A \in \mathbb{C}^{n\times n}$ be a matrix with canonical form for congruence

$$\begin{aligned} C_A = & J_{p_1}(0) \oplus J_{p_2}(0) \oplus \cdots \oplus J_{p_a}(0) \\ & \oplus \Gamma_{q_1} \oplus \Gamma_{q_2} \oplus \cdots \oplus \Gamma_{q_b} \\ & \oplus H_{2r_1}(\mu_1) \oplus H_{2r_2}(\mu_2) \oplus \cdots \oplus H_{2r_c}(\mu_c). \end{aligned}$$

Then the codimension of the orbit of A for the action of congruence, i.e., the dimension of the solution space of $XA + AX^T = 0$, depends only on C_A. It can be computed as the sum

$$c_{\text{Total}} = c_0 + c_1 + c_2 + i_{00} + i_{11} + i_{22} + i_{01} + i_{02} + i_{12}\,.$$

Codimension		
$c_0 \to \left\lceil \frac{k}{2} \right\rceil$	$c_1 \to \left\lfloor \frac{k}{2} \right\rfloor$	$c_2 \to \begin{cases} k\,, & \text{if } \mu \neq (-1)^k \\ k + 2\left\lceil \frac{k}{2} \right\rceil\,, & \text{if } \mu = (-1)^k \end{cases}$

Interaction (same type)	Interaction (different type)
$i_{00} \to \begin{cases} \ell\,, & \ell \text{ even} \\ k\,, & \ell \text{ odd and } k \neq \ell \\ k+1\,, & \ell \text{ odd and } k = \ell \end{cases}$	$i_{01} \to \begin{cases} 0\,, & k \text{ even} \\ \ell\,, & k \text{ odd} \end{cases}$
$i_{11} \to \begin{cases} 0\,, & k, \ell \text{ different parity} \\ \min\{k,\ell\}\,, & k, \ell \text{ same parity} \end{cases}$	$i_{02} \to \begin{cases} 0\,, & k \text{ even} \\ 2\ell\,, & k \text{ odd} \end{cases}$
$i_{22} \to \begin{cases} 4\min\{k,\ell\}\,, & \mu = \widetilde{\mu} = \pm 1 \\ 2\min\{k,\ell\}\,, & \mu = \widetilde{\mu} \neq \pm 1 \\ 2\min\{k,\ell\}\,, & \mu \neq \widetilde{\mu},\ \mu\widetilde{\mu} = 1 \\ 0\,, & \mu \neq \widetilde{\mu},\ \mu\widetilde{\mu} \neq 1 \end{cases}$	$i_{12} \to \begin{cases} 2\min\{k,\ell\}\,, & \mu = (-1)^{k+1} \\ 0\,, & \mu \neq (-1)^{k+1} \end{cases}$

► Explicit solution found by De Terán & D (LAA, 2011) in all cases, except for the case eq. corresp. to codim. of two special type II blocks:

$$X H_{2k}((-1)^k) + H_{2k}((-1)^k) X^T = 0\,.$$

This solved by S. R. García & A. L. Shoemaker, Lin. Alg. Appl., 2012.

Codimensions and interactions of canonical blocks

Codimension		
$c_0 \to \lceil \frac{k}{2} \rceil$	$c_1 \to \lfloor \frac{k}{2} \rfloor$	$c_2 \to \begin{cases} k\,, & \text{if } \mu \neq (-1)^k \\ k + 2\lceil \frac{k}{2} \rceil\,, & \text{if } \mu = (-1)^k \end{cases}$

Interaction (same type)	Interaction (different type)
$i_{00} \to \begin{cases} \ell\,, & \ell \text{ even} \\ k\,, & \ell \text{ odd and } k \neq \ell \\ k+1\,, & \ell \text{ odd and } k = \ell \end{cases}$	$i_{01} \to \begin{cases} 0\,, & k \text{ even} \\ \ell\,, & k \text{ odd} \end{cases}$
$i_{11} \to \begin{cases} 0\,, & k,\ell \text{ different parity} \\ \min\{k,\ell\}\,, & k,\ell \text{ same parity} \end{cases}$	$i_{02} \to \begin{cases} 0\,, & k \text{ even} \\ 2\ell\,, & k \text{ odd} \end{cases}$
$i_{22} \to \begin{cases} 4\min\{k,\ell\}\,, & \mu = \widetilde{\mu} = \pm 1 \\ 2\min\{k,\ell\}\,, & \mu = \widetilde{\mu} \neq \pm 1 \\ 2\min\{k,\ell\}\,, & \mu \neq \widetilde{\mu},\ \mu\widetilde{\mu} = 1 \\ 0\,, & \mu \neq \widetilde{\mu},\ \mu\widetilde{\mu} \neq 1 \end{cases}$	$i_{12} \to \begin{cases} 2\min\{k,\ell\}\,, & \mu = (-1)^{k+1} \\ 0\,, & \mu \neq (-1)^{k+1} \end{cases}$

► Explicit solution found by De Terán & D (LAA, 2011) in all cases, except for the case eq. corresp. to codim. of two special type II blocks:

$$X H_{2k}((-1)^k) + H_{2k}((-1)^k) X^T = 0\,.$$

This solved by S. R. García & A. L. Shoemaker, Lin. Alg. Appl., 2012.

Generic = codimension zero

Theorem (De Terán & D, Lin. Alg. Appl., 2011)

The minimal codimension for a congruence orbit in $\mathbb{C}^{n\times n}$ *is* $\lfloor n/2 \rfloor$.

Generic canonical **structure** for congruence is not given by a single orbit!!

Similarity orbits (JCF): There is no generic JCF with fixed eigenvalues.

- The **generic Jordan structure** is $J_1(\lambda_1) \oplus \cdots \oplus J_1(\lambda_n)$, with $\lambda_1, \ldots, \lambda_n$ different (not fixed)

Generic = codimension zero

Theorem (De Terán & D, Lin. Alg. Appl., 2011)

The minimal codimension for a congruence orbit in $\mathbb{C}^{n\times n}$ *is* $\lfloor n/2 \rfloor$.

Generic canonical **structure** for congruence is not given by a single orbit!!

Similarity orbits (JCF): There is no generic JCF with fixed eigenvalues.

- The **generic Jordan structure** is $J_1(\lambda_1) \oplus \cdots \oplus J_1(\lambda_n)$, with $\lambda_1, \ldots, \lambda_n$ different (not fixed)

Definition (Arnold, 1971)

Given $A \in \mathbb{C}^{n\times n}$ with Jordan Canonical Form

$$J_A = J_{\lambda_1} \oplus \cdots \oplus J_{\lambda_d} \,,$$

where

$$J_{\lambda_i} := J_{n_{i,1}}(\lambda_i) \oplus \cdots \oplus J_{n_{i,q_i}}(\lambda_i), \qquad \text{for } i = 1,\ldots,d \text{ and } \lambda_i \neq \lambda_j \text{ if } i \neq j,$$

the **similarity bundle** of A is

$$\mathcal{B}_s(A) = \bigcup_{\substack{\lambda_i' \in \mathbb{C},\ i=1,\ldots,d \\ \lambda_i' \neq \lambda_j',\ i\neq j}} \mathcal{O}_s\left(J_{\lambda_1'} \oplus \cdots \oplus J_{\lambda_d'}\right)$$

Generic canonical structure for congruence (III): Congruence bundles

Definition (De Terán & D, Lin. Alg. Appl., 2011)

Given $A \in \mathbb{C}^{n\times n}$ with canonical form for congruence

$$C_A = \bigoplus_{i=1}^{a} J_{p_i}(0) \;\oplus\; \bigoplus_{i=1}^{b} \Gamma_{q_i} \;\oplus\; \bigoplus_{i=1}^{t} \mathcal{H}(\mu_i), \quad \mu_i \neq \mu_j,\, \mu_i \neq 1/\mu_j \text{ if } i \neq j,$$

where

$$\mathcal{H}(\mu_i) = H_{2r_{i,1}}(\mu_i) \oplus H_{2r_{i,2}}(\mu_i) \oplus \cdots \oplus H_{2r_{i,g_i}}(\mu_i), \qquad \text{for } i = 1, \ldots, t,$$

the **congruence bundle** of A is

$$\mathcal{B}(A) = \bigcup_{\substack{\mu_i' \in \mathbb{C},\ i=1,\ldots,t \\ \mu_i' \neq \mu_j',\ \mu_i'\mu_j' \neq 1, i\neq j}} \mathcal{O}\left(\bigoplus_{i=1}^{a} J_{p_i}(0) \;\oplus\; \bigoplus_{i=1}^{b} \Gamma_{q_i} \;\oplus\; \bigoplus_{i=1}^{t} \mathcal{H}(\mu_i') \right).$$

(same structure as C_A but unfixed complex values μ in type II blocks)

If t=number of different $\mu' s$ appearing in type II blocks of C_A, then $\mathrm{codim}(\mathcal{B}(A)) = \mathrm{codim}(\mathcal{O}(A)) - t\,.$

Theorem (De Terán & D, Lin. Alg. Appl., 2011)

The following bundles for congruence in $\mathbb{C}^{n\times n}$ have codimension zero

1. n *even*

$$G_n = \mathcal{B}\left(H_2(\mu_1) \oplus H_2(\mu_2) \oplus \cdots \oplus H_2(\mu_{n/2})\right),$$

with $\mu_i \neq \pm 1$, $i = 1, \ldots, n/2$, $\mu_i \neq \mu_j$ *and* $\mu_i \neq 1/\mu_j$ *if* $i \neq j$.

