The term **nonlinear eigenvalue problem** is not used in a unique way in the literature.
The term **nonlinear eigenvalue problem** is not used in a unique way in the literature.

On the one hand: Parameter dependent nonlinear (with respect to the state variable) operator equations

\[ T(\lambda, u) = 0 \]

discussing

- positivity of solutions
- multiplicity of solution
- dependence of solutions on the parameter; bifurcation
- (change of) stability of solutions
For $D \subset \mathbb{C}$ let $T(\lambda)$, $\lambda \in D$ be a family of linear operators on a Hilbert space $\mathcal{H}$ (more generally a family of closed linear operators on a Banach space). Find $\lambda \in D$ and $x \neq 0$ such that $T(\lambda)x = 0$. Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding (right) eigenvector.
For $D \subset \mathbb{C}$ let $T(\lambda)$, $\lambda \in D$ be a family of linear operators on a Hilbert space $\mathcal{H}$ (more generally a family of closed linear operators on a Banach space).

Find $\lambda \in D$ and $x \neq 0$ such that

$$T(\lambda)x = 0.$$ 

Then $\lambda$ is called an eigenvalue of $T(\cdot)$, and $x$ a corresponding (right) eigenvector.
Outline

1. Examples
2. Basic Properties
3. Rayleigh Functional
4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method
5. Invariant pairs
Outline

1. Examples
2. Basic Properties
3. Rayleigh Functional
4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method
5. Invariant pairs
Equations of motion arising in dynamic analysis of structures (with a finite number of degrees of freedom) are

\[ M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = f(t) \]

where \( q \) are the Lagrangean coordinates, \( M \) is the mass matrix, \( K \) the stiffness matrix, \( C \) the viscous damping matrix, and \( f \) an external force vector.
Example 1: Vibration of structures

Equations of motion arising in dynamic analysis of structures (with a finite number of degrees of freedom) are

\[ M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = f(t) \]

where \( q \) are the Lagrangean coordinates, \( M \) is the mass matrix, \( K \) the stiffness matrix, \( C \) the viscous damping matrix, and \( f \) an external force vector.

Suppose that the system is exited by a time harmonic force \( f(t) = f_0 e^{i\omega_0 t} \).
Example 1: Vibration of structures

Equations of motion arising in dynamic analysis of structures (with a finite number of degrees of freedom) are

$$M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = f(t)$$

where $q$ are the Lagrangean coordinates, $M$ is the mass matrix, $K$ the stiffness matrix, $C$ the viscous damping matrix, and $f$ an external force vector.

Suppose that the system is exited by a time harmonic force $f(t) = f_0 e^{i\omega_0 t}$.

If $C = \alpha K + \beta M$ (modal damping) and $(x_j, \mu_j), j = 1, \ldots, n$ denotes an orthonormal eigensystem of $Kx = \mu Mx$ (i.e. $x_k^T M x_j = \delta_{jk}$) then the periodic response of the system is

$$q(t) = e^{i\omega_0 t} \sum_{j=1}^{n} \frac{x_j^T f_0}{\mu_j - \omega_0^2 + i\omega_0 (\alpha \mu_j + \beta)} x_j.$$

Usually, one gets good approximations taking into account only a small number of eigenvalues in the vicinity of $i\omega_0$ (truncated mode superposition).
Example 2: Viscoelastic model of damping

Using a viscoelastic constitutive relation to describe the material behavior in the equations of motion yields a rational eigenvalue problem in the case of free vibrations.

\[ T(\omega)x = (\omega^2 M + K - K \sum_{j=1}^{1} b_j \omega \Delta K_j) x = 0 \]

where \( M \) is the consistent mass matrix, \( K \) is the stiffness matrix with the instantaneous elastic material parameters used in Hooke's law, and \( \Delta K_j \) collects the contributions of damping from elements with relaxation parameter \( b_j \).
Example 2: Viscoelastic model of damping

Using a viscoelastic constitutive relation to describe the material behavior in the equations of motion yields a rational eigenvalue problem in the case of free vibrations.

Discretizing by finite elements yields (cf. Hager & Wiberg 2000)

\[ T(\omega) x := \left( \omega^2 M + K - \sum_{j=1}^{K} \frac{1}{1 + b_j \omega} \Delta K_j \right) x = 0 \]

where \( M \) is the consistent mass matrix, \( K \) is the stiffness matrix with the instantaneous elastic material parameters used in Hooke’s law, and \( \Delta K_j \) collects the contributions of damping from elements with relaxation parameter \( b_j \).
Semiconductor nanostructures have attracted tremendous interest in the past few years because of their special physical properties and their potential for applications in micro- and optoelectronic devices.
Example 3: Electronic structure of quantum dots

Semiconductor nanostructures have attracted tremendous interest in the past few years because of their special physical properties and their potential for applications in micro- and optoelectronic devices.

