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
The **matrix Sylvester equation**

\[
AX - XB = C, \quad A \in \mathbb{C}^{m \times m}, \quad B \in \mathbb{C}^{n \times n}
\]

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**

\[ AX - XB = C, \quad A \in \mathbb{C}^{m \times m}, \quad B \in \mathbb{C}^{n \times n} \]

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**

\[ AX - XB = C, \quad A \in \mathbb{C}^{m \times m}, \ B \in \mathbb{C}^{n \times n} \]

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**

\[ AX - XB = C, \quad A \in \mathbb{C}^{m \times m}, \ B \in \mathbb{C}^{n \times n} \]

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**

\[ AX - XB = C, \quad A \in \mathbb{C}^{m \times m}, \; B \in \mathbb{C}^{n \times n} \]

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**

\[ AX - XB = C, \quad A \in \mathbb{C}^{m \times m}, \quad B \in \mathbb{C}^{n \times n} \] 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.
Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

\[
AX + X^T B = C, \quad A \in \mathbb{C}^{m \times n}, \ B \in \mathbb{C}^{n \times m}
\]

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

\[
Gx = -\lambda G^T x, \quad G \in \mathbb{C}^{n \times n}.
\]

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);

The spectrum of palindromic eigenproblems has the symmetry \((\lambda, 1/\lambda)\).
Recently, the matrix Sylvester equation for congruence or T-Sylvester equation

\[ AX + X^T B = C, \quad A \in \mathbb{C}^{m \times n}, \ B \in \mathbb{C}^{n \times m} \]

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

\[ Gx = -\lambda G^T x, \quad G \in \mathbb{C}^{n \times n}. \]

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);

The spectrum of palindromic eigenproblems has the symmetry \((\lambda, 1/\lambda)\).
Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

\[ AX + X^T B = C, \quad A \in \mathbb{C}^{m \times n}, \quad B \in \mathbb{C}^{n \times m} \]

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

\[ Gx = -\lambda G^T x, \quad G \in \mathbb{C}^{n \times n}. \]

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);

The spectrum of palindromic eigenproblems has the symmetry \((\lambda, 1/\lambda)\).
Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

\[
AX + X^T B = C, \quad A \in \mathbb{C}^{m \times n}, \ B \in \mathbb{C}^{n \times m}
\]

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

\[
Gx = -\lambda G^T x, \quad G \in \mathbb{C}^{n \times n}.
\]

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);

The spectrum of palindromic eigenproblems has the symmetry \((\lambda, 1/\lambda)\)
Recently, the **matrix Sylvester equation for congruence** or T-Sylvester equation

\[ AX + X^T B = C, \quad A \in \mathbb{C}^{m \times n}, \quad B \in \mathbb{C}^{n \times m} \]

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

\[ Gx = -\lambda G^T x, \quad G \in \mathbb{C}^{n \times n}. \]

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);

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


This talk is my “personal journey” through Sylvester eq. for congruence.
Abstract (III)

\[ AX - XB = C \quad \text{vs.} \quad AX + X^T B = C \]

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


This talk is my “personal journey” trough Sylvester eq. for congruence.
Abstract (III)

\[ AX - XB = C \quad \text{vs.} \quad AX + X^T B = C \]

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


This talk is my “personal journey” through Sylvester eq. for congruence.
Abstract (III)

\[ AX - XB = C \quad \text{vs.} \quad AX + X^T B = C \]

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


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


Both interesting, both related, but DIFFERENT!!!
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


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


but sometimes both cases simultaneously: $* = T$ 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 $AX + X^*B = C$
   - Motivation
   - Consistency of the Sylvester equation for $*$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions

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

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

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

5 Conclusions
An important particular case of $AX + X^*B = C$

$AX + X^*A^* = C$ \hspace{1cm} ($* = T$ or $*$)

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

An important particular case of \( AX + X^* B = C \)

\[
AX + X^* A^* = C \quad (\star = T \text{ or } \star)
\]

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

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

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

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

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

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

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

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

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

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

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

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

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

An important particular case of \( AX + X^*B = C \)

\[
AX + X^*A^* = C \quad (\star = T \text{ or } *)
\]

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

Another important particular case of \( AX + X^* B = C \)

\[
AX + X^* A = 0, \quad A \in \mathbb{C}^{n \times n} \quad (\ast = T \text{ or } \ast)
\]

  - 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 \( AX + X^*B = C \)

- **Kressner & Schröder & Watkins, Numer. Algor., (2009)**: Same for \( \star = \ast \).
- **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. \( AXB + CX^*D = E \).
References for general equation $AX + X^*B = C$

- **Kressner & Schröder & Watkins, Numer. Algor., (2009):** Same for $\star = \ast$.
- **Chiang & Chu & Lin, Appl. Math. Comp., (2012):** Repeat some of the results above in other language and considers more general eqs. $AXB + CX^*D = E$.
References for general equation \( AX + X^*B = C \)


- **Kressner & Schröder & Watkins, Numer. Algor., (2009):** Same for \( \star = \ast \).

- **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. \( AXB + CX^*D = E \).


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


- **Kressner & Schröder & Watkins, Numer. Algor., (2009):** Same for $\star = \ast$.

- **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. $AXB + CX^*D = E$.


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


- **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. $AXB + CX^*D = E$.