2. n *odd*

$$G_n = \mathcal{B}\left(H_2(\mu_1) \oplus H_2(\mu_2) \oplus \cdots \oplus H_2(\mu_{(n-1)/2}) \oplus \Gamma_1\right),$$

with $\mu_i \neq \pm 1$, $i = 1, \ldots, (n-1)/2$, $\mu_i \neq \mu_j$ *and* $\mu_i \neq 1/\mu_j$ *if* $i \neq j$.

Then G_n is the **generic canonical structure for congruence** *in $\mathbb{C}^{n\times n}$ (with unspecified values $\mu_1, \mu_2, \ldots$).*

Summary of section 3

Given $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times m}$, and $C \in \mathbb{C}^{m\times m}$, we study the equations

$$\boxed{A\,X + X^\star B = C\,,} \qquad (X^\star = X^T \text{ or } X^*),$$

where $X \in \mathbb{C}^{n\times m}$ is the unknown to be determined. More precisely:

1. Necessary and sufficient conditions for consistency (Wimmer 1994, De Terán & D., Elect. J. Lin. Alg., 2011 (2)).
2. Necessary and sufficient conditions for uniqueness of solutions (Byers, Kressner, Schröder, Watkins, 2006, 2009).
3. Efficient and stable numerical algorithm for computing the unique solution (De Terán & D., Elect. J. Lin. Alg., 2011 (2)).

We establish parallelisms/differences with **well-known Sylvester equation**

$$\boxed{A\,X - X\,B = C\,,} \qquad A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}, C \in \mathbb{C}^{m\times n}.$$

Summary of section 3

Given $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times m}$, and $C \in \mathbb{C}^{m\times m}$, we study the equations

$$\boxed{A\,X + X^\star\,B = C\,,} \qquad (X^\star = X^T \text{ or } X^*),$$

where $X \in \mathbb{C}^{n\times m}$ is the unknown to be determined. More precisely:

1. Necessary and sufficient conditions for consistency (Wimmer 1994, De Terán & D., Elect. J. Lin. Alg., 2011 (2)).
2. Necessary and sufficient conditions for uniqueness of solutions (Byers, Kressner, Schröder, Watkins, 2006, 2009).
3. Efficient and stable numerical algorithm for computing the unique solution (De Terán & D., Elect. J. Lin. Alg., 2011 (2)).

We establish parallelisms/differences with **well-known Sylvester equation**

$$\boxed{A\,X - X\,B = C\,,} \qquad A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}, C \in \mathbb{C}^{m\times n}.$$

Summary of section 3

Given $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times m}$, and $C \in \mathbb{C}^{m\times m}$, we study the equations

$$\boxed{A\,X + X^\star B = C\,,} \qquad (X^\star = X^T \text{ or } X^*),$$

where $X \in \mathbb{C}^{n\times m}$ is the unknown to be determined. More precisely:

1. Necessary and sufficient conditions for consistency (Wimmer 1994, De Terán & D., Elect. J. Lin. Alg., 2011 (2)).
2. Necessary and sufficient conditions for uniqueness of solutions (Byers, Kressner, Schröder, Watkins, 2006, 2009).
3. Efficient and stable numerical algorithm for computing the unique solution (De Terán & D., Elect. J. Lin. Alg., 2011 (2)).

We establish parallelisms/differences with **well-known Sylvester equation**

$$\boxed{A\,X - X\,B = C\,,} \qquad A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}, C \in \mathbb{C}^{m\times n}\,.$$

KEY: Canonical forms for congruence DO not work for $AX + X^T B = C$

- $AX + X^T B = C$, with $A \neq B$.
- $A = Q_A C_A Q_A^T$ and $B = Q_B C_B Q_B^T$.
- $$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$
- $$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$
- $$C_A \boxed{Q_A^T X Q_B^{-T}} + \boxed{Q_A^{-1} X^T Q_B} C_B = Q_A^{-1} C Q_B^{-T}$$
- But,
$$\boxed{\left(Q_A^T X Q_B^{-T}\right)^T} \neq \boxed{Q_A^{-1} X^T Q_B}$$
- with equality only if
$$Q_A = Q_B$$

KEY: Canonical forms for congruence DO not work for $AX + X^T B = C$

- $AX + X^T B = C$, with $A \neq B$.
- $A = Q_A C_A Q_A^T$ and $B = Q_B C_B Q_B^T$.
- $$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$
- $$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$
- $$C_A \boxed{Q_A^T X Q_B^{-T}} + \boxed{Q_A^{-1} X^T Q_B} C_B = Q_A^{-1} C Q_B^{-T}$$
- But,
$$\boxed{\left(Q_A^T X Q_B^{-T}\right)^T} \neq \boxed{Q_A^{-1} X^T Q_B}$$
- with equality only if
$$Q_A = Q_B$$

KEY: Canonical forms for congruence DO not work for $AX + X^T B = C$

- $AX + X^T B = C$, with $A \neq B$.
- $A = Q_A C_A Q_A^T$ and $B = Q_B C_B Q_B^T$.
- $$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$
- $$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$
- $$C_A \boxed{Q_A^T X Q_B^{-T}} + \boxed{Q_A^{-1} X^T Q_B} C_B = Q_A^{-1} C Q_B^{-T}$$
- But,
$$\boxed{\left(Q_A^T X Q_B^{-T}\right)^T} \neq \boxed{Q_A^{-1} X^T Q_B}$$
- with equality only if
$$Q_A = Q_B$$

KEY: Canonical forms for congruence DO not work for $AX + X^T B = C$

- $AX + X^T B = C$, with $A \neq B$.
- $A = Q_A C_A Q_A^T$ and $B = Q_B C_B Q_B^T$.
- $$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$
- $$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$
- $$C_A \boxed{Q_A^T X Q_B^{-T}} + \boxed{Q_A^{-1} X^T Q_B} C_B = Q_A^{-1} C Q_B^{-T}$$
- But,
$$\boxed{\left(Q_A^T X Q_B^{-T}\right)^T} \neq \boxed{Q_A^{-1} X^T Q_B}$$
- with equality only if
$$Q_A = Q_B$$

KEY: Canonical forms for congruence DO not work for $AX + X^T B = C$

- $AX + X^T B = C$, with $A \neq B$.
- $A = Q_A C_A Q_A^T$ and $B = Q_B C_B Q_B^T$.
- $$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$
- $$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$
- $$C_A \boxed{Q_A^T X Q_B^{-T}} + \boxed{Q_A^{-1} X^T Q_B} C_B = Q_A^{-1} C Q_B^{-T}$$
- But,
$$\boxed{\left(Q_A^T X Q_B^{-T}\right)^T} \neq \boxed{Q_A^{-1} X^T Q_B}$$
- with equality only if
$$Q_A = Q_B$$

KEY: Canonical forms for congruence DO not work for $AX + X^T B = C$

- $AX + X^T B = C$, with $A \neq B$.
- $A = Q_A C_A Q_A^T$ and $B = Q_B C_B Q_B^T$.
- $$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$
- $$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$
- $$C_A \boxed{Q_A^T X Q_B^{-T}} + \boxed{Q_A^{-1} X^T Q_B} C_B = Q_A^{-1} C Q_B^{-T}$$
- But, $$\boxed{\left(Q_A^T X Q_B^{-T}\right)^T} \neq \boxed{Q_A^{-1} X^T Q_B}$$
- with equality only if $$Q_A = Q_B$$

KEY: Canonical forms for congruence DO not work for $AX + X^T B = C$

- $AX + X^T B = C$, with $A \neq B$.
- $A = Q_A C_A Q_A^T$ and $B = Q_B C_B Q_B^T$.
- $$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$
- $$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$
- $$C_A \boxed{Q_A^T X Q_B^{-T}} + \boxed{Q_A^{-1} X^T Q_B} C_B = Q_A^{-1} C Q_B^{-T}$$
- But,
$$\boxed{\left(Q_A^T X Q_B^{-T}\right)^T} \neq \boxed{Q_A^{-1} X^T Q_B}$$
- with equality only if
$$Q_A = Q_B$$

In stark contrast with classical Sylvester $AX - XB = C$,

where "canonical forms" for similarity work both if $A = B$ or if $A \neq B$:

- $AX - XB = C$.
- $A = Q_A J_A Q_A^{-1}$ and $B = Q_B J_B Q_B^{-1}$, with J_A and J_B JCFs.
- $$Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C$$
- $$J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$$
- $$J_A \, Q_A^{-1} X Q_B - Q_A^{-1} X Q_B \, J_B = Q_A^{-1} C Q_B$$

"Canonical forms" to be used:

1. For theory: JCF.
2. For computations: Schur form

In stark contrast with classical Sylvester $AX - XB = C$,

where "canonical forms" for similarity work both if $A = B$ or if $A \neq B$:

- $AX - XB = C$.
- $A = Q_A J_A Q_A^{-1}$ and $B = Q_B J_B Q_B^{-1}$, with J_A and J_B JCFs.
- $$Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C$$
- $$J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$$
- $$J_A \boxed{Q_A^{-1} X Q_B} - \boxed{Q_A^{-1} X Q_B} J_B = Q_A^{-1} C Q_B$$

"Canonical forms" to be used:

1. For theory: JCF.
2. For computations: Schur form

In stark contrast with classical Sylvester $AX - XB = C$,

where "canonical forms" for similarity work both if $A = B$ or if $A \neq B$:

- $AX - XB = C$.
- $A = Q_A J_A Q_A^{-1}$ and $B = Q_B J_B Q_B^{-1}$, with J_A and J_B JCFs.
- $$Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C$$
- $$J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$$
- $$J_A \boxed{Q_A^{-1} X Q_B} - \boxed{Q_A^{-1} X Q_B} J_B = Q_A^{-1} C Q_B$$

"Canonical forms" to be used:

1. For theory: JCF.
2. For computations: Schur form

In stark contrast with classical Sylvester $AX - XB = C$,

where "canonical forms" for similarity work both if $A = B$ or if $A \neq B$:

- $AX - XB = C$.
- $A = Q_A J_A Q_A^{-1}$ and $B = Q_B J_B Q_B^{-1}$, with J_A and J_B JCFs.
- $$Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C$$
- $$J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$$
- $$J_A \boxed{Q_A^{-1} X Q_B} - \boxed{Q_A^{-1} X Q_B} J_B = Q_A^{-1} C Q_B$$

"Canonical forms" to be used:

1. For theory: JCF.
2. For computations: Schur form

In stark contrast with classical Sylvester $AX - XB = C$,

where "canonical forms" for similarity work both if $A = B$ or if $A \neq B$:

- $AX - XB = C$.
- $A = Q_A J_A Q_A^{-1}$ and $B = Q_B J_B Q_B^{-1}$, with J_A and J_B JCFs.
- $$Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C$$
- $$J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$$
- $$J_A \boxed{Q_A^{-1} X Q_B} - \boxed{Q_A^{-1} X Q_B} J_B = Q_A^{-1} C Q_B$$

"Canonical forms" to be used:

1. For theory: JCF.
2. For computations: Schur form

In stark contrast with classical Sylvester $AX - XB = C$,

where "canonical forms" for similarity work both if $A = B$ or if $A \neq B$:

- $AX - XB = C$.
- $A = Q_A J_A Q_A^{-1}$ and $B = Q_B J_B Q_B^{-1}$, with J_A and J_B JCFs.
- $$Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C$$
- $$J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$$
- $$J_A \boxed{Q_A^{-1} X Q_B} - \boxed{Q_A^{-1} X Q_B} J_B = Q_A^{-1} C Q_B$$

"Canonical forms" to be used:

1. For theory: JCF.
2. For computations: Schur form

Equivalence of pencil $A - \lambda B^T$

- $AX + X^T B = C$, with $A \neq B$.
- $A - \lambda B^T = PRQ - \lambda PSQ = P(R - \lambda S)Q$, with P and Q nonsingular.
- $$PRQX + X^T Q^T S^T P^T = C$$
- $$RQXP^{-T} + P^{-1} X^T Q^T S^T = P^{-1} C P^{-T}$$
- $$R\, QXP^{-T} + \left(QXP^{-T}\right)^T S^T = P^{-1} C P^{-T}$$

"Canonical forms" for pencils to be used:

1. For theory: KCF (Kronecker Canonical Form).
2. For computations: Generalized Schur form.

Equivalence of pencil $A - \lambda B^T$

- $AX + X^T B = C$, with $A \neq B$.
- $A - \lambda B^T = PRQ - \lambda PSQ = P(R - \lambda S)Q$, with P and Q nonsingular.
- $$PRQX + X^T Q^T S^T P^T = C$$
- $$RQXP^{-T} + P^{-1} X^T Q^T S^T = P^{-1} C P^{-T}$$
- $$R\, QXP^{-T} + \left(QXP^{-T}\right)^T S^T = P^{-1} C P^{-T}$$

"Canonical forms" for pencils to be used:

1. For theory: KCF (Kronecker Canonical Form).
2. For computations: Generalized Schur form.

Equivalence of pencil $A - \lambda B^T$

- $AX + X^T B = C$, with $A \neq B$.
- $A - \lambda B^T = PRQ - \lambda PSQ = P(R - \lambda S)Q$, with P and Q nonsingular.
- $$PRQX + X^T Q^T S^T P^T = C$$
- $$RQXP^{-T} + P^{-1}X^T Q^T S^T = P^{-1}CP^{-T}$$
- $$R\,\boxed{QXP^{-T}} + \boxed{(QXP^{-T})^T}\,S^T = P^{-1}CP^{-T}$$

"Canonical forms" for pencils to be used:

1. For theory: KCF (Kronecker Canonical Form).
2. For computations: Generalized Schur form.

The proper transformation for $AX + X^T B = C$

Equivalence of pencil $A - \lambda B^T$

- $AX + X^T B = C$, with $A \neq B$.
- $A - \lambda B^T = PRQ - \lambda PSQ = P(R - \lambda S)Q$, with P and Q nonsingular.
- $$PRQX + X^T Q^T S^T P^T = C$$
- $$RQXP^{-T} + P^{-1}X^T Q^T S^T = P^{-1}CP^{-T}$$
- $$R\, QXP^{-T} + (QXP^{-T})^T S^T = P^{-1}CP^{-T}$$

"Canonical forms" for pencils to be used:

1. For theory: KCF (Kronecker Canonical Form).
2. For computations: Generalized Schur form.

Equivalence of pencil $A - \lambda B^T$

- $AX + X^T B = C$, with $A \neq B$.
- $A - \lambda B^T = PRQ - \lambda PSQ = P(R - \lambda S)Q$, with P and Q nonsingular.
- $$PRQX + X^T Q^T S^T P^T = C$$
- $$RQXP^{-T} + P^{-1} X^T Q^T S^T = P^{-1} C P^{-T}$$
- $$R \boxed{QXP^{-T}} + \boxed{\left(QXP^{-T}\right)^T} S^T = P^{-1} C P^{-T}$$

"Canonical forms" for pencils to be used:

1. For theory: KCF (Kronecker Canonical Form).
2. For computations: Generalized Schur form.

Equivalence of pencil $A - \lambda B^T$

- $AX + X^T B = C$, with $A \neq B$.
- $A - \lambda B^T = PRQ - \lambda PSQ = P(R - \lambda S)Q$, with P and Q nonsingular.
- $$PRQX + X^T Q^T S^T P^T = C$$
- $$RQXP^{-T} + P^{-1}X^T Q^T S^T = P^{-1}CP^{-T}$$
- $$R\,\boxed{QXP^{-T}} + \boxed{\left(QXP^{-T}\right)^T} S^T = P^{-1}CP^{-T}$$

"Canonical forms" for pencils to be used:

1. For theory: KCF (Kronecker Canonical Form).
2. For computations: Generalized Schur form.

Outline

1. **Previous and related work**
2. **The equation $AX^T + XA = 0$**
   - Motivation: Orbits and the computation of canonical forms
   - Strategy for solving $AX^T + XA = 0$
   - The canonical form for congruence
   - The solution of $AX^T + XA = 0$
   - Generic canonical structure for congruence
3. **The general equation $AX + X^\star B = C$**
   - Motivation
   - Consistency of the Sylvester equation for $\star$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions
4. **General solution of $AX + X^\star B = 0$**
5. **Conclusions**

It is well known that given a block upper triangular matrix (computed by the QR-algorithm for eigenvalues), then

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} A & \boxed{C - (AX - XB)} \\ 0 & B \end{bmatrix}.$$

Therefore, to find a solution of the **Sylvester equation** $AX - XB = C$ allows us to block-diagonalize block-triangular matrices via similarity

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & \boxed{0} \\ 0 & B \end{bmatrix}.$$

This is indeed done in practice in numerical algorithms (LAPACK, MATLAB) to compute bases of invariant subspaces (eigenvectors) of matrices, via the classical Bartels-Stewart algorithm (Comm ACM, 1972) or level-3 BLAS variants of it Jonsson-Kågström (ACM TMS, 2002).

It is well known that given a block upper triangular matrix (computed by the QR-algorithm for eigenvalues), then

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} A & \boxed{C - (AX - XB)} \\ 0 & B \end{bmatrix}.$$

Therefore, to find a solution of the **Sylvester equation** $AX - XB = C$ allows us to block-diagonalize block-triangular matrices via similarity

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & \boxed{0} \\ 0 & B \end{bmatrix}.$$

This is indeed done in practice in numerical algorithms (LAPACK, MATLAB) to compute bases of invariant subspaces (eigenvectors) of matrices, via the classical Bartels-Stewart algorithm (Comm ACM, 1972) or level-3 BLAS variants of it Jonsson-Kågström (ACM TMS, 2002).

Motivation for studying $AX + X^{\star}B = C$ (I)

It is well known that given a block upper triangular matrix (computed by the QR-algorithm for eigenvalues), then

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} A & \boxed{C - (AX - XB)} \\ 0 & B \end{bmatrix}.$$

Therefore, to find a solution of the **Sylvester equation** $AX - XB = C$ allows us to block-diagonalize block-triangular matrices via similarity

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & \boxed{0} \\ 0 & B \end{bmatrix}.$$

This is indeed done in practice in numerical algorithms (LAPACK, MATLAB) to compute bases of invariant subspaces (eigenvectors) of matrices, via the classical Bartels-Stewart algorithm (Comm ACM, 1972) or level-3 BLAS variants of it Jonsson-Kågström (ACM TMS, 2002).

Structured numerical algorithms for linear palindromic eigenproblems ($Z + \lambda Z^\star$) compute an **anti-triangular Schur form** via unitary $\star$-congruence:

Theorem (Kressner, Schröder, Watkins (Numer. Alg., 2009) and Mackey^2, Mehl, Mehrmann (NLAA, 2009))

Let $Z \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$M = U^\star Z\, U = \begin{bmatrix} * & \cdots & \cdots & * \\ \vdots & & \iddots & 0 \\ \vdots & \iddots & \iddots & \vdots \\ * & 0 & \cdots & 0 \end{bmatrix}$$

M can be computed via structure-preserving methods (Kressner, Schröder, Watkins (Numer. Alg., 2009)) or (Mackey^2, Mehl, Mehrmann (NLAA, 2009)) and compute eigenvalues of $Z + \lambda Z^\star$ with exact pairing $\lambda, 1/\lambda^\star$.