In such nanostructures, the free carriers are confined to a small region of space by potential barriers, and if the size of this region is less than the electron wavelength, the electronic states become quantized at discrete energy levels.
Example 3: Electronic structure of quantum dots

Semiconductor nanostructures have attracted tremendous interest in the past few years because of their special physical properties and their potential for applications in micro- and optoelectronic devices.

In such nanostructures, the free carriers are confined to a small region of space by potential barriers, and if the size of this region is less than the electron wavelength, the electronic states become quantized at discrete energy levels.

The ultimate limit of low dimensional structures is the quantum dot, in which the carriers are confined in all three directions, thus reducing the degrees of freedom to zero.

Therefore, a quantum dot can be thought of as an artificial atom.
Electron Spectroscopy Group, Fritz-Haber-Institute, Berlin
Example 3 ct.

Governing equation: Schrödinger equation

\[-\nabla \cdot \left( \frac{\hbar^2}{2m(x, E)} \nabla \Phi \right) + V(x)\Phi = E\Phi, \quad x \in \Omega_q \cup \Omega_m\]

where $\hbar$ is the reduced Planck constant, $m(x, E)$ is the electron effective mass depending on the energy state $E$, and $V(x)$ is the confinement potential.
Example 3 ct.

Governing equation: Schrödinger equation

\[ -\nabla \cdot \left( \frac{\hbar^2}{2m(x, E)} \nabla \Phi \right) + V(x)\Phi = E\Phi, \quad x \in \Omega_q \cup \Omega_m \]

where \( \hbar \) is the reduced Planck constant, \( m(x, E) \) is the electron effective mass depending on the energy state \( E \), and \( V(x) \) is the confinement potential.

\( m \) and \( V \) are discontinous across the heterojunction.
Example 3 ct.

**Governing equation:** Schrödinger equation

\[-\nabla \cdot \left( \frac{\hbar^2}{2m(x,E)} \nabla \Phi \right) + V(x)\Phi = E\Phi, \quad x \in \Omega_q \cup \Omega_m\]

where \(\hbar\) is the reduced Planck constant, \(m(x,E)\) is the electron effective mass depending on the energy state \(E\), and \(V(x)\) is the confinement potential.

\(m\) and \(V\) are discontinuous across the heterojunction.

**Boundary and interface conditions**

\[\Phi = 0 \quad \text{on outer boundary of matrix } \Omega_m\]

**BenDaniel–Duke condition**

\[\frac{1}{m_m} \frac{\partial \Phi}{\partial n} \bigg|_{\partial \Omega_m} = \frac{1}{m_q} \frac{\partial \Phi}{\partial n} \bigg|_{\partial \Omega_q} \quad \text{on interface}\]
Example 3 ct.

Governing equation: Schrödinger equation

\[-\nabla \cdot \left( \frac{\hbar^2}{2m(x, E)} \nabla \Phi \right) + V(x)\Phi = E\Phi, \quad x \in \Omega_q \cup \Omega_m\]

where \(\hbar\) is the reduced Planck constant, \(m(x, E)\) is the electron effective mass depending on the energy state \(E\), and \(V(x)\) is the confinement potential.

\(m\) and \(V\) are discontinuous across the heterojunction.

Boundary and interface conditions

\[\Phi = 0 \quad \text{on outer boundary of matrix } \Omega_m\]

BenDaniel–Duke condition

\[\frac{1}{m_m} \frac{\partial \Phi}{\partial n} \bigg|_{\partial \Omega_m} = \frac{1}{m_q} \frac{\partial \Phi}{\partial n} \bigg|_{\partial \Omega_q} \quad \text{on interface}\]

Details in Talk tomorrow
Classes of nonlinear eigenproblems

Quadratic eigenvalue problems
Examples

Classes of nonlinear eigenproblems

Quadratic eigenvalue problems

Recent survey: Tisseur, Meerbergen 2001

\[ T(\lambda) = \lambda^2 A + \lambda B + C, \quad A, B, C \in \mathbb{C}^{n \times n} \]

- Dynamic analysis of structures
- Stability analysis of flows in fluid mechanics
- Signal processing
- Vibration of spinning structures
- Vibration of fluid-solid structures (Conca et al. 1992)
- Lateral buckling analysis (Prandtl 1899)
- Corner singularities of anisotropic material (Wendland et al. 1992)
- Constrained least squares problems (Golub 1973)
- Regularization of total least squares problems (Sima et al. 2004)
Classes of nonlinear eigenproblems

Quadratic eigenvalue problems
Polynomial eigenvalue problems
Classes of nonlinear eigenproblems