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

- **Kressner & Schröder & Watkins, Numer. Algor., (2009):** Same for $\star = \ast$.
- **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. $AXB + CX^*D = E$.
References for general equation $AX + X^*B = C$

- **Kressner & Schröder & Watkins, Numer. Algor., (2009):** Same for $\star = \ast$.
- **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. $AXB + CX^*D = E$.
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 \( AX + X*B = C \)
   - Motivation
   - Consistency of the Sylvester equation for \(*\)-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions

4. General solution of \( AX + X*B = 0 \)

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

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

5 Conclusions
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)).
- \( Ax = \lambda B x \) (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.
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.
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.
Congruence, equivalence, and similarity. Orbits

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

$O(A) = \{ PAP^T : P \text{ nonsingular} \}$ \hspace{1cm} Congruence orbit of $A$

$O_s(A) = \{ PAP^{-1} : P \text{ nonsingular} \}$ \hspace{1cm} Similarity orbit of $A$

$O_e(A - \lambda B) = \{ P(A - \lambda B)Q : P, Q \text{ nonsing.} \}$ \hspace{1cm} Equivalency orbit of $A - \lambda B$

Similarity/equivalency orbits


- 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


- 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.
Congruence, equivalence, and similarity. Orbits

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

$\mathcal{O}(A) = \{ PAP^T : P \text{ nonsingular} \}$ \hspace{1cm} Congruence orbit of $A$

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

$\mathcal{O}_e(A - \lambda B) = \{ P(A - \lambda B)Q : P, Q \text{ nonsing.} \}$ \hspace{1cm} Equivalency orbit of $A - \lambda B$

Similarity/equivalency orbits


- 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**

- 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.
Congruence, equivalence, and similarity. Orbits

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**


- 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 \( A, B \in \mathbb{C}^{n \times n} \)

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

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

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

**Similarity/equivalency orbits**


- 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.

---

**F. M. Dopico** (U. Carlos III, Madrid)  
Sylvester equation for congruence  
Edinburgh, 2013
Given \( A, B \in \mathbb{C}^{n \times n} \)

\[
O(A) = \{ PAP^T : P \text{ nonsingular} \} \quad \text{Congruence orbit of } A
\]

\[
O_s(A) = \{ PAP^{-1} : P \text{ nonsingular} \} \quad \text{Similarity orbit of } A
\]

\[
O_e(A - \lambda B) = \{ P(A - \lambda B)Q : P, Q \text{ nonsing.} \} \quad \text{Equivalency orbit of } A - \lambda B
\]

**Similarity/equivalency orbits**


- 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.**
Codimension of the tangent space

\[ T_{\mathcal{O}(A)}(A) = \{XA + AX^T : X \in \mathbb{C}^{n \times n}\} \]

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

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

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!)
Codimension of the tangent space

\[ T_{O(A)}(A) = \{ XA + AX^T : X \in \mathbb{C}^{n \times n} \} \]
\[ T_{O_s(A)}(A) = \{ XA - AX : X \in \mathbb{C}^{n \times n} \} \]

Then:

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

(b) \( \text{codim } O_s(A) = \text{codim } T_{O_s(A)}(A) = \dim(\text{solution space of } XA - AX = 0) \)

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!)
Codimension of the tangent space

\[ T_{O(A)}(A) = \{ XA + AX^T : X \in \mathbb{C}^{n \times n} \} \quad \text{Tangent space of } O(A) \text{ at } A \]
\[ T_{O_s(A)}(A) = \{ XA - AX : X \in \mathbb{C}^{n \times n} \} \quad \text{Tangent space of } O_s(A) \text{ at } A \]

Then:

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

(b) \( \text{codim } O_s(A) = \text{codim } T_{O_s(A)}(A) = \text{dim}(\text{solution space of } XA - AX = 0) \)

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!)
Codimension of the tangent space

\[ T_{\mathcal{O}(A)}(A) = \{ XA + AX^T : X \in \mathbb{C}^{n \times n} \} \]

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

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

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!)
Codimension of the tangent space

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

\[ T_{O_s(A)}(A) = \{ XA - AX : X \in \mathbb{C}^{n \times n} \} \]
Tangent space of \( O_s(A) \) at \( A \)

Then:

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

(b) \( \text{codim } O_s(A) = \text{codim } T_{O_s(A)}(A) = \text{dim}(\text{solution space of } XA - AX = 0) \)

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!)
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 $AX +X^*B = C$
   - Motivation
   - Consistency of the Sylvester equation for $\ast$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions

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

5 Conclusions
Getting a simpler equation via congruence

Notation: \( S_A = \{ X \in \mathbb{C}^{n \times n} : AX^T +XA = 0 \} \) (solution space)

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

\[
B (PXP^{-1})^T + (PXP^{-1}) B = 0
\]

and \( S_A = P^{-1} S_B P \).

In particular: \( \dim S_A = \dim S_B \)

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: \( X = P^{-1}YP \).
Getting a simpler equation via congruence

Notation: \( S_A = \{ X \in \mathbb{C}^{n \times n} : AX^T +XA = 0 \} \) (solution space)

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

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

and \( S_A = P^{-1}S_BP \).

In particular: \( \dim S_A = \dim S_B \)

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: \( X = P^{-1}YP \).
Getting a simpler equation via congruence

Notation: \( S_A = \{ X \in \mathbb{C}^{n \times n} : AX^T + XA = 0 \} \) (solution space)

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

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

and \( S_A = P^{-1}S_B P \).

In particular: \( \dim S_A = \dim S_B \)

Procedure to solve \( AX^T + XA = 0 \):

1. Set \( C_A = PAP^T \), the canonical form of \( A \) for congruence !?.
2. Solve \( C_AY^T + YC_A = 0 \).
3. Undo the change: \( X = P^{-1}YP \).
Getting a simpler equation via congruence

Notation: \( S_A = \{ X \in \mathbb{C}^{n \times n} : AX^T +XA = 0 \} \) (solution space)

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

\[
B (PXP^{-1})^T + (PXP^{-1}) B = 0
\]

and \( S_A = P^{-1}S_B P \).

In particular: \( \text{dim} \ S_A = \text{dim} \ S_B \)

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: \( X = P^{-1}YP \).
Getting a simpler equation via congruence

Notation: \( S_A = \{ X \in \mathbb{C}^{n \times n} : AX^T +XA = 0 \} \) (solution space)

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

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

and \( S_A = P^{-1}S_B P \).