Structured numerical algorithms for linear palindromic eigenproblems ($Z + \lambda Z^\star$) compute an **anti-triangular Schur form** via unitary $\star$-congruence:

Theorem (Kressner, Schröder, Watkins (Numer. Alg., 2009) and Mackey^2, Mehl, Mehrmann (NLAA, 2009))

Let $Z \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$M = U^\star\, Z\, U = \begin{bmatrix} * & \cdots & \cdots & * \\ \vdots & & \iddots & 0 \\ \vdots & \iddots & \iddots & \vdots \\ * & 0 & \cdots & 0 \end{bmatrix}$$

M can be computed via structure-preserving methods (Kressner, Schröder, Watkins (Numer. Alg., 2009)) or (Mackey^2, Mehl, Mehrmann (NLAA, 2009)) and compute eigenvalues of $Z + \lambda Z^\star$ with exact pairing $\lambda, 1/\lambda^\star$.

Structured numerical algorithms for linear palindromic eigenproblems ($Z + \lambda Z^{\star}$) compute an **anti-triangular Schur form** via unitary $\star$-congruence:

Theorem (Kressner, Schröder, Watkins (Numer. Alg., 2009) and Mackey^2, Mehl, Mehrmann (NLAA, 2009))

Let $Z \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$M = U^{\star}\,Z\,U = \begin{bmatrix} * & \cdots & \cdots & * \\ \vdots & & \iddots & 0 \\ \vdots & \iddots & \iddots & \vdots \\ * & 0 & \cdots & 0 \end{bmatrix}$$

M can be computed via structure-preserving methods (Kressner, Schröder, Watkins (Numer. Alg., 2009)) or (Mackey^2, Mehl, Mehrmann (NLAA, 2009)) and compute eigenvalues of $Z + \lambda Z^{\star}$ with exact pairing $\lambda, 1/\lambda^{\star}$.

Motivation for studying $AX + X^\star B = C$ (III)

Given a block upper ANTI-triangular matrix (computed via structured algorithms for linear palindromic eigenproblems, when the matrix is real or several eigenvalues form a cluster), then

$$\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}^\star \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} = \begin{bmatrix} C - (AX + X^\star B) & A \\ B & 0 \end{bmatrix}.$$

Therefore, to find a solution of the **Sylvester equation for $\star$-congruence** allows us to block-**ANTI**-diagonalize block-**ANTI**-triangular matrices via $\star$-congruence

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

and to compute deflating subspaces of palindromic pencils.

Given a block upper ANTI-triangular matrix (computed via structured algorithms for linear palindromic eigenproblems, when the matrix is real or several eigenvalues form a cluster), then

$$\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}^\star \begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} = \begin{bmatrix} C - (A\,X + X^\star B) & A \\ B & 0 \end{bmatrix}.$$

Therefore, to find a solution of the **Sylvester equation for $\star$-congruence** allows us to block-**ANTI**-diagonalize block-**ANTI**-triangular matrices via $\star$-congruence

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

and to compute deflating subspaces of palindromic pencils.

Theorem (Wimmer (LAA, 1994), De Terán and D. (ELA, 2011))

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

$$AX + X^\star B = C$$

if and only if

$$\begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \quad and \quad \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$$ *are $\star$-congruent.*

Remarks:

- The implication $\Longrightarrow$ very easy: done in previous slide.
- The implication $\Longleftarrow$ more challenging.
- Wimmer proved in 1994 the result, for $\mathbb{F} = \mathbb{C}$ and $\star = *$, without any reference to palindromic eigenproblems.
- His motivation was the study of standard Sylvester equations with Hermitian solutions.

Theorem (Wimmer (LAA, 1994), De Terán and D. (ELA, 2011))

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

$$AX + X^\star B = C$$

if and only if

$$\begin{bmatrix} C & A \\ B & 0 \end{bmatrix} \quad and \quad \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$$ *are $\star$-congruent.*

Remarks:

- The implication $\Longrightarrow$ very easy: done in previous slide.
- The implication $\Longleftarrow$ more challenging.
- Wimmer proved in 1994 the result, for $\mathbb{F} = \mathbb{C}$ and $\star = *$, without any reference to palindromic eigenproblems.
- His motivation was the study of standard Sylvester equations with Hermitian solutions.

Theorem (Roth (Proc. AMS, 1952))

Let $\mathbb{F}$ be any field and let $A \in \mathbb{F}^{m\times m}$, $B \in \mathbb{F}^{n\times n}$, $C \in \mathbb{F}^{m\times n}$ be given. There is some $X \in \mathbb{F}^{m\times n}$ such that

$$A\,X \;-\; X\,B = C$$

if and only if

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

Outline

1. Previous and related work
2. The equation $AX^T + XA = 0$
   - Motivation: Orbits and the computation of canonical forms
   - Strategy for solving $AX^T + XA = 0$
   - The canonical form for congruence
   - The solution of $AX^T + XA = 0$
   - Generic canonical structure for congruence
3. **The general equation $AX + X^\star B = C$**
   - Motivation
   - Consistency of the Sylvester equation for $\star$-congruence
   - **Uniqueness of solutions**
   - Efficient and stable algorithm to compute unique solutions
4. General solution of $AX + X^\star B = 0$
5. Conclusions

Remarks:

- If the matrices $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times m}$ **are rectangular** ($m \neq n$), then the equation **does not have a unique solution for every right-hand side** C,
- that is, **the operator**

$$\begin{array}{ccc} \mathbb{F}^{n\times m} & \longrightarrow & \mathbb{F}^{m\times m} \\ X & \longmapsto & A\,X + X^\star B \end{array}$$

  is never invertible.
- It is of course possible that $m > n$ and that for particular A, B and C, a solution exists and is unique,
- but **we restrict ourselves here to the square case** $m = n$.

Remarks:

- If the matrices $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times m}$ **are rectangular** ($m \neq n$), then the equation **does not have a unique solution for every right-hand side C**,
- that is, **the operator**

$$\begin{array}{ccc} \mathbb{F}^{n\times m} & \longrightarrow & \mathbb{F}^{m\times m} \\ X & \longmapsto & AX + X^\star B \end{array}$$

is never invertible.

- It is of course possible that $m > n$ and that for particular A, B and C, a solution exists and is unique,
- but **we restrict ourselves here to the square case $m = n$**.

Remarks:

- If the matrices $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times m}$ **are rectangular** $(m \neq n)$, then the equation **does not have a unique solution for every right-hand side C**,
- that is, **the operator**

$$\begin{array}{ccc} \mathbb{F}^{n\times m} & \longrightarrow & \mathbb{F}^{m\times m} \\ X & \longmapsto & AX + X^\star B \end{array}$$

  is never invertible.
- It is of course possible that $m > n$ and that for particular A, B and C, a solution exists and is unique,
- but **we restrict ourselves here to the square case $m = n$**.

Remarks:

- If the matrices $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times m}$ **are rectangular** ($m \neq n$), then the equation **does not have a unique solution for every right-hand side C**,
- that is, **the operator**

$$\begin{array}{ccc} \mathbb{F}^{n\times m} & \longrightarrow & \mathbb{F}^{m\times m} \\ X & \longmapsto & A\,X + X^\star B \end{array}$$

  is never invertible.
- It is of course possible that $m > n$ and that for particular A, B and C, a solution exists and is unique,
- but **we restrict ourselves here to the square case $m = n$**.

Uniqueness of solutions of $A\,X + X^\star\,B = C$ (II)

Definition: a set $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is *$\star$-reciprocal free* if $\lambda_i \neq 1/\lambda_j^\star$ for any $1 \leq i, j \leq n$. We admit 0 and/or ∞ as elements of $\{\lambda_1, \ldots, \lambda_n\}$.

Theorem (Byers, Kressner (SIMAX, 2006), Kressner, Schröder, Watkins, (Num. Alg., 2009))

Let $A, B \in \mathbb{C}^{n\times n}$ be given. Then:

- *$A\,X + X^T\,B = C$ has a unique solution X for every right-hand side $C \in \mathbb{C}^{n\times n}$ if and only if the following conditions hold:*
  1) *The pencil $A - \lambda B^T$ is regular, and*
  2) *the set of eigenvalues of $A - \lambda B^T \setminus \{1\}$ is T-reciprocal free and if 1 is an eigenvalue of $A - \lambda B^T$, then it has algebraic multiplicity 1.*
- *$A\,X + X^*\,B = C$ has a unique solution X for every right-hand side $C \in \mathbb{C}^{n\times n}$ if and only if the following conditions hold:*
  1) *The pencil $A - \lambda B^*$ is regular, and*
  2) *the set of eigenvalues of $A - \lambda B^*$ is $*$-reciprocal free.*

Definition: a set $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is *$\star$-reciprocal free* if $\lambda_i \neq 1/\lambda_j^\star$ for any $1 \leq i, j \leq n$. We admit 0 and/or ∞ as elements of $\{\lambda_1, \ldots, \lambda_n\}$.

Theorem (Byers, Kressner (SIMAX, 2006), Kressner, Schröder, Watkins, (Num. Alg., 2009))

Let $A, B \in \mathbb{C}^{n\times n}$ be given. Then:

- *$AX + X^T B = C$ has a unique solution X for every right-hand side $C \in \mathbb{C}^{n\times n}$ if and only if the following conditions hold:*
  1) *The pencil $A - \lambda B^T$ is regular, and*
  2) *the set of eigenvalues of $A - \lambda B^T \setminus \{1\}$ is T-reciprocal free and if 1 is an eigenvalue of $A - \lambda B^T$, then it has algebraic multiplicity 1.*
- *$AX + X^* B = C$ has a unique solution X for every right-hand side $C \in \mathbb{C}^{n\times n}$ if and only if the following conditions hold:*
  1) *The pencil $A - \lambda B^*$ is regular, and*
  2) *the set of eigenvalues of $A - \lambda B^*$ is $*$-reciprocal free.*

Definition: a set $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is *$\star$-reciprocal free* if $\lambda_i \neq 1/\lambda_j^\star$ for any $1 \leq i, j \leq n$. We admit 0 and/or ∞ as elements of $\{\lambda_1, \ldots, \lambda_n\}$.