Quadratic eigenvalue problems
Polynomial eigenvalue problems

\[ T(\lambda) = \sum_{j=0}^{k} \lambda^j A_j, \quad A_j \in \mathbb{C}^{n \times n} \]

- Optimal control problems (Mehrmann 1991)
- Singularities in linear elasticity near vertex of a cone (Kozlov et al. 1992)
- Dynamic element discretization of linear eigenproblems (V. 1987)
- Least squares element methods (Rothe 1989)
- System of coupled Euler-Bernoulli and Timoshenko beams (Balakrishnan et al. 2004)
- Nonlinear integrated optics (Botchev et al. 2008)
- Electronic states of quantum dots (Hwang et al. 2005)
Classes of nonlinear eigenproblems

Quadratic eigenvalue problems
Polynomial eigenvalue problems
Rational eigenvalue problems
Classes of nonlinear eigenproblems

Quadratic eigenvalue problems
Polynomial eigenvalue problems
Rational eigenvalue problems

\[ T(\lambda) = \lambda^2 M + K + \sum_{m=1}^{\text{NMAT}} \frac{1}{1 + b_m \lambda} \Delta K_m \]

- Vibration of structures with viscoelastic damping
- Vibration of sandwich plates (Soni 1981)
- Dynamics of plates with concentrated masses (Andreev et al. 1988)
- Vibration of fluid-solid structures (Conca et al. 1989)
- Vibration of rails exited by fast trains (Mehrmann, Watkins 2003)
- Singularities in hydrodynamics near vertex of a cone (Kozlov et al. 1994)
- Exact condensation of linear eigenvalue problems (Wittrick, Williams 1971)
- Electronic states of semiconductor heterostructures (Luttinger, Kohn 1954)
Classes of nonlinear eigenproblems

- Quadratic eigenvalue problems
- Polynomial eigenvalue problems
- Rational eigenvalue problems
- General nonlinear eigenproblems
Classes of nonlinear eigenproblems

- Quadratic eigenvalue problems
- Polynomial eigenvalue problems
- Rational eigenvalue problems
- General nonlinear eigenproblems

\[ T(\lambda) = K - \lambda M + i \sum_{j=1}^{p} \sqrt{\lambda - \sigma_j} W_j \]

- Exact dynamic element methods
- Nonlinear eigenproblems in accelerator design (Igarashi et al. 1995)
- Vibrations of poroelastic structures (Dazel et al. 2002)
- Vibro-acoustic behavior of piezoelectric/poroelastic structures (Batifol et al. 2007)
- Nonlinear integrated optics (Dirichlet-to-Neumann) (Botchev 2008)
- Stability of acoustic pressure levels in combustion chambers (van Leeuwen 2007)
Classes of nonlinear eigenproblems

- Quadratic eigenvalue problems
- Polynomial eigenvalue problems
- Rational eigenvalue problems
- General nonlinear eigenproblems
- Exponential dependence on eigenvalues
Classes of nonlinear eigenproblems

Quadratic eigenvalue problems
Polynomial eigenvalue problems
Rational eigenvalue problems
General nonlinear eigenproblems
Exponential dependence on eigenvalues

\[
\frac{d}{dt}s(t) = A_0 s(t) + \sum_{j=1}^{p} A_j s(t - \tau_j), \quad s(t) = xe^{\lambda t}
\]

\[\Rightarrow T(\lambda) = -\lambda I + A_0 + \sum_{j=1}^{p} e^{-\tau_j \lambda} A_j\]

- Stability of time-delay systems (Jarlebring 2008)
Classes of nonlinear eigenproblems

Quadratic eigenvalue problems
Polynomial eigenvalue problems
Rational eigenvalue problems
General nonlinear eigenproblems
Exponential dependence on eigenvalues
Two real parameter complex nonlinear eigenproblem
Classes of nonlinear eigenproblems

Quadratic eigenvalue problems
Polynomial eigenvalue problems
Rational eigenvalue problems
General nonlinear eigenproblems
Exponential dependence on eigenvalues
Two real parameter complex nonlinear eigenproblem

\[ T(\omega, \tau)x = (i\omega M + A + e^{-i\omega \tau} B)x = 0 \]

- Loss of stability for delay differential equations (Meerbergen, Schröder, Voss 2013)
Outline