In particular: \( \text{dim} \ S_A = \text{dim} \ S_B \)

**Procedure to solve** \( AX^T +XA = 0 \):

1. Set \( C_A = PAP^T \), the canonical form of \( A \) for congruence !?.
2. Solve \( C_AY^T +YC_A = 0 \).
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 $AX +X*B = C$
   - Motivation
   - Consistency of the Sylvester equation for $*$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions

4 General solution of $AX +X*B = 0$

5 Conclusions
The canonical form for congruence

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.

\((\text{Type 0}) \quad J_k(0) = \begin{bmatrix} 0 & 1 \\ \vdots & \ddots & \ddots \\ 0 & 1 \\ 0 & \end{bmatrix}_{k \times k}^{k+1}

\((\text{Type I}) \quad \Gamma_k = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & -1 \\ \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 \\ -1 & -1 \\ 1 & 1 \\ 0 & \end{bmatrix}_{k \times k}, \quad \Gamma_1 = [1]

\((\text{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 \( \neq 2 \).


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


- **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)

  For complex matrices. **Six types of blocks.**

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


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


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


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


---

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.**


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


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.

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

  For complex matrices. **Six types of blocks.**

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


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


Corbas and Williams, J. Pure Appl. Algebra (2001), canonical forms over algebraically closed fields with characteristic not 2.


Corbas and Williams, J. Pure Appl. Algebra (2001), canonical forms over algebraically closed fields with characteristic not 2.


Corbas and Williams, J. Pure Appl. Algebra (2001), canonical forms over algebraically closed fields with characteristic not 2.


Proofs: first based on quivers and second constructive and based only on basic Matrix Analysis.

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


- **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.

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

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

5. Conclusions
Remember...

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**: $X = P^{-1}YP$.  

F. M. Dopico (U. Carlos III, Madrid) Sylvester equation for congruence Edinburgh, 2013 22 / 61
Partition into blocks to solve \( X C_A + C_A X^T = 0 \)  (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 \( X = \begin{bmatrix} X_{11} & \cdots & X_{1s} \\ \vdots & \ddots & \vdots \\ X_{s1} & \cdots & X_{ss} \end{bmatrix} \) conformally with \( C_A \).

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

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

Then:

\[
\dim S_A = \text{codim } \mathcal{O}(A) = \sum_i \text{codim } D_i + \sum_{i \neq j} \text{inter } (D_i, D_j)
\]
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 \( H_{2k}(\mu) \) \( (\text{Canonical form of } A) \)

Partition \( X = \begin{bmatrix} X_{11} & \ldots & X_{1s} \\ \vdots & \ddots & \vdots \\ X_{s1} & \ldots & X_{ss} \end{bmatrix} \) conformally with \( C_A \).

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

- \( i = j \): \( X_{ii} D_i + D_i X_{ii}^T = 0 \) \( \rightarrow \) \( \text{codim } D_i \) \( (\text{codimension}) \)

- \( i \neq j \): \( (i, j) \) \( X_{ij} D_j + D_i X_{ji}^T = 0 \) \( \rightarrow \) \( \text{inter } (D_i, D_j) \) \( (\text{interaction}) \)

Then:

\[ \dim S_A = \text{codim } \mathcal{O}(A) = \sum_i \text{codim } D_i + \sum_{i \neq j} \text{inter } (D_i, D_j) \]
Partition into blocks to solve $XC_A + C_A X^T = 0$ \hspace{1cm} (I)

Set $C_A = D_1 \oplus \cdots \oplus D_s$, $D_i = J_k(0)$, $\Gamma_k$, or $H_{2k}(\mu)$ \hspace{1cm} (Canonical form of $A$)

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

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

- $i = j$: $X_{ii}D_i + D_i X_{ii}^T = 0 \quad \rightarrow \text{codim } D_i$ (codimension)
- $i \neq j$: $(i, j)$ $X_{ij}D_j + D_j X_{ji}^T = 0$ \hspace{1cm} $(j, i)$ $X_{ji}D_i + D_i X_{ij}^T = 0 \quad \rightarrow \text{inter } (D_i, D_j)$ (interaction)

Then:

$$\dim S_A = \text{codim } O(A) = \sum_i \text{codim } D_i + \sum_{i \neq j} \text{inter } (D_i, D_j)$$
Partition into blocks to solve $XC_A + C_A X^T = 0$ (II)

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), $\Gamma_k$ (type I), or $H_{2k}(\mu)$ (type II)
(3 different types of eqs.)

(b) $XD_1 + D_2 Y^T = 0$
$YD_2 + D_1 X^T = 0$ (system of two matrix equations)

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


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

\[
C_A = J_{p_1}(0) \oplus J_{p_2}(0) \oplus \cdots \oplus J_{p_a}(0) \\
\oplus \Gamma_{q_1} \oplus \Gamma_{q_2} \oplus \cdots \oplus \Gamma_{q_b} \\
\oplus H_{2r_1}(\mu_1) \oplus H_{2r_2}(\mu_2) \oplus \cdots \oplus H_{2r_c}(\mu_c).
\]

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**

\[
c_{\text{Total}} = c_0 + c_1 + c_2 + i_{00} + i_{11} + i_{22} + i_{01} + i_{02} + i_{12}.
\]
Codimensions and interactions of canonical blocks

**Codimension**

\[
c_0 \rightarrow \left\lceil \frac{k}{2} \right\rceil \quad c_1 \rightarrow \left\lceil \frac{k}{2} \right\rceil \quad c_2 \rightarrow \begin{cases} k, & \text{if } \mu \neq (-1)^k \\ k + 2 \left\lceil \frac{k}{2} \right\rceil, & \text{if } \mu = (-1)^k \end{cases}
\]

**Interaction (same type)**

\[
i_{00} \rightarrow \begin{cases} \ell, & \ell \text{ even} \\ k, & \ell \text{ odd and } k \neq \ell \\ k + 1, & \ell \text{ odd and } k = \ell \end{cases}
\]

\[
i_{11} \rightarrow \begin{cases} 0, & k, \ell \text{ different parity} \\ \min\{k, \ell\}, & k, \ell \text{ same parity} \\ 4 \min\{k, \ell\}, & \mu = \tilde{\mu} = \pm 1 \\ 2 \min\{k, \ell\}, & \mu = \tilde{\mu} \neq \pm 1 \\ 2 \min\{k, \ell\}, & \mu \neq \tilde{\mu}, \mu\tilde{\mu} = 1 \\ 0, & \mu \neq \tilde{\mu}, \mu\tilde{\mu} \neq 1 \end{cases}
\]