Theorem (Byers, Kressner (SIMAX, 2006), Kressner, Schröder, Watkins, (Num. Alg., 2009))

Let $A, B \in \mathbb{C}^{n\times n}$ be given. Then:

- $AX + X^T B = C$ *has a unique solution X for every right-hand side $C \in \mathbb{C}^{n\times n}$ if and only if the following conditions hold:*
  1) *The pencil $A - \lambda B^T$ is regular, and*
  2) *the set of eigenvalues of $A - \lambda B^T \setminus \{1\}$ is T-reciprocal free and if 1 is an eigenvalue of $A - \lambda B^T$, then it has algebraic multiplicity 1.*
- $AX + X^* B = C$ *has a unique solution X for every right-hand side $C \in \mathbb{C}^{n\times n}$ if and only if the following conditions hold:*
  1) *The pencil $A - \lambda B^*$ is regular, and*
  2) *the set of eigenvalues of $A - \lambda B^*$ is $*$-reciprocal free.*

Theorem

Let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be given. Then:

- *$AX - XB = C$ has a unique solution X for every right-hand side $C \in \mathbb{C}^{m\times n}$ if and only if A and B have no eigenvalues in common.*

Remark: Comparison of both results brings to our attention **a key difference** that appears always between solution methods for $AX + X^\star B = C$ and $AX - XB = C$:

- In $AX + X^\star B = C$, one starts by dealing with the eigenproblem of $A - \lambda B^\star$, that is, one deals from the very beginning **simultaneously with A and B**.
- By contrast in $AX - XB = C$, one starts by dealing **independently** with the eigenproblems **of A and B**.

Theorem

Let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be given. Then:

- $A\,X \;-\; X\,B = C$ *has a unique solution X for every right-hand side $C \in \mathbb{C}^{m\times n}$ if and only if A and B have no eigenvalues in common.*

Remark: Comparison of both results brings to our attention **a key difference** that appears always between solution methods for $AX + X^\star B = C$ and $AX - XB = C$:

- In $AX + X^\star B = C$, one starts by dealing with the eigenproblem of $A - \lambda B^\star$, that is, one deals from the very beginning **simultaneously with A and B**.
- By contrast in $AX - XB = C$, one starts by dealing **independently** with the eigenproblems **of A and B**.

Theorem

Let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be given. Then:

- *$A\,X \;-\; X\,B = C$ has a unique solution X for every right-hand side $C \in \mathbb{C}^{m\times n}$ if and only if A and B have no eigenvalues in common.*

Remark: Comparison of both results brings to our attention **a key difference** that appears always between solution methods for $AX + X^\star B = C$ and $AX - XB = C$:

- In $AX + X^\star B = C$, one starts by dealing with the eigenproblem of $A - \lambda B^\star$, that is, one deals from the very beginning **simultaneously with A and B**.
- By contrast in $AX - XB = C$, one starts by dealing **independently** with the eigenproblems **of A and B**.

Theorem

Let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be given. Then:

- *$AX - XB = C$ has a unique solution X for every right-hand side $C \in \mathbb{C}^{m\times n}$ if and only if A and B have no eigenvalues in common.*

Remark: Comparison of both results brings to our attention **a key difference** that appears always between solution methods for $AX + X^{\star}B = C$ and $AX - XB = C$:

- In $AX + X^{\star}B = C$, one starts by dealing with the eigenproblem of $A - \lambda B^{\star}$, that is, one deals from the very beginning **simultaneously with A and B**.
- By contrast in $AX - XB = C$, one starts by dealing **independently** with the eigenproblems **of A and B**.

- In this section in $A\,X\;+\;X^{\star}\,B = C$ all matrices are in $\mathbb{C}^{n\times n}$ and the solution is unique for every C.
- $A\,X\;+\;X^{\star}\,B = C$ is equivalent to a linear system for the n^2 entries of X if $\star = T$ and to a linear system for the $2\,n^2$ entries of $(\mathcal{R}e\,X\,,\,\mathcal{I}m\,X)$ if $\star = *$. From now on, we say simply "linear system" for X.
- Then, it is possible to use Gaussian elimination on the equivalent system, but it costs $O(n^6)$ flops, which is not feasible except for small n.
- **IDEA: transform $A\,X\;+\;X^{\star}\,B = C$ into an equation of the same type but with much simpler coefficients instead of A and B and that can be easily solved to get a total cost of $O(n^3)$ flops.**
- To this purpose, use **QZ algorithm** to compute in $O(n^3)$ flops the **generalized Schur decomposition** of

$$A - \lambda B^{\star} = U(R - \lambda S)V\,, \quad \text{where} \quad \begin{cases} R,\,S & \text{are upper triangular} \\ U,\,V & \text{are unitary matrices} \end{cases}$$

If A, B real matrices: use real arithmetic to get *quasi-triangular* R. We do not consider this for brevity.

The fundamental transformation

- In this section in $A\,X \,+\, X^\star\,B = C$ all matrices are in $\mathbb{C}^{n\times n}$ and the solution is unique for every C.
- $A\,X \,+\, X^\star\,B = C$ is equivalent to a linear system for the n^2 entries of X if $\star = T$ and to a linear system for the $2\,n^2$ entries of $(\mathcal{R}e\,X\,,\,\mathcal{I}m\,X)$ if $\star = *$. From now on, we say simply "linear system" for X.
- Then, it is possible to use Gaussian elimination on the equivalent system, but it costs $O(n^6)$ flops, which is not feasible except for small n.
- **IDEA: transform $A\,X \,+\, X^\star\,B = C$ into an equation of the same type but with much simpler coefficients instead of A and B and that can be easily solved to get a total cost of $O(n^3)$ flops.**
- To this purpose, use **QZ algorithm** to compute in $O(n^3)$ flops the **generalized Schur decomposition** of

$$A - \lambda B^\star = U(R - \lambda S)V\,, \quad \text{where} \quad \begin{cases} R,\, S & \text{are upper triangular} \\ U,\, V & \text{are unitary matrices} \end{cases}$$

If A, B real matrices: use real arithmetic to get *quasi-triangular* R. We do not consider this for brevity.

The fundamental transformation

- In this section in $A\,X + X^\star\,B = C$ all matrices are in $\mathbb{C}^{n\times n}$ and the solution is unique for every C.
- $A\,X + X^\star\,B = C$ is equivalent to a linear system for the n^2 entries of X if $\star = T$ and to a linear system for the $2\,n^2$ entries of $(\mathcal{R}e\,X\,,\,\mathcal{I}m\,X)$ if $\star = *$. From now on, we say simply "linear system" for X.
- Then, it is possible to use Gaussian elimination on the equivalent system, but it costs $O(n^6)$ flops, which is not feasible except for small n.
- **IDEA: transform $A\,X + X^\star\,B = C$ into an equation of the same type but with much simpler coefficients instead of A and B and that can be easily solved to get a total cost of $O(n^3)$ flops.**
- To this purpose, use **QZ algorithm** to compute in $O(n^3)$ flops the **generalized Schur decomposition** of

$$A - \lambda B^\star = U(R - \lambda S)V\,, \quad \text{where} \quad \begin{cases} R,\,S & \text{are upper triangular} \\ U,\,V & \text{are unitary matrices} \end{cases}$$

If A, B real matrices: use real arithmetic to get *quasi-triangular* R. We do not consider this for brevity.

- In this section in $A\,X + X^\star\,B = C$ all matrices are in $\mathbb{C}^{n\times n}$ and the solution is unique for every C.
- $A\,X + X^\star\,B = C$ is equivalent to a linear system for the n^2 entries of X if $\star = T$ and to a linear system for the $2\,n^2$ entries of $(\mathcal{R}e\,X\,,\,\mathcal{I}m\,X)$ if $\star = *$. From now on, we say simply "linear system" for X.
- Then, it is possible to use Gaussian elimination on the equivalent system, but it costs $O(n^6)$ flops, which is not feasible except for small n.
- **IDEA: transform $A\,X + X^\star\,B = C$ into an equation of the same type but with much simpler coefficients instead of A and B and that can be easily solved to get a total cost of $O(n^3)$ flops.**
- To this purpose, use **QZ algorithm** to compute in $O(n^3)$ flops the **generalized Schur decomposition** of

$$A - \lambda B^\star = U(R - \lambda S)V\,, \quad \text{where} \quad \begin{cases} R,\,S & \text{are upper triangular} \\ U,\,V & \text{are unitary matrices} \end{cases}$$

If A, B real matrices: use real arithmetic to get *quasi-triangular* R. We do not consider this for brevity.

- In this section in $A\,X + X^\star B = C$ all matrices are in $\mathbb{C}^{n\times n}$ and the solution is unique for every C.
- $A\,X + X^\star B = C$ is equivalent to a linear system for the n^2 entries of X if $\star = T$ and to a linear system for the $2\,n^2$ entries of $(\mathcal{R}e\,X\,,\,\mathcal{I}m\,X)$ if $\star = *$. From now on, we say simply "linear system" for X.
- Then, it is possible to use Gaussian elimination on the equivalent system, but it costs $O(n^6)$ flops, which is not feasible except for small n.
- **IDEA: transform $A\,X + X^\star B = C$ into an equation of the same type but with much simpler coefficients instead of A and B and that can be easily solved to get a total cost of $O(n^3)$ flops.**
- To this purpose, use **QZ algorithm** to compute in $O(n^3)$ flops the **generalized Schur decomposition** of

$$A - \lambda B^\star = U(R - \lambda S)V\,, \quad \text{where} \quad \begin{cases} R,\,S & \text{are upper triangular} \\ U,\,V & \text{are unitary matrices} \end{cases}$$

If A, B real matrices: use real arithmetic to get *quasi-triangular* R. We do not consider this for brevity.