1. Examples

2. Basic Properties

3. Rayleigh Functional

4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method

5. Invariant pairs
Definitions

If $T(\hat{\lambda})x = 0$ has a nontrivial solution $\hat{x} \neq 0$, then $\hat{\lambda}$ is an eigenvalue of $T(\cdot)$ and $\hat{x}$ is a corresponding (right) eigenvector. Every solution $\hat{y} \neq 0$ of the adjoint equation $T(\hat{\lambda})^*y = 0$ is a left eigenvector, and the vector-scalar-vector triplet $(\hat{y}, \hat{\lambda}, \hat{x})$ is called eigentriplet of $T(\cdot)$.
If $T(\hat{\lambda})x = 0$ has a nontrivial solution $\hat{x} \neq 0$, then $\hat{\lambda}$ is an eigenvalue of $T(\cdot)$ and $\hat{x}$ is a corresponding (right) eigenvector. Every solution $\hat{y} \neq 0$ of the adjoint equation $T(\hat{\lambda})^*y = 0$ is a left eigenvector, and the vector-scalar-vector triplet $(\hat{y}, \hat{\lambda}, \hat{x})$ is called eigentriplet of $T(\cdot)$.

An eigenvalue $\hat{\lambda}$ of $T(\cdot)$ has the algebraic multiplicity $k$, if

$$\left. \frac{d^\ell}{d\lambda^\ell} (\det T(\lambda)) \right|_{\lambda = \hat{\lambda}} = 0, \text{ for } \ell = 0, \ldots, k - 1 \text{ and } \left. \frac{d^k}{d\lambda^k} (\det T(\lambda)) \right|_{\lambda = \hat{\lambda}} \neq 0.$$
Definitions

If \( T(\hat{\lambda})x = 0 \) has a nontrivial solution \( \hat{x} \neq 0 \), then \( \hat{\lambda} \) is an eigenvalue of \( T(\cdot) \) and \( \hat{x} \) is a corresponding (right) eigenvector. Every solution \( \hat{y} \neq 0 \) of the adjoint equation \( T(\hat{\lambda})^* y = 0 \) is a left eigenvector, and the vector-scalar-vector triplet \( (\hat{y}, \hat{\lambda}, \hat{x}) \) is called eigentriplet of \( T(\cdot) \).

An eigenvalue \( \hat{\lambda} \) of \( T(\cdot) \) has the algebraic multiplicity \( k \), if

\[
\frac{d^\ell}{d\lambda^\ell} (\det T(\lambda)) \bigg|_{\lambda=\hat{\lambda}} = 0, \text{ for } \ell = 0, \ldots, k - 1 \quad \text{and} \quad \frac{d^k}{d\lambda^k} (\det T(\lambda)) \bigg|_{\lambda=\hat{\lambda}} \neq 0.
\]

An eigenvalue \( \hat{\lambda} \) is simple, if its algebraic multiplicity is 1.
Definitions

If $T(\hat{\lambda})x = 0$ has a nontrivial solution $\hat{x} \neq 0$, then $\hat{\lambda}$ is an eigenvalue of $T(\cdot)$ and $\hat{x}$ is a corresponding (right) eigenvector. Every solution $\hat{y} \neq 0$ of the adjoint equation $T(\hat{\lambda})^*y = 0$ is a left eigenvector, and the vector-scalar-vector triplet $(\hat{y}, \hat{\lambda}, \hat{x})$ is called eigentriplet of $T(\cdot)$.

An eigenvalue $\hat{\lambda}$ of $T(\cdot)$ has the algebraic multiplicity $k$, if

$$\frac{d^\ell}{d\lambda^\ell} (\det T(\lambda)) \bigg|_{\lambda = \hat{\lambda}} = 0, \text{ for } \ell = 0, \ldots, k - 1 \text{ and } \frac{d^k}{d\lambda^k} (\det T(\lambda)) \bigg|_{\lambda = \hat{\lambda}} \neq 0.$$

An eigenvalue $\hat{\lambda}$ is simple, if its algebraic multiplicity is 1.

The geometric multiplicity of an eigenvalue $\hat{\lambda}$ is the dimension of the kernel of $T(\hat{\lambda})$. 


Definitions

If $T(\hat{\lambda})x = 0$ has a nontrivial solution $\hat{x} \neq 0$, then $\hat{\lambda}$ is an eigenvalue of $T(\cdot)$ and $\hat{x}$ is a corresponding (right) eigenvector. Every solution $\hat{y} \neq 0$ of the adjoint equation $T(\hat{\lambda})*y = 0$ is a left eigenvector, and the vector-scalar-vector triplet $(\hat{y}, \hat{\lambda}, \hat{x})$ is called eigentriplet of $T(\cdot)$.

An eigenvalue $\hat{\lambda}$ of $T(\cdot)$ has the algebraic multiplicity $k$, if

$$\frac{d^\ell}{d\lambda^\ell} (\det T(\lambda)) \bigg|_{\lambda=\hat{\lambda}} = 0, \text{ for } \ell = 0, \ldots, k - 1 \text{ and } \frac{d^k}{d\lambda^k} (\det T(\lambda)) \bigg|_{\lambda=\hat{\lambda}} \neq 0.$$

An eigenvalue $\hat{\lambda}$ is simple, if its algebraic multiplicity is 1.