**Interaction (different type)**

\[
i_{01} \rightarrow \begin{cases} 0, & k \text{ even} \\ \ell, & k \text{ odd} \end{cases}
\]

\[
i_{02} \rightarrow \begin{cases} 0, & k \text{ even} \\ 2\ell, & k \text{ odd} \end{cases}
\]

\[
i_{12} \rightarrow \begin{cases} 2 \min\{k, \ell\}, & \mu = (-1)^{k+1} \\ 0, & \mu \neq (-1)^{k+1} \end{cases}
\]

- 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:

\[
XH_{2k}((-1)^k) + H_{2k}((-1)^k)X^T = 0.
\]

### Codimensions and interactions of canonical blocks

**Codimension**

<table>
<thead>
<tr>
<th>$c_0$</th>
<th>$\left\lceil \frac{k}{2} \right\rceil$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>$\left\lfloor \frac{k}{2} \right\rfloor$</td>
</tr>
</tbody>
</table>
| $c_2$ | $\begin{cases} 
k, & \text{if } \mu \neq (-1)^k \\
 k + 2 \left\lceil \frac{k}{2} \right\rceil, & \text{if } \mu = (-1)^k 
\end{cases}$ |

**Interaction (same type)**

| $i_{00}$ | $\begin{cases} 
\ell, & \ell \text{ even} \\
 k, & \ell \text{ odd and } k \neq \ell \\
 k + 1, & \ell \text{ odd and } k = \ell 
\end{cases}$ |
| $i_{11}$ | $\begin{cases} 
0, & k, \ell \text{ different parity} \\
 4 \min\{k, \ell\}, & k, \ell \text{ same parity} \\
 2 \min\{k, \ell\}, & \mu = \tilde{\mu} = \pm 1 \\
 2 \min\{k, \ell\}, & \mu = \tilde{\mu} \neq \pm 1 \\
 0, & \mu \neq \tilde{\mu}, \mu \tilde{\mu} = 1 
\end{cases}$ |

**Interaction (different type)**

| $i_{01}$ | $\begin{cases} 
0, & k \text{ even} \\
 \ell, & k \text{ odd} 
\end{cases}$ |
| $i_{02}$ | $\begin{cases} 
0, & k \text{ even} \\
 2\ell, & k \text{ odd} 
\end{cases}$ |
| $i_{12}$ | $\begin{cases} 
2 \min\{k, \ell\}, & \mu = (-1)^{k+1} \\
 0, & \mu \neq (-1)^{k+1} 
\end{cases}$ |

---

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

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

5. Conclusions
Generic canonical structure for congruence (I)

Generic = codimension zero


*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 canonical structure for congruence (I)

**Generic** = codimension zero

---


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

\[
J_{\lambda_i} := J_{n_i,1}(\lambda_i) \oplus \cdots \oplus J_{n_i,q_i}(\lambda_i), \quad \text{for } i = 1, \ldots, d \text{ and } \lambda_i \neq \lambda_j \text{ if } i \neq j,
\]

the similarity bundle of \( A \) is

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

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

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

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


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 $\mu_1, \mu_2, \ldots$).
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 $AX + X^*B = C$
   - Motivation
   - Consistency of the Sylvester equation for $*$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions

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

5. Conclusions
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

$$AX + X^* B = C,$$

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


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

$$AX - XB = C,$$  $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$, $C \in \mathbb{C}^{m \times n}$.
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

$$AX + X^*B = C,$$

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)).


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

$$AX - XB = C,$$

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

$$AX + X^* B = C,$$

$(X^* = X^T$ 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)).


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**

$$AX - XB = C,$$

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

$A = Q_A C_A Q_A^T$ and $B = Q_B C_B Q_B^T$.

$$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$

$$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$

But,

$$(Q_A^T X Q_B^{-T})^T \neq Q_A^{-1} X^T Q_B$$

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 $B = Q_B C_B Q_B^T$.

$$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$

$$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$

$$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$

- But,

$$\left(Q_A^T X Q_B^{-T}\right)^T \neq Q_A^{-1} X^T Q_B$$

- 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 $B = Q_B C_B Q_B^T$.

$$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$

$$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$

$$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$

But,

$$\left( Q_A^T X Q_B^{-T} \right)^T \neq Q_A^{-1} X^T Q_B$$

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 $B = Q_B C_B Q_B^T$.

$$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$

$$C_A Q_A^T XQ_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$

But,

$$(Q_A^T X Q_B^{-T})^T \neq Q_A^{-1} X^T Q_B$$

with equality only if $Q_A = Q_B$. 

F. M. Dopico (U. Carlos III, Madrid)  Sylvester equation for congruence  Edinburgh, 2013  34 / 61
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 $B = Q_B C_B Q_B^T$.

$$Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C$$

$$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$

$$C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}$$

- But,

$$\left(Q_A^T X Q_B^{-T}\right)^T \neq Q_A^{-1} X^T Q_B$$

- 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 \( B = Q_B C_B Q_B^T \).
- 
  \[ Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C \]
- 
  \[ C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T} \]
- 
  \[ C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T} \]
- But, 
  
  \[ (Q_A^T X Q_B^{-T})^T \neq Q_A^{-1} X^T Q_B \]
  
  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 \( B = Q_B C_B Q_B^T \).

\[
Q_A C_A Q_A^T X + X^T Q_B C_B Q_B^T = C
\]

\[
C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}
\]

\[
C_A Q_A^T X Q_B^{-T} + Q_A^{-1} X^T Q_B C_B = Q_A^{-1} C Q_B^{-T}
\]