INPUT: $A, B, C \in \mathbb{C}^{n\times n}$
OUTPUT: $X \in \mathbb{C}^{n\times n}$

Step 1. Compute via QZ algorithm R, S, U and V such that

$$A = URV\,, \quad B^\star = USV\,, \text{ where } \begin{cases} R,\ S & \text{are upper triangular} \\ U,\ V & \text{are unitary matrices} \end{cases}$$

Step 2. Compute $E = U^* C\,(U^\star)^*$

Step 3. Solve for $W \in \mathbb{C}^{n\times n}$ the transformed equation

$$RW + W^\star S^\star = E$$

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

INPUT: $A, B, C \in \mathbb{C}^{n\times n}$
OUTPUT: $X \in \mathbb{C}^{n\times n}$

Step 1. Compute via QZ algorithm R, S, U and V such that

$$A = URV\,, \quad B^\star = USV\,, \text{ where } \begin{cases} R,\ S & \text{are upper triangular} \\ U,\ V & \text{are unitary matrices} \end{cases}$$

Step 2. Compute $E = U^*\,C\,(U^\star)^*$

Step 3. How to solve for $W \in \mathbb{C}^{n\times n}$ the transformed equation

$$RW + W^\star S^\star = E \text{ ?}$$

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

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w_{11}^\star & w_{21}^\star & w_{31}^\star & w_{41}^\star \\ w_{12}^\star & w_{22}^\star & w_{32}^\star & w_{42}^\star \\ w_{13}^\star & w_{23}^\star & w_{33}^\star & w_{43}^\star \\ w_{14}^\star & w_{24}^\star & w_{34}^\star & w_{44}^\star \end{bmatrix} \begin{bmatrix} s_{11}^\star & 0 & 0 & 0 \\ s_{12}^\star & s_{22}^\star & 0 & 0 \\ s_{13}^\star & s_{23}^\star & s_{33}^\star & 0 \\ s_{14}^\star & s_{24}^\star & s_{34}^\star & s_{44}^\star \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (4,4)-entry, then we get

$$r_{44} \, w_{44} + w_{44}^\star \, s_{44}^\star = e_{44} \, ,$$

a scalar equation that allows us to determine w_{44}.

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w^{\star}_{11} & w^{\star}_{21} & w^{\star}_{31} & w^{\star}_{41} \\ w^{\star}_{12} & w^{\star}_{22} & w^{\star}_{32} & w^{\star}_{42} \\ w^{\star}_{13} & w^{\star}_{23} & w^{\star}_{33} & w^{\star}_{43} \\ w^{\star}_{14} & w^{\star}_{24} & w^{\star}_{34} & w^{\star}_{44} \end{bmatrix} \begin{bmatrix} s^{\star}_{11} & 0 & 0 & 0 \\ s^{\star}_{12} & s^{\star}_{22} & 0 & 0 \\ s^{\star}_{13} & s^{\star}_{23} & s^{\star}_{33} & 0 \\ s^{\star}_{14} & s^{\star}_{24} & s^{\star}_{34} & s^{\star}_{44} \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (4,4)-entry, then we get

$$r_{44} \; \boxed{w_{44}} + \boxed{w^{\star}_{44}} \; s^{\star}_{44} = e_{44} \; ,$$

a scalar equation that allows us to determine $\boxed{w_{44}}$.

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w^\star_{11} & w^\star_{21} & w^\star_{31} & w^\star_{41} \\ w^\star_{12} & w^\star_{22} & w^\star_{32} & w^\star_{42} \\ w^\star_{13} & w^\star_{23} & w^\star_{33} & w^\star_{43} \\ w^\star_{14} & w^\star_{24} & w^\star_{34} & w^\star_{44} \end{bmatrix} \begin{bmatrix} s^\star_{11} & 0 & 0 & 0 \\ s^\star_{12} & s^\star_{22} & 0 & 0 \\ s^\star_{13} & s^\star_{23} & s^\star_{33} & 0 \\ s^\star_{14} & s^\star_{24} & s^\star_{34} & s^\star_{44} \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (3,4) and (4,3) entries, then we get

$$\begin{aligned} s_{33}\, w_{34} + w^\star_{43}\, r^\star_{44} &= e^\star_{43} - s_{34}\, w_{44} \\ r_{33}\, w_{34} + w^\star_{43}\, s^\star_{44} &= e_{34} - r_{34}\, w_{44} \end{aligned},$$

a 2×2 system of scalar equations that allows us to determine w_{34} and w_{43} **simultaneously**.

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w_{11}^\star & w_{21}^\star & w_{31}^\star & w_{41}^\star \\ w_{12}^\star & w_{22}^\star & w_{32}^\star & w_{42}^\star \\ w_{13}^\star & w_{23}^\star & w_{33}^\star & w_{43}^\star \\ w_{14}^\star & w_{24}^\star & w_{34}^\star & w_{44}^\star \end{bmatrix} \begin{bmatrix} s_{11}^\star & 0 & 0 & 0 \\ s_{12}^\star & s_{22}^\star & 0 & 0 \\ s_{13}^\star & s_{23}^\star & s_{33}^\star & 0 \\ s_{14}^\star & s_{24}^\star & s_{34}^\star & s_{44}^\star \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (3,4) and (4,3) entries, then we get

$$\begin{aligned} s_{33}\, w_{34} + w_{43}^\star\, r_{44}^\star &= e_{43}^\star - s_{34}\, w_{44} \\ r_{33}\, w_{34} + w_{43}^\star\, s_{44}^\star &= e_{34} - r_{34}\, w_{44} \end{aligned},$$

a 2×2 system of scalar equations that allows us to determine w_{34} and w_{43} **simultaneously**.

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & \boxed{w_{34}} \\ w_{41} & w_{42} & \boxed{w_{43}} & \boxed{w_{44}} \end{bmatrix}$$

$$+ \begin{bmatrix} w_{11}^\star & w_{21}^\star & w_{31}^\star & w_{41}^\star \\ w_{12}^\star & w_{22}^\star & w_{32}^\star & w_{42}^\star \\ w_{13}^\star & w_{23}^\star & w_{33}^\star & \boxed{w_{43}^\star} \\ w_{14}^\star & w_{24}^\star & \boxed{w_{34}^\star} & \boxed{w_{44}^\star} \end{bmatrix} \begin{bmatrix} s_{11}^\star & 0 & 0 & 0 \\ s_{12}^\star & s_{22}^\star & 0 & 0 \\ s_{13}^\star & s_{23}^\star & s_{33}^\star & 0 \\ s_{14}^\star & s_{24}^\star & s_{34}^\star & s_{44}^\star \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (2,4) and (4,2) entries, then we get

$$\begin{aligned} s_{22}\, w_{24} + w_{42}^\star\, r_{44}^\star &= e_{42}^\star - s_{23}\, w_{34} - s_{24}\, w_{44} \\ r_{22}\, w_{24} + w_{42}^\star\, s_{44}^\star &= e_{24} - r_{23}\, w_{34} - r_{24}\, w_{44} \end{aligned},$$

a 2×2 system of scalar equations that allows us to determine w_{24} and w_{42} **simultaneously**.

Algorithm to solve the transformed equation $RW + W^\star S^\star = E$ (I)

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w^\star_{11} & w^\star_{21} & w^\star_{31} & w^\star_{41} \\ w^\star_{12} & w^\star_{22} & w^\star_{32} & w^\star_{42} \\ w^\star_{13} & w^\star_{23} & w^\star_{33} & w^\star_{43} \\ w^\star_{14} & w^\star_{24} & w^\star_{34} & w^\star_{44} \end{bmatrix} \begin{bmatrix} s^\star_{11} & 0 & 0 & 0 \\ s^\star_{12} & s^\star_{22} & 0 & 0 \\ s^\star_{13} & s^\star_{23} & s^\star_{33} & 0 \\ s^\star_{14} & s^\star_{24} & s^\star_{34} & s^\star_{44} \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (2,4) and (4,2) entries, then we get

$$\begin{aligned} s_{22}\, w_{24} + w^\star_{42}\, r^\star_{44} &= e^\star_{42} - s_{23}\, w_{34} - s_{24}\, w_{44} \\ r_{22}\, w_{24} + w^\star_{42}\, s^\star_{44} &= e_{24} - r_{23}\, w_{34} - r_{24}\, w_{44} \end{aligned} ,$$

a 2×2 system of scalar equations that allows us to determine w_{24} and w_{42} **simultaneously**.

Algorithm to solve the transformed equation $RW + W^\star S^\star = E$ (I)

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w^\star_{11} & w^\star_{21} & w^\star_{31} & w^\star_{41} \\ w^\star_{12} & w^\star_{22} & w^\star_{32} & w^\star_{42} \\ w^\star_{13} & w^\star_{23} & w^\star_{33} & w^\star_{43} \\ w^\star_{14} & w^\star_{24} & w^\star_{34} & w^\star_{44} \end{bmatrix} \begin{bmatrix} s^\star_{11} & 0 & 0 & 0 \\ s^\star_{12} & s^\star_{22} & 0 & 0 \\ s^\star_{13} & s^\star_{23} & s^\star_{33} & 0 \\ s^\star_{14} & s^\star_{24} & s^\star_{34} & s^\star_{44} \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (1,4) and (4,1) entries, then we get

$$\begin{aligned} s_{11}\, w_{14} + w^\star_{41}\, r^\star_{44} &= e^\star_{41} - s_{12}\, w_{24} - s_{13}\, w_{34} - s_{14}\, w_{44} \\ r_{11}\, w_{14} + w^\star_{41}\, s^\star_{44} &= e_{14} - r_{12}\, w_{24} - r_{13}\, w_{34} - r_{14}\, w_{44} \end{aligned},$$

a 2×2 system of scalar equations that allows us to determine w_{14} and w_{41} **simultaneously**.