The geometric multiplicity of an eigenvalue $\hat{\lambda}$ is the dimension of the kernel of $T(\hat{\lambda})$.

An eigenvalue $\hat{\lambda}$ is called semi–simple, if its algebraic and geometric multiplicities coincide.
For $A \in \mathbb{C}^{n \times n}$ and $T(\lambda) := \lambda I - A$ the terms (left and right) eigenvector, eigenpair, eigentriplet, spectrum, algebraic and geometric multiplicity, and semi–simple have their standard meaning.
For $A \in \mathbb{C}^{n \times n}$ and $T(\lambda) := \lambda I - A$ the terms (left and right) eigenvector, eigenpair, eigentriplet, spectrum, algebraic and geometric multiplicity, and semi–simple have their standard meaning.

The following facts known for linear eigenproblems are not true for nonlinear eigenproblem:

1. eigenvectors corresponding to distinct eigenvalues are linearly independent:
For $A \in \mathbb{C}^{n \times n}$ and $T(\lambda) := \lambda I - A$ the terms (left and right) eigenvector, eigenpair, eigentriplet, spectrum, algebraic and geometric multiplicity, and semi–simple have their standard meaning.

The following facts known for linear eigenproblems are not true for nonlinear eigenproblem:

1. eigenvectors corresponding to distinct eigenvalues are linearly independent:

   **Example 1** For the quadratic eigenvalue problem $T(\lambda)x = 0$ with
   
   
   $$T(\lambda) := \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} + \lambda \begin{bmatrix} 7 & -5 \\ 10 & -8 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
   
   the distinct eigenvalues $\lambda = 1$ and $\lambda = 2$ share the eigenvector $[1, 2]^T$. 
2. Left and right eigenvectors corresponding to distinct eigenvalues are orthogonal.
2. Left and right eigenvectors corresponding to distinct eigenvalues are orthogonal.

Example 2

\[ T(\lambda)x := \begin{bmatrix} e^{i\lambda^2} & 1 \\ 1 & 1 \end{bmatrix} x = 0 \]

has a countable set of eigenvalues \( \sqrt{2k\pi}, k \in \mathbb{N}_0 \).
2. Left and right eigenvectors corresponding to distinct eigenvalues are orthogonal.

Example 2

\[ T(\lambda)x := \begin{bmatrix} e^{i\lambda^2} & 1 \\ 1 & 1 \end{bmatrix} x = 0 \]

has a countable set of eigenvalues \( \sqrt{2k\pi}, k \in \mathbb{N}_0 \).

\( \hat{\lambda} = 0 \) is an algebraically double and geometrically simple eigenvalue with left and right eigenvectors \( \hat{x} = \hat{y} = [1; -1]^T \), and \( \hat{y}^H T'(0)\hat{x} = 0 \).
2. Left and right eigenvectors corresponding to distinct eigenvalues are orthogonal.

Example 2

\[ T(\lambda)x := \begin{bmatrix} e^{i\lambda^2} & 1 \\ 1 & 1 \end{bmatrix} x = 0 \]

has a countable set of eigenvalues \( \sqrt{2k\pi}, \ k \in \mathbb{N}_0 \).

\( \hat{\lambda} = 0 \) is an algebraically double and geometrically simple eigenvalue with left and right eigenvectors \( \hat{x} = \hat{y} = [1; -1]^T \), and \( \hat{y}^H T'(0)\hat{x} = 0 \).

Every \( \hat{\lambda}_k = \sqrt{2k\pi} \) is algebraically and geometrically simple with the same eigenvectors \( \hat{x}, \hat{y} \) as before, and

\[ \hat{y}^H T'(\hat{\lambda}_k)\hat{x} = 2\sqrt{2k\pi}i \neq 0. \]
3. The algebraic multiplicities of all eigenvalues sum up to the dimension of the problem, whereas for nonlinear problems there may exist an infinite number of eigenvalues (cf. Example in item 2.), and an eigenvalue may have any algebraic multiplicity.
3. The algebraic multiplicities of all eigenvalues sum up to the dimension of the problem, whereas for nonlinear problems there may exist an infinite number of eigenvalues (cf. Example in item 2.), and an eigenvalue may have any algebraic multiplicity.