But,

\[
(Q_A^T X Q_B^{-T})^T \neq Q_A^{-1} X^T Q_B
\]

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.

\[
Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C
\]

\[
J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B
\]

“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
\]

- $J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$
- $J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$

“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.

$$Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C$$

$$J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$$

“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.

\[
Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C
\]

\[
J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B
\]

\[
J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B
\]

“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.

\[
Q_A J_A Q_A^{-1} X - X Q_B J_B Q_B^{-1} = C
\]

- $J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$

- $J_A Q_A^{-1} X Q_B - Q_A^{-1} X Q_B J_B = Q_A^{-1} C Q_B$

“Canonical forms” to be used:

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

$$PRQX + X^T Q^T S^T P^T = C$$

$$R QXP^{-T} + P^{-1}X^T Q^T S^T = P^{-1}CP^{-T}$$

$$R QXP^{-T} + (QXP^{-T})^T S^T = P^{-1}CP^{-T}$$

"Canonical forms" for pencils to be used:

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

\[ PRQX + X^T Q^T S^T P^T = C \]

\[ R Q X P^{-T} + P^{-1} X^T Q^T S^T = P^{-1} C P^{-T} \]

\[ R \begin{pmatrix} Q X P^{-T} \\ (Q X P^{-T})^T \end{pmatrix} + S^T = P^{-1} C P^{-T} \]

“Canonical forms” for pencils to be used:

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

$$PRQX + X^T Q^T S^T P^T = C$$

$$RQP^{-T} + P^{-1}X^T Q^T S^T = P^{-1}CP^{-T}$$

$$RQP^{-T} + (QXP^{-T})^T S^T = P^{-1}CP^{-T}$$

“Canonical forms” for pencils to be used:

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

$$PRQX + X^T Q^T S^T P^T = C$$

$$RQP^{-T} + P^{-1}X^T Q^T S^T = P^{-1}CP^{-T}$$

“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.

\[
PRQX + X^T Q^T S^T P^T = C
\]

\[
RQP^{-T} + P^{-1}X^T Q^T S^T = P^{-1}CP^{-T}
\]

\[
RQP^{-T} + (QXP^{-T})^T S^T = P^{-1}CP^{-T}
\]

“Canonical forms” for pencils to be used:

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

$$PRQX + X^T Q^T S^T P^T = C$$

$$RQXP^{-T} + P^{-1}X^T Q^T S^T = P^{-1}CP^{-T}$$

$$RQXP^{-T} + (QXP^{-T})^T S^T = P^{-1}CP^{-T}$$

“Canonical forms” for pencils to be used:

1. For theory: KCF (Kronecker Canonical Form).
2. For computations: Generalized Schur form.
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 $AX +X^*B = C$
   - Motivation
   - Consistency of the Sylvester equation for $\star$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions
4. General solution of $AX +X^*B = 0$
5. Conclusions
Motivation for studying $AX + X^* B = C$ (I)

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

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

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

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

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 $AX + X^* B = C$ (I)

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

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

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

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

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 $AX + X^*B = C$ (I)

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

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

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

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

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 $AX + X^* B = C$ (II)

Structured numerical algorithms for linear palindromic eigenproblems $(Z + \lambda Z^*)$ 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

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

$M$ can be computed via structure-preserving methods (Kressner, Schröder, Watkins (Numer. Alg., 2009)) or (Mackey$^2$, Mehl, Mehrmann (NLAA, 2009)) and compute eigenvalues of $Z + \lambda Z^*$ with exact pairing $\lambda, 1/\lambda^*$. 
Motivation for studying $AX + X^* B = C$ (II)

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

**Theorem (Kressner, Schröder, Watkins (Numer. Alg., 2009) and Mackey\textsuperscript{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

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

$M$ can be computed via structure-preserving methods (Kressner, Schröder, Watkins (Numer. Alg., 2009)) or (Mackey\textsuperscript{2}, Mehl, Mehrmann (NLAA, 2009)) and compute eigenvalues of $Z + \lambda Z^*$ with exact pairing $\lambda, 1/\lambda^*$. 
Motivation for studying $AX + X^*B = C$ (II)

Structured numerical algorithms for linear palindromic eigenproblems $(Z + \lambda Z^*)$ 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

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

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

F. M. Dopico (U. Carlos III, Madrid)
Motivation for studying $AX + X^* B = C$ (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

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

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

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

and to compute deflating subspaces of palindromic pencils.
Motivation for studying $AX + X^* B = C$ (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

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

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

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

and to compute **deflating subspaces of palindromic pencils**.
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 $AX + X*B = C$
   - Motivation
   - Consistency of the Sylvester equation for $*$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions

4 General solution of $AX + X*B = 0$

5 Conclusions
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

$$AX + X^* B = C$$

if and only if

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

are $\star$-congruent.

Remarks:

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

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

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

$$AX + X^* B = C$$

if and only if

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

are $*$-congruent.

**Remarks:**

- The implication $\Longrightarrow$ very easy: done in previous slide.
- The implication $\Longleftarrow$ more challenging.
- Wimmer proved in 1994 the result, for $F = \mathbb{C}$ and $* = *$, without any reference to palindromic eigenproblems.
- His motivation was the study of standard Sylvester equations with Hermitian solutions.
Let $F$ be any field and let $A \in F^{m \times m}$, $B \in F^{n \times n}$, $C \in F^{m \times n}$ be given. There is some $X \in F^{m \times n}$ such that

$$AX - XB = C$$

if and only if

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

are similar.
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
Uniqueness of solutions of $AX + X^* B = C$ (I)

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

$$\begin{align*}
\mathbb{F}^{n \times m} & \rightarrow \mathbb{F}^{m \times n} \\
X & \mapsto AX + X^* B
\end{align*}$$

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 $m = n$. 