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w_{11}^\star & w_{21}^\star & w_{31}^\star & w_{41}^\star \\ w_{12}^\star & w_{22}^\star & w_{32}^\star & w_{42}^\star \\ w_{13}^\star & w_{23}^\star & w_{33}^\star & w_{43}^\star \\ w_{14}^\star & w_{24}^\star & w_{34}^\star & w_{44}^\star \end{bmatrix} \begin{bmatrix} s_{11}^\star & 0 & 0 & 0 \\ s_{12}^\star & s_{22}^\star & 0 & 0 \\ s_{13}^\star & s_{23}^\star & s_{33}^\star & 0 \\ s_{14}^\star & s_{24}^\star & s_{34}^\star & s_{44}^\star \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (1,4) and (4,1) entries, then we get

$$\begin{array}{rcl} s_{11}\, w_{14} + w_{41}^\star\, r_{44}^\star & = & e_{41}^\star - s_{12}\, w_{24} - s_{13}\, w_{34} - s_{14}\, w_{44} \\ r_{11}\, w_{14} + w_{41}^\star\, s_{44}^\star & = & e_{14} - r_{12}\, w_{24} - r_{13}\, w_{34} - r_{14}\, w_{44} \end{array},$$

a 2×2 system of scalar equations that allows us to determine w_{14} and w_{41} **simultaneously**.

Algorithm to solve the transformed equation $RW + W^\star S^\star = E$ (I)

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w_{11}^\star & w_{21}^\star & w_{31}^\star & w_{41}^\star \\ w_{12}^\star & w_{22}^\star & w_{32}^\star & w_{42}^\star \\ w_{13}^\star & w_{23}^\star & w_{33}^\star & w_{43}^\star \\ w_{14}^\star & w_{24}^\star & w_{34}^\star & w_{44}^\star \end{bmatrix} \begin{bmatrix} s_{11}^\star & 0 & 0 & 0 \\ s_{12}^\star & s_{22}^\star & 0 & 0 \\ s_{13}^\star & s_{23}^\star & s_{33}^\star & 0 \\ s_{14}^\star & s_{24}^\star & s_{34}^\star & s_{44}^\star \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (1:3,1:3) submatrices , then we get

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} + \begin{bmatrix} w_{11}^\star & w_{21}^\star & w_{31}^\star \\ w_{12}^\star & w_{22}^\star & w_{32}^\star \\ w_{13}^\star & w_{23}^\star & w_{33}^\star \end{bmatrix} \begin{bmatrix} s_{11}^\star & 0 & 0 \\ s_{12}^\star & s_{22}^\star & 0 \\ s_{13}^\star & s_{23}^\star & s_{33}^\star \end{bmatrix}$$

$$= \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} - \begin{bmatrix} r_{14} \\ r_{24} \\ r_{34} \end{bmatrix} \begin{bmatrix} w_{41} & w_{42} & w_{43} \end{bmatrix} - \begin{bmatrix} w_{41}^\star \\ w_{42}^\star \\ w_{43}^\star \end{bmatrix} \begin{bmatrix} s_{14}^\star & s_{24}^\star & s_{34}^\star \end{bmatrix}$$

which is a 3×3 **matrix equation of the same type as the original one**.

We illustrate with 4×4 example for simplicity:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

$$+ \begin{bmatrix} w_{11}^\star & w_{21}^\star & w_{31}^\star & w_{41}^\star \\ w_{12}^\star & w_{22}^\star & w_{32}^\star & w_{42}^\star \\ w_{13}^\star & w_{23}^\star & w_{33}^\star & w_{43}^\star \\ w_{14}^\star & w_{24}^\star & w_{34}^\star & w_{44}^\star \end{bmatrix} \begin{bmatrix} s_{11}^\star & 0 & 0 & 0 \\ s_{12}^\star & s_{22}^\star & 0 & 0 \\ s_{13}^\star & s_{23}^\star & s_{33}^\star & 0 \\ s_{14}^\star & s_{24}^\star & s_{34}^\star & s_{44}^\star \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{bmatrix}$$

If we equate the (1:3,1:3) submatrices , then we get

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} + \begin{bmatrix} w_{11}^\star & w_{21}^\star & w_{31}^\star \\ w_{12}^\star & w_{22}^\star & w_{32}^\star \\ w_{13}^\star & w_{23}^\star & w_{33}^\star \end{bmatrix} \begin{bmatrix} s_{11}^\star & 0 & 0 \\ s_{12}^\star & s_{22}^\star & 0 \\ s_{13}^\star & s_{23}^\star & s_{33}^\star \end{bmatrix}$$

$$= \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} - \begin{bmatrix} r_{14} \\ r_{24} \\ r_{34} \end{bmatrix} \begin{bmatrix} w_{41} & w_{42} & w_{43} \end{bmatrix} - \begin{bmatrix} w_{41}^\star \\ w_{42}^\star \\ w_{43}^\star \end{bmatrix} \begin{bmatrix} s_{14}^\star & s_{24}^\star & s_{34}^\star \end{bmatrix}$$

which is a 3×3 **matrix equation of the same type as the original one**.

- **Cost:** $2n^3 + O(n^2)$ flops for simplified system and a total cost $76n^3 + O(n^2)$ flops for the whole algorithm for $AX + X^\star B = C$.
- **Forward stable algorithm.**
- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:
  1. Compute independently triang. Schur forms T_A and T_B of A and B.
  2. Solve $T_A\,Y - Y\,T_B = D$ for Y.
  3. Recover X from Y.
- Same flavor, but also relevant differences.

Remarks on algorithm to solve $A\,X + X^\star\,B = C$

- **Cost:** $2n^3 + O(n^2)$ flops for simplified system and a total cost $76n^3 + O(n^2)$ flops for the whole algorithm for $AX + X^\star B = C$.
- **Forward stable algorithm.**
- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:
  1. Compute independently triang. Schur forms T_A and T_B of A and B.
  2. Solve $T_A\,Y - Y\,T_B = D$ for Y.
  3. Recover X from Y.
- Same flavor, but also relevant differences.

Remarks on algorithm to solve $A\,X\,+\,X^\star\,B = C$

- **Cost:** $2n^3 + O(n^2)$ flops for simplified system and a total cost $76n^3 + O(n^2)$ flops for the whole algorithm for $AX + X^\star B = C$.
- **Forward stable algorithm.**
- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:
  1. Compute independently triang. Schur forms T_A and T_B of A and B.
  2. Solve $T_A\,Y - Y\,T_B = D$ for Y.
  3. Recover X from Y.
- Same flavor, but also relevant differences.

Remarks on algorithm to solve $A\,X + X^\star\,B = C$

- **Cost:** $2n^3 + O(n^2)$ flops for simplified system and a total cost $76n^3 + O(n^2)$ flops for the whole algorithm for $AX + X^\star B = C$.
- **Forward stable algorithm.**
- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:
  1. Compute independently triang. Schur forms T_A and T_B of A and B.
  2. Solve $T_A\,Y - Y\,T_B = D$ for Y.
  3. Recover X from Y.
- Same flavor, but also relevant differences.

Remarks on algorithm to solve $A\,X + X^\star B = C$

- **Cost:** $2n^3 + O(n^2)$ flops for simplified system and a total cost $76n^3 + O(n^2)$ flops for the whole algorithm for $AX + X^\star B = C$.
- **Forward stable algorithm.**
- The algorithm should be compared with Bartels-Stewart algorithm for Sylvester equation $AX - XB = C$:
  1. Compute independently triang. Schur forms T_A and T_B of A and B.
  2. Solve $T_A\,Y - Y\,T_B = D$ for Y.
  3. Recover X from Y.
- Same flavor, but also relevant differences.

Outline

Theoretical method to solve $AX + X^\star B = 0$ (I)

- In case of consistency, but "nonuniqueness", general solution of $AX + X^\star B = C$ is $X = X_p + X_h$, where
  1. X_p is a particular solution and
  2. X_h is the general solution of $AX + X^\star B = 0$.

  The latter found by De Terán, D., Guillery, Montealegre, Reyes, Lin. Alg. Appl., 2013
- Summer REU program 2011, U. of California at S. Barbara, M.I. Bueno.
- I do not know any clear application for this problem.

- In case of consistency, but "nonuniqueness", general solution of $AX + X^\star B = C$ is $X = X_p + X_h$, where
  1. X_p is a particular solution and
  2. X_h is the general solution of $AX + X^\star B = 0$.

  The latter found by De Terán, D., Guillery, Montealegre, Reyes, Lin. Alg. Appl., 2013
- Summer REU program 2011, U. of California at S. Barbara, M.I. Bueno.
- I do not know any clear application for this problem.

- In case of consistency, but "nonuniqueness", general solution of $AX + X^\star B = C$ is $X = X_p + X_h$, where
  1. X_p is a particular solution and
  2. X_h is the general solution of $AX + X^\star B = 0$.

  The latter found by De Terán, D., Guillery, Montealegre, Reyes, Lin. Alg. Appl., 2013
- Summer REU program 2011, U. of California at S. Barbara, M.I. Bueno.
- I do not know any clear application for this problem.