Example 3:

\[ T(\lambda) = [\lambda^k], \quad k \in \mathbb{N} \]

has the eigenvalue \( \hat{\lambda} = 0 \) with algebraic multiplicity \( k \).
The following facts known for linear eigenproblems are also true for nonlinear eigenproblem:
The following facts known for linear eigenproblems are also true for nonlinear eigenproblem:

1. If $\hat{\lambda}$ is an algebraically simple eigenvalue of $T(\cdot)$, then $\hat{\lambda}$ is geometrically simple.
The following facts known for linear eigenproblems are also true for nonlinear eigenproblem:

1. If $\hat{\lambda}$ is an algebraically simple eigenvalue of $T(\cdot)$, then $\hat{\lambda}$ is geometrically simple.

2. Let $(\hat{y}, \hat{\lambda}, \hat{x})$ be an eigentriplet of $T(\cdot)$. The $\hat{\lambda}$ is algebraically simple if and only if $\hat{\lambda}$ is geometrically simple and

   $$\hat{y}^HT'(\lambda)\hat{x} \neq 0.$$
The following facts known for linear eigenproblems are also true for nonlinear eigenproblem:

1. If $\hat{\lambda}$ is an algebraically simple eigenvalue of $T(\cdot)$, then $\hat{\lambda}$ is geometrically simple.

2. Let $(\hat{y}, \hat{\lambda}, \hat{x})$ be an eigentriplet of $T(\cdot)$. The $\hat{\lambda}$ is algebraically simple if and only if $\hat{\lambda}$ is geometrically simple and

$$\hat{y}^H T'(\lambda) \hat{x} \neq 0.$$ 

The proofs (cf. Schreiber 2008) require only basic linear algebra, but are quite lengthy.
3. If $T(\cdot)$ is real symmetric and $\lambda$ is a real eigenvalue, then left and right eigenvectors corresponding to $\lambda$ coincide.
3. If $T(\cdot)$ is real symmetric and $\lambda$ is a real eigenvalue, then left and right eigenvectors corresponding to $\lambda$ coincide.

4. If $T(\cdot)$ is complex and $x$ is a right eigenvector, then $\overline{x}$ is a left eigenvector corresponding to the same eigenvalue.
3. If $T(\cdot)$ is real symmetric and $\lambda$ is a real eigenvalue, then left and right eigenvectors corresponding to $\lambda$ coincide.

4. If $T(\cdot)$ is complex and $x$ is a right eigenvector, then $\overline{x}$ is a left eigenvector corresponding to the same eigenvalue.

5. If $T(\cdot)$ is Hermitian, then eigenvalues are real (and left and right eigenvectors corresponding to $\lambda$ coincide), or they come in pairs, i.e. if $(y, \lambda, x)$ is an eigentriplet of $T(\cdot)$, then this is also true for $(x, \overline{\lambda}, y)$. 
3. If $T(\cdot)$ is real symmetric and $\lambda$ is a real eigenvalue, then left and right eigenvectors corresponding to $\lambda$ coincide.

4. If $T(\cdot)$ is complex and $x$ is a right eigenvector, then $\overline{x}$ is a left eigenvector corresponding to the same eigenvalue.

5. If $T(\cdot)$ is Hermitian, then eigenvalues are real (and left and right eigenvectors corresponding to $\lambda$ coincide), or they come in pairs, i.e. if $(y, \lambda, x)$ is an eigentriplet of $T(\cdot)$, then this is also true for $(x, \overline{\lambda}, y)$.

The proofs require only basic linear algebra they are not lengthy, but are left as an exercise.
Theorem (Neumeier 1985)
Let $T(\cdot)$ be differentiable and let $\hat{\lambda}$ be a simple eigenvalue of $T(\cdot)$. Let $\hat{x}$ be a right eigenvector normalized such that $v^Hx = 1$ for some vector $v$. Then the matrix

$$B := T(\hat{\lambda}) + T'(\hat{\lambda})\hat{x}v^H$$

is nonsingular.
Theorem (Neumeier 1985)
Let $T(\cdot)$ be differentiable and let $\hat{\lambda}$ be a simple eigenvalue of $T(\cdot)$. Let $\hat{x}$ be a right eigenvector normalized such that $v^Hx = 1$ for some vector $v$. Then the matrix

$$B := T(\hat{\lambda}) + T'(\hat{\lambda})\hat{x}v^H$$

is nonsingular.

Representations for the resolvent $(A - \lambda I)^{-1}$ for an arbitrary matrix $A$ can be derived by several available decompositions, such as the spectral decomposition, Schur decomposition, and singular value decomposition. It turns out that in this case the norm of the inverse shifted operator depends on the eigenvalue gap $(\min_{\lambda_i \in \sigma(A)}(\lambda_i - \lambda))^{-1}$. For the nonlinear case we have
**Facts**

**Theorem (Neumeier 1985)**
Let $T(\cdot)$ be differentiable and let $\hat{\lambda}$ be a simple eigenvalue of $T(\cdot)$. Let $\hat{x}$ be a right eigenvector normalized such that $v^Hx = 1$ for some vector $v$. Then the matrix

$$B := T(\hat{\lambda}) + T'(\hat{\lambda})\hat{x}v^H$$

is nonsingular.