F. M. Dopico (U. Carlos III, Madrid)  Sylvester equation for congruence  Edinburgh, 2013  45 / 61
Uniqueness of solutions of $AX + X^* B = C$ (I)

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

$$
\mathbb{F}^{n \times m} \rightarrow \mathbb{F}^{m \times m}
X \mapsto AX + X^* B
$$

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 $AX + X^* B = C$ (I)

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

\[
\mathbb{F}^{n \times m} \longrightarrow \mathbb{F}^{m \times m}
\]

\[
X \longmapsto AX + X^* B
\]

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 \( AX + X^* B = C \) \((I)\)

**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**

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

**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 \( AX + X^* B = C \) (II)

**Definition:** a set \( \{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C} \) is \( \star \)-reciprocal free if \( \lambda_i \neq 1/\lambda_j^* \) for any \( 1 \leq i, j \leq n \). We admit 0 and/or \( \infty \) 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 \( \star \)-reciprocal free.
Uniqueness of solutions of $AX + X^* B = C$ (II)

**Definition:** a set $\{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C}$ is $\star$-reciprocal free if $\lambda_i \neq 1/\lambda_j^*$ for any $1 \leq i, j \leq n$. We admit 0 and/or $\infty$ 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 $\star$-reciprocal free.
Uniqueness of solutions of \( A X + X^* B = C \) (II)

**Definition:** a set \( \{\lambda_1, \ldots, \lambda_n\} \subset \mathbb{C} \) is \((\ast)-reciprocal\ free\) if \( \lambda_i \neq 1/\lambda_j^* \) for any \( 1 \leq i, j \leq n \). We admit 0 and/or \( \infty \) 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:

1. \( AX + X^T B = C \) 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.

2. \( AX + X^* B = C \) 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 \((\ast)\)-reciprocal free.
...to be compared with uniqueness conditions for standard Sylvester eq

**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 $AX + X^*B = C$ and $AX - XB = C$:

- In $AX + X^*B = C$, one starts by dealing with the eigenproblem of $A - \lambda B^*$, 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 \( AX + XB = C \) and \( AX - XB = C \):

- In \( AX + XB = C \), one starts by dealing with the eigenproblem of \( A - \lambda B \), 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 \( AX + X^*B = C \) and \( AX - XB = C \):

- In \( AX + X^*B = C \), one starts by dealing with the eigenproblem of \( A - \lambda B^* \), 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 $AX + X^*B = C$ and $AX - XB = C$:

- In $AX + X^*B = C$, one starts by dealing with the eigenproblem of $A - \lambda B^*$, 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 $B$. 
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 $AX + X^*B = C$
   - Motivation
   - Consistency of the Sylvester equation for $*$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions
4. General solution of $AX + X^*B = 0$
5. Conclusions
In this section in $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

$AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\Re X, \Im X)$ if $\star = \ast$. 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 $AX + X^* B = C$ 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

$$A - \lambda B^* = U(R - \lambda S)V,$$

where

$$\begin{cases} R, S \\
U, V \end{cases}$$

are upper triangular

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 $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

- $AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\text{Re} X, \text{Im} X)$ if $\star = \ast$. 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 $AX + X^* B = C$ 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

$$A - \lambda B^* = U(R - \lambda S)V,$$

where $\{ R, S \}$ are upper triangular

$$U, V$$

are unitary matrices

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 $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

- $AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\text{Re} X, \text{Im} X)$ if $\star = \ast$. 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 $AX + X^* B = C$ 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

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

- 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 $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

$AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\mathcal{R}e\, X, \mathcal{I}m\, X)$ if $\star = \ast$. 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 $AX + X^* B = C$ 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

$$A - \lambda B^* = U(R - \lambda S)V,$$

where $\{R, S\}$ are upper triangular and $U, V$ are unitary matrices.

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 $AX + X^* B = C$ all matrices are in $\mathbb{C}^{n \times n}$ and the solution is unique for every $C$.

- $AX + X^* B = C$ is equivalent to a linear system for the $n^2$ entries of $X$ if $\star = T$ and to a linear system for the $2n^2$ entries of $(\text{Re} X, \text{Im} X)$ if $\star = \ast$. 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 $AX + X^* B = C$ 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

$$A - \lambda B^* = U(R - \lambda S)V,$$

where $\{ R, S \}$ are upper triangular and $U, V$ are unitary matrices.

If $A, B$ real matrices: use real arithmetic to get **quasi-triangular** $R$. We do not consider this for brevity.
Algorithm to solve $AX + X^*B = C$ in $O(n^3)$ flops

**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

$$A = URV, \quad B^* = USV,$$

where $\left\{ \begin{array}{l} R, S \text{ are upper triangular} \\ U, V \text{ are unitary matrices} \end{array} \right.$$

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

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

$$RW + W^* S^* = E$$

**Step 4.** Compute $X = V^* W U^*$
Algorithm to solve $AX + X^* B = C$ in $O(n^3)$ flops

**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

$$A = URV, \quad B^* = USV,$$

where $\{R, S\}$ are upper triangular and $U, V$ are unitary matrices.

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

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

$$RW + W^* S^* = E?$$

**Step 4.** Compute $X = V^* W U^*$
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

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

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

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

a scalar equation that allows us to determine $w_{44}$.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ 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}^* & w_{12}^* & w_{31}^* & w_{41}^* \\
    w_{12}^* & w_{22}^* & w_{32}^* & w_{42}^* \\
    w_{13}^* & w_{23}^* & w_{33}^* & w_{43}^* \\
    w_{14}^* & w_{24}^* & w_{34}^* & w_{44}^*
\end{bmatrix}
\begin{bmatrix}
    s_{11}^* & 0 & 0 & 0 \\
    s_{12}^* & s_{22}^* & 0 & 0 \\
    s_{13}^* & s_{23}^* & s_{33}^* & 0 \\
    s_{14}^* & s_{24}^* & s_{34}^* & s_{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

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

a scalar equation that allows us to determine $w_{44}$. 
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