- **KEY IDEA:** If $E - \lambda F^\star$ is the *Kronecker Canonical form (KCF)* of $A - \lambda B^\star$ for strict equivalence, i.e.,

  $$A - \lambda B^\star = P(E - \lambda F^\star)Q, \qquad \text{with } P \text{ and } Q \text{ nonsingular,}$$

  then $AX + X^\star B = 0$ can be transformed into

  $$EY + Y^\star F = 0, \qquad \text{with } Y = QXP^{-\star}.$$

- If $E = E_1 \oplus \cdots \oplus E_d$, $F^\star = F_1^\star \oplus \cdots \oplus F_d^\star$, and $Y = [Y_{ij}]$ is partitioned into blocks accordingly, then this equation decouples in

  $$E_i Y_{ii} + Y_{ii}^\star F_i = 0 \quad \text{and} \quad \left\{ \begin{array}{l} E_i Y_{ij} + Y_{ji}^\star F_j = 0 \\ E_j Y_{ji} + Y_{ij}^\star F_i = 0 \end{array} \right. , \quad (1 \le i < j \le d).$$

- Since KCF has 4 types of blocks, this produces **14** different types of matrix (systems) equations, whose explicit solutions have been found.

- Much more complicated general solution than standard Sylvester eq: $AX - XB = 0$, which depends on JCF of A and B and requires to solve only one type of equation.

- **KEY IDEA:** If $E - \lambda F^\star$ is the *Kronecker Canonical form (KCF)* of $A - \lambda B^\star$ for strict equivalence, i.e.,

$$A - \lambda B^\star = P(E - \lambda F^\star)Q, \qquad \text{with } P \text{ and } Q \text{ nonsingular},$$

then $AX + X^\star B = 0$ can be transformed into

$$EY + Y^\star F = 0, \qquad \text{with } Y = QXP^{-\star}.$$

- If $E = E_1 \oplus \cdots \oplus E_d$, $F^\star = F_1^\star \oplus \cdots \oplus F_d^\star$, and $Y = [Y_{ij}]$ is partitioned into blocks accordingly, then this equation decouples in

$$E_i Y_{ii} + Y_{ii}^\star F_i = 0 \quad \text{and} \quad \left\{ \begin{array}{l} E_i Y_{ij} + Y_{ji}^\star F_j = 0 \\ E_j Y_{ji} + Y_{ij}^\star F_i = 0 \end{array} \right. , \quad (1 \le i < j \le d).$$

- Since KCF has 4 types of blocks, this produces **14** different types of matrix (systems) equations, whose explicit solutions have been found.
- Much more complicated general solution than standard Sylvester eq: $AX - XB = 0$, which depends on JCF of A and B and requires to solve only one type of equation.

- **KEY IDEA:** If $E - \lambda F^\star$ is the *Kronecker Canonical form (KCF)* of $A - \lambda B^\star$ for strict equivalence, i.e.,

$$A - \lambda B^\star = P(E - \lambda F^\star)Q, \qquad \text{with } P \text{ and } Q \text{ nonsingular,}$$

then $AX + X^\star B = 0$ can be transformed into

$$EY + Y^\star F = 0, \qquad \text{with } Y = QXP^{-\star}.$$

- If $E = E_1 \oplus \cdots \oplus E_d$, $F^\star = F_1^\star \oplus \cdots \oplus F_d^\star$, and $Y = [Y_{ij}]$ is partitioned into blocks accordingly, then this equation decouples in

$$E_i Y_{ii} + Y_{ii}^\star F_i = 0 \quad \text{and} \quad \left\{ \begin{array}{l} E_i Y_{ij} + Y_{ji}^\star F_j = 0 \\ E_j Y_{ji} + Y_{ij}^\star F_i = 0 \end{array} \right. , \quad (1 \leq i < j \leq d).$$

- Since KCF has 4 types of blocks, this produces **14** different types of matrix (systems) equations, whose explicit solutions have been found.

- Much more complicated general solution than standard Sylvester eq: $AX - XB = 0$, which depends on JCF of A and B and requires to solve only one type of equation.

- **KEY IDEA:** If $E - \lambda F^\star$ is the *Kronecker Canonical form (KCF)* of $A - \lambda B^\star$ for strict equivalence, i.e.,

$$A - \lambda B^\star = P(E - \lambda F^\star)Q, \qquad \text{with } P \text{ and } Q \text{ nonsingular,}$$

then $AX + X^\star B = 0$ can be transformed into

$$EY + Y^\star F = 0, \qquad \text{with } Y = QXP^{-\star}.$$

- If $E = E_1 \oplus \cdots \oplus E_d$, $F^\star = F_1^\star \oplus \cdots \oplus F_d^\star$, and $Y = [Y_{ij}]$ is partitioned into blocks accordingly, then this equation decouples in

$$E_i Y_{ii} + Y_{ii}^\star F_i = 0 \quad \text{and} \quad \left\{ \begin{array}{l} E_i Y_{ij} + Y_{ji}^\star F_j = 0 \\ E_j Y_{ji} + Y_{ij}^\star F_i = 0 \end{array} \right. , \quad (1 \leq i < j \leq d).$$

- Since KCF has 4 types of blocks, this produces **14** different types of matrix (systems) equations, whose explicit solutions have been found.
- Much more complicated general solution than standard Sylvester eq: $AX - XB = 0$, which depends on JCF of A and B and requires to solve only one type of equation.

Theorem

Let $G, H \in \mathbb{C}^{m\times n}$. Then $G - \lambda H$ is strictly equivalent to a direct sum of pencils of the following types

"Finite blocks": $J_k(\lambda_i - \lambda) := \begin{bmatrix} \lambda_i - \lambda & 1 & & \\ & \lambda_i - \lambda & 1 & \\ & & \ddots & \ddots \\ & & & \lambda_i - \lambda \end{bmatrix}$ *are $k \times k$.*

"Infinite blocks": $N_\ell = \begin{bmatrix} 1 & \lambda & & \\ & 1 & \lambda & \\ & & \ddots & \ddots \\ & & & 1 \end{bmatrix}$ *are $\ell \times \ell$.*

"Right singular blocks": $L_p := \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \end{bmatrix}$ *are $p \times (p+1)$.*

"Left singular blocks": *transposes of right singular blocks.*

Theorem

Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$. If the pencil $A - \lambda B^T$ has the KCF

$$
\begin{aligned}
E - \lambda F^T = L_{\varepsilon_1} \oplus L_{\varepsilon_2} \oplus \cdots \oplus L_{\varepsilon_a} \\
\oplus L_{\eta_1}^T \oplus L_{\eta_2}^T \oplus \cdots \oplus L_{\eta_b}^T \\
\oplus N_{u_1} \oplus N_{u_2} \oplus \cdots \oplus N_{u_c} \\
\oplus J_{k_1}(\lambda_1 - \lambda) \oplus J_{k_2}(\lambda_2 - \lambda) \oplus \cdots \oplus J_{k_d}(\lambda_d - \lambda)\,.
\end{aligned}
$$

Then the dimension of the solution space of the matrix equation

$$AX + X^T B = 0$$

depends only on $E - \lambda F^T$ and is

Breakdown of the dimension count for $AX + X^T B = 0$ (II)

Theorem

$$\text{dimension} = \sum_{i=1}^{a} \varepsilon_i + \sum_{\lambda_i=1} \lfloor k_i/2 \rfloor + \sum_{\lambda_j=-1} \lceil k_j/2 \rceil$$

$$+ \sum_{\substack{i,j=1\\ i<j}}^{a} (\varepsilon_i + \varepsilon_j) + \sum_{\substack{i<j\\ \lambda_i\lambda_j=1}} \min\{k_i, k_j\}$$

$$+ \sum_{\epsilon_i \le \eta_j} (\eta_j - \varepsilon_i + 1)$$

$$+ a \sum_{i=1}^{c} u_i + a \sum_{i=1}^{d} k_i + \sum_{\substack{i,j\\ \lambda_j=0}} \min\{u_i, k_j\}$$

- The method in this section can be applied when $A = B$ (orbits).
- Same results are obtained but expressed in different ways.
- What method is better?

Last important comment on solution of $AX + X^T B = 0$

- The method in this section can be applied when $A = B$ (orbits).
- Same results are obtained but expressed in different ways.
- What method is better?

Last important comment on solution of $AX + X^T B = 0$

- The method in this section can be applied when $A = B$ (orbits).
- Same results are obtained but expressed in different ways.
- What method is better?

Conclusions

- Many questions related to the Sylvester equation for $\star$-congruence $AX + X^{\star}B = C$ are nowadays well-understood.
- This equation appears in several applications and is related to "congruence problems".
- Connections with classical Sylvester equation $AX - XB = C$ but also relevant differences.
- Several problems still remain open. Among them, I consider the most relevant:
  - Eigenvalues of the operator $X \longmapsto AX + X^{T}B$.
  - Hasse diagram for inclusion of closures of congruence orbits.

Conclusions

- Many questions related to the Sylvester equation for $\star$-congruence $AX + X^{\star}B = C$ are nowadays well-understood.
- This equation appears in several applications and is related to "congruence problems".
- Connections with classical Sylvester equation $AX - XB = C$ but also relevant differences.
- Several problems still remain open. Among them, I consider the most relevant:
  - Eigenvalues of the operator $X \longmapsto AX + X^{T}B$.
  - Hasse diagram for inclusion of closures of congruence orbits.

Conclusions

- Many questions related to the Sylvester equation for $\star$-congruence $AX + X^{\star}B = C$ are nowadays well-understood.
- This equation appears in several applications and is related to "congruence problems".
- Connections with classical Sylvester equation $AX - XB = C$ but also relevant differences.
- Several problems still remain open. Among them, I consider the most relevant:
  - Eigenvalues of the operator $X \longmapsto AX + X^{T}B$.
  - Hasse diagram for inclusion of closures of congruence orbits.