Representations for the resolvent $(A - \lambda I)^{-1}$ for an arbitrary matrix $A$ can be derived by several available decompositions, such as the spectral decomposition, Schur decomposition, and singular value decomposition. It turns out that in this case the norm of the inverse shifted operator depends on the eigenvalue gap $(\min_{\lambda_i \in \sigma(A)} (\lambda_i - \lambda))^{-1}$. For the nonlinear case we have

**Theorem (Schreiber 2008)**
Let $(\hat{y}, \hat{\lambda}, \hat{x})$ be an eigentriplet with simple eigenvalue $\hat{\lambda}$. Then for sufficiently small $|\lambda - \hat{\lambda}|$

$$T(\lambda)^{-1} = \frac{1}{\lambda - \hat{\lambda}} \frac{\hat{x}\hat{y}^H}{\hat{y}^HT'(\hat{\lambda})\hat{x}} + O(1).$$
Outline

1. Examples

2. Basic Properties

3. Rayleigh Functional

4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method

5. Invariant pairs
In this section we collect results on the existence and approximation properties of a Rayleigh functional in a vicinity of eigenvectors corresponding to algebraically simple eigenvalues. Rayleigh functionals will play a completely different, but also important role in the talk on variational characterizations of eigenvalues tomorrow.
In this section we collect results on the existence and approximation properties of a Rayleigh functional in a vicinity of eigenvectors corresponding to algebraically simple eigenvalues. Rayleigh functionals will play a completely different, but also important role in the talk on variational characterizations of eigenvalues tomorrow.

Let $T : \mathbb{C} \supset D \to \mathbb{C}^{n \times n}$ be differentiable in an open set $D$ with Lipschitz continuous derivative.

Let $(\hat{\lambda}, \hat{x})$ be an eigenpair of $T(\cdot)$, and define neighborhoods

$B(\hat{\lambda}, \tau) := \{ \lambda \in \mathbb{C} : |\lambda - \hat{\lambda}| < \tau \}$ and

$\mathcal{K}_\varepsilon(\hat{x}) := \{ x \in \mathbb{C}^n : \angle(\text{span}\{x\}, \text{span}\{\hat{x}\}) \leq \varepsilon \}$ of $\hat{\lambda}$ and $\hat{x}$, respectively.
Rayleigh Functional

In this section we collect results on the existence and approximation properties of a Rayleigh functional in a vicinity of eigenvectors corresponding to algebraically simple eigenvalues. Rayleigh functionals will play a completely different, but also important roll in the talk on variational characterizations of eigenvalues tomorrow.

Let $T : \mathbb{C} \supset D \to \mathbb{C}^{n \times n}$ be differentiable in an open set $D$ with Lipschitz continuous derivative.

Let $(\hat{\lambda}, \hat{x})$ be an eigenpair of $T(\cdot)$, and define neighborhoods

$B(\hat{\lambda}, \tau) := \{ \lambda \in \mathbb{C} : |\lambda - \hat{\lambda}| < \tau \}$ and

$\mathcal{K}_\varepsilon(\hat{x}) := \{ x \in \mathbb{C}^n : \angle(\text{span}\{x\}, \text{span}\{\hat{x}\}) \leq \varepsilon \}$ of $\hat{\lambda}$ and $\hat{x}$, respectively.

Let $(\hat{y}, \hat{\lambda}, \hat{x})$ be an eigentriplet of $T(\cdot)$.

$p : \mathcal{K}_\varepsilon(\hat{x}) \times \mathcal{K}_\varepsilon(\hat{y}) \to B(\hat{\lambda}, \tau)$ is a two-sided Rayleigh functional if the following conditions hold for every $x \in \mathcal{K}_\varepsilon(\hat{x})$ and every $y \in \mathcal{K}_\varepsilon(\hat{y})$:

(i) $p(\alpha x, \beta y) = p(x, y)$ for every $\alpha, \beta \in \mathbb{C} \setminus \{0\}$

(ii) $y^H T(p(x, y))x = 0$

(iii) $y^H T'(p(x, y))x \neq 0$. 
Theorem 1 (Schwetlick, Schreiber 2012)
Let \((\hat{y}, \hat{\lambda}, \hat{x})\) be an eigentriplet of \(T(\cdot)\) with \(\|\hat{x}\| = \|\hat{y}\| = 1\), and assume that \(\hat{y}^H T'(\hat{\lambda}) \hat{x} \neq 0\).