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

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

\[
s_{33} w_{34} + w_{43}^* r_{44}^* = e_{43}^* - s_{34} w_{44}^*
\]

\[
r_{33} w_{34} + w_{43}^* s_{44}^* = e_{34}^* - r_{34} w_{44}^*
\]

a $2 \times 2$ system of scalar equations that allows us to determine $w_{34}$ and $w_{43}$ simultaneously.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

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

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

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

a $2 \times 2$ system of scalar equations that allows us to determine $w_{34}$ and $w_{43}$ simultaneously.
We illustrate with $4 \times 4$ example for simplicity:

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

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

\[
\begin{align*}
  s_{22} w_{24} + w_{42}^* r_{44}^* &= e_{42}^* - s_{23}^* w_{34}^* - s_{24}^* w_{44}^* \\
  r_{22} w_{24} + w_{42}^* s_{44}^* &= e_{24}^* - r_{23}^* w_{34}^* - r_{24}^* w_{44}^*
\end{align*}
\]

a $2 \times 2$ system of scalar equations that allows us to determine $w_{24}$ and $w_{42}$ simultaneously.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

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

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

$$
s_{22} w_{24} + w_{42}^* r_{44}^* = e_{42}^* - s_{23},
$$

$$
r_{22} w_{24} + w_{42}^* s_{44}^* = e_{24} - r_{23},
$$

a $2 \times 2$ system of scalar equations that allows us to determine $w_{24}$ and $w_{42}$ simultaneously.
Algorithm to solve the transformed equation \( RW + W^* S^* = E \) (I)

We illustrate with \( 4 \times 4 \) example for simplicity:

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

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

\[
\begin{align*}
  s_{11} & \quad w_{14} & + & \quad w_{14}^* & \quad r_{11}  \\
  r_{11} & \quad w_{14} & + & \quad w_{14}^* & = e_{41}^* - s_{12} & \quad w_{24} & - s_{13} & \quad w_{34} & - s_{14} & \quad w_{44}
\end{align*}
\]

a \( 2 \times 2 \) system of scalar equations that allows us to determine \( w_{14} \) and \( w_{41} \) simultaneously.
Algorithm to solve the transformed equation \( RW + W^* S^* = E \) (I)

We illustrate with \( 4 \times 4 \) example for simplicity:

\[
\begin{bmatrix}
    r_{11} & r_{12} & r_{13} & r_{14} \\
    0 & r_{22} & r_{23} & r_{24} \\
    0 & 0 & r_{33} & r_{34} \\
    0 & 0 & 0 & r_{44}
\end{bmatrix}
+ \begin{bmatrix}
    w^*_{11} & w^*_{12} & w^*_{13} & w^*_{14} \\
    w^*_{12} & w^*_{22} & w^*_{23} & w^*_{24} \\
    w^*_{13} & w^*_{23} & w^*_{33} & w^*_{34} \\
    w^*_{14} & w^*_{24} & w^*_{34} & w^*_{44}
\end{bmatrix}
= \begin{bmatrix}
    s^*_{11} & 0 & 0 & 0 \\
    s^*_{12} & s^*_{22} & 0 & 0 \\
    s^*_{13} & s^*_{23} & s^*_{33} & 0 \\
    s^*_{14} & s^*_{24} & s^*_{34} & s^*_{44}
\end{bmatrix}
\]

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

\[
s_{11} \begin{bmatrix} w_{14} \\ r_{14} \end{bmatrix} + \begin{bmatrix} w^*_{14} \\ w^*_{41} \end{bmatrix} \begin{bmatrix} r^*_{44} \\ s^*_{44} \end{bmatrix} = \begin{bmatrix} e^*_{14} - s_{12} \\ e_{14} - r_{12} \end{bmatrix} \begin{bmatrix} w_{24} \\ w_{24} \end{bmatrix} - \begin{bmatrix} w_{24} \\ w_{24} \end{bmatrix} \begin{bmatrix} e_{13} \\ e_{13} \end{bmatrix} + \begin{bmatrix} e_{33} \\ e_{33} \end{bmatrix} \begin{bmatrix} w_{34} \\ w_{34} \end{bmatrix} - \begin{bmatrix} w_{34} \\ w_{34} \end{bmatrix} \begin{bmatrix} e_{14} \\ e_{14} \end{bmatrix},
\]

a \( 2 \times 2 \) system of scalar equations that allows us to determine \( w_{14} \) and \( w_{41} \) simultaneously.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

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

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

$$
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} \\
  0 & r_{22} & r_{23} \\
  0 & 0 & r_{33}
\end{bmatrix}
- \begin{bmatrix}
  r_{14} \\
  r_{24} \\
  r_{34}
\end{bmatrix}
= \begin{bmatrix}
  w_{41}^* & w_{42}^* & w_{43}^* \\
  w_{42}^* & w_{43}^* & w_{44}^*
\end{bmatrix}
$$

which is a $3 \times 3$ matrix equation of the same type as the original one.
Algorithm to solve the transformed equation $RW + W^* S^* = E$ (I)

We illustrate with $4 \times 4$ example for simplicity:

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

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

$$
\begin{bmatrix}
  r_{11} & r_{12} & r_{13} \\
  0 & r_{22} & r_{23} \\
  0 & 0 & r_{33}
\end{bmatrix}
- 
\begin{bmatrix}
  e_{11} & e_{12} & e_{13} \\
  e_{21} & e_{22} & e_{23} \\
  e_{31} & e_{32} & e_{33}
\end{bmatrix}
= 
\begin{bmatrix}
  w^*_{11} & w^*_{21} & w^*_{31} \\
  w^*_{12} & w^*_{22} & w^*_{32} \\
  w^*_{13} & w^*_{23} & w^*_{33}
\end{bmatrix}
- 
\begin{bmatrix}
  s^*_{11} & 0 & 0 \\
  s^*_{12} & s^*_{22} & 0 \\
  s^*_{13} & s^*_{23} & s^*_{33}
\end{bmatrix}
$$

which is a $3 \times 3$ matrix equation of the same type as the original one.
Remarks on algorithm to solve $AX + X^* B = C$

- **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 $AX + X^* B = C$.

- **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 $AX + X^*B = C$

- **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 $AX + X^*B = C$.

- **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 $AX + X^* B = C$

- **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 $AX + X^* B = C$.

- **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 $AX + X^* B = C$

- **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 $AX + X^* B = C$.

- **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 $AX + X^* B = C$

- **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 $AX + X^* B = C$.

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

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

5 Conclusions
Theoretical method to solve $AX + X^* B = 0$ (I)

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

The latter found by De Terán, D., Guillery, Montealegre, Reyes, Lin. Alg. Appl., 2013

- I do not know any clear application for this problem.
Theoretical method to solve $AX + X^*B = 0$ (I)

In case of consistency, but “nonuniqueness”, general solution of $AX + X^*B = C$ is $X = X_p + X_h$, where

1. $X_p$ is a particular solution and
2. $X_h$ is the general solution of $AX + X^*B = 0$.

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 
\[ AX + X^*B = C \] is \( X = X_p + X_h \), where

1. \( X_p \) is a particular solution and
2. \( X_h \) is the general solution of \( AX + X^*B = 0 \).

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.
Theoretical method to solve $AX + X^* B = 0$ (II)

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

$$A - \lambda B^* = P(E - \lambda F^*)Q,$$

with $P$ and $Q$ nonsingular,

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

$$EY + Y^* F = 0,$$

with $Y = QXP^{*-*}$.

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

$$E_iY_{ii} + Y_{ii}^* F_i = 0$$

and

$$\begin{cases} E_iY_{ij} + Y_{ji}^* F_j = 0 \\ E_jY_{ji} + Y_{ij}^* F_i = 0 \end{cases}, \quad (1 \leq i < j \leq d).$$

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.
Theoretical method to solve $AX + X^*B = 0$ (II)

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

  $$A - \lambda B^* = P(E - \lambda F^*)Q,$$

  with $P$ and $Q$ nonsingular,

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

  $$EY + Y^*F = 0,$$

  with $Y = QXP^{-*}$.

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

  $$E_iY_{ii} + Y_{ii}^*F_i = 0$$

  and

  $$\begin{cases}
  E_iY_{ij} + Y_{ji}^*F_j = 0 \\
  E_jY_{ji} + Y_{ij}^*F_i = 0
  \end{cases}, \quad (1 \leq i < j \leq d).$$

- 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.
Theoretical method to solve $AX + X^* B = 0$ (II)

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

$$A - \lambda B^* = P(E - \lambda F^*)Q,$$

with $P$ and $Q$ nonsingular,

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

$$EY + Y^* F = 0,$$

with $Y = QXP^{-*}$.

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

$$E_i Y_{ii} + Y_{ii}^* F_i = 0$$

and

$$\begin{cases} E_i Y_{ij} + Y_{ji}^* F_j = 0 \\ E_j Y_{ji} + Y_{ij}^* F_i = 0 \end{cases}, \quad (1 \leq i < j \leq d).$$

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.
Theoretical method to solve $AX + X^* B = 0$ (II)

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

$$A - \lambda B^* = P(E - \lambda F^*)Q,$$

with $P$ and $Q$ nonsingular,

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

$$EY + Y^* F = 0,$$

with $Y = QXP^{-*}$.

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

$$E_i Y_{ii} + Y_{ii}^* F_i = 0$$

and

$$E_i Y_{ij} + Y_{ij}^* F_j = 0$$

$$E_j Y_{ji} + Y_{ji}^* F_i = 0,$$

$(1 \leq i < j \leq d)$.

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.
The Kronecker Canonical Form of a Matrix Pencil

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

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

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

- **“Left singular blocks”:** transposes of right singular blocks.
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

$$E - \lambda F^T = L_{\epsilon_1} \oplus L_{\epsilon_2} \oplus \cdots \oplus L_{\epsilon_a}$$
$$\oplus L_{\eta_1}^T \oplus L_{\eta_2}^T \oplus \cdots \oplus L_{\eta_b}^T$$
$$\oplus N_{u_1} \oplus N_{u_2} \oplus \cdots \oplus N_{u_c}$$
$$\oplus J_{k_1}(\lambda_1 - \lambda) \oplus J_{k_2}(\lambda_2 - \lambda) \oplus \cdots \oplus J_{k_d}(\lambda_d - \lambda).$$

Then the dimension of the solution space of the matrix equation

$$AX + X^TB = 0$$

depends only on $E - \lambda F^T$ and is
Breakdown of the dimension count for $AX + X^T B = 0$ (II)

**Theorem**

\[
\text{dimension} = \sum_{i=1}^{a} \varepsilon_i + \sum_{\lambda_i = 1} \lfloor k_i / 2 \rfloor + \sum_{\lambda_j = -1} \lfloor k_j / 2 \rfloor \\
+ \sum_{\substack{i,j = 1 \atop i < j}}^{a} (\varepsilon_i + \varepsilon_j) + \sum_{\substack{i < j \atop \lambda_i \lambda_j = 1}} \min\{k_i, k_j\} \\
+ \sum_{\varepsilon_i \leq \eta_j} (\eta_j - \varepsilon_i + 1) \\
+ a \sum_{i=1}^{c} u_i + a \sum_{i=1}^{d} k_i + \sum_{\substack{i,j \atop \lambda_j = 0}} \min\{u_i, k_j\}\]
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?
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?
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?
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 $AX + X^*B = C$
   - Motivation
   - Consistency of the Sylvester equation for $\ast$-congruence
   - Uniqueness of solutions
   - Efficient and stable algorithm to compute unique solutions

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

5. 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 $X \mapsto AX + X^T B$.
- 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 $X \mapsto AX + X^T B$.
- Hasse diagram for inclusion of closures of congruence orbits.
Many questions related to the **Sylvester equation for ⋆-congruence**
\[ AX + X^* B = C \] 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 \( X \mapsto AX + X^T B \).
- Hasse diagram for inclusion of closures of congruence orbits.