Then there exist \(\varepsilon > 0\) and \(\tau > 0\) such that the two-sided Rayleigh functional \(p(\cdot)\) is defined in \(K_{\varepsilon}(\hat{x}) \times K_{\varepsilon}(\hat{y})\), and it holds that

\[
|p(x, y) - \hat{\lambda}| \leq \frac{8}{3} \|T(\hat{\lambda})\| \frac{\tan \xi \tan \eta}{|\hat{y}^H T'(\hat{\lambda}) \hat{x}|}
\]

where \(\xi := \angle(\text{span}\{x\}, \text{span}\{\hat{x}\})\) and \(\eta := \angle(\text{span}\{y\}, \text{span}\{\hat{y}\})\).
Rayleigh Functional

**Theorem 1** (Schwetlick, Schreiber 2012)
Let \((\hat{y}, \hat{\lambda}, \hat{x})\) be an eigentriplet of \(T(\cdot)\) with \(\|\hat{x}\| = \|\hat{y}\| = 1\), and assume that \(\hat{y}^H T'(\hat{\lambda})\hat{x} \neq 0\).

Then there exist \(\varepsilon > 0\) and \(\tau > 0\) such that the two-sided Rayleigh functional \(p(\cdot)\) is defined in \(K_{\varepsilon}(\hat{x}) \times K_{\varepsilon}(\hat{y})\), and it holds that

\[
|p(x, y) - \hat{\lambda}| \leq \frac{8}{3} \frac{\|T(\hat{\lambda})\|}{|\hat{y}^H T'(\hat{\lambda})\hat{x}|} \tan \xi \tan \eta,
\]

where \(\xi := \angle(\text{span}\{x\}, \text{span}\{\hat{x}\})\) and \(\eta := \angle(\text{span}\{y\}, \text{span}\{\hat{y}\})\).

**Theorem 2** (Schwetlick, Schreiber 2012)
Under the conditions of Thm. 1 let \(\xi < \pi/3\) and \(\eta < \pi/3\). Then

\[
|p(x, y) - \hat{\lambda}| \leq \frac{32}{3} \frac{\|T(\hat{\lambda})\|}{|\hat{y}^H T'(\hat{\lambda})\hat{x}|} \|x - \hat{x}\| \|y - \hat{y}\|,
\]

where \(\xi := \angle(\text{span}\{x\}, \text{span}\{\hat{x}\})\).
Theorem 3 (Schwetlick, Schreiber 2012)
Unter the conditions of Thm. 1 the two–sided Rayleigh functional is stationary at \((\hat{x}, \hat{y})\), i.e.

\[
|p(\hat{x} + s, \hat{y} + t) - \hat{\lambda}| = \mathcal{O}(\|s\| + \|t\|)^2).
\]
**Theorem 3** (Schwetlick, Schreiber 2012)
Unter the conditions of Thm. 1 the two–sided Rayleigh functional is stationary at \((\hat{x}, \hat{y})\), i.e.

\[
|p(\hat{x} + s, \hat{y} + t) - \hat{\lambda}| = O((\|s\| + \|t\|)^2).
\]

**Definition**

\[
p : \mathcal{K}_{\varepsilon} \to B(\hat{\lambda}, \tau)
\]

is a (one-sided) Rayleigh functional if the following conditions hold:

(i) \(p(\alpha x) = p(x)\) for every \(\alpha \in \mathbb{C}, \alpha \neq 0\)

(ii) \(x^H T(p(x))x = 0\) for every \(x \in \mathcal{K}_{\varepsilon}(\hat{x})\)

(iii) \(x^H T'(p(x))x \neq 0\) for every \(x \in \mathcal{K}_{\varepsilon}(\hat{x})\).
Theorem 3 (Schwetlick, Schreiber 2012)
Let \((\hat{\lambda}, \hat{x})\) be an eigenpair of \(T(\cdot)\) with \(\|\hat{x}\| = 1\) and \(\hat{x}^H T'(\hat{\lambda}) \hat{x} \neq 0\), and suppose that \(T(\hat{\lambda}) = T(\hat{\lambda})^H\).

Then there exist \(\varepsilon > 0\) and \(\tau > 0\) such that the one-sided Rayleigh functional \(p(\cdot)\) is defined in \(K_{\varepsilon}(\hat{x})\), and it holds that

\[
|p(x) - \hat{\lambda}| \leq \frac{8}{3} \frac{\|T(\hat{\lambda})\|}{|\hat{x}^H T'(\hat{\lambda}) \hat{x}|} \tan^2 \xi,
\]

where \(\xi := \angle(\text{span}\{x\}, \text{span}\{\hat{x}\})\).
Rayleigh Functional

**Theorem 3** (Schwetlick, Schreiber 2012)
Let $(\hat{\lambda}, \hat{x})$ be an eigenpair of $T(\cdot)$ with $\|\hat{x}\| = 1$ and $\hat{x}^H T'(\hat{\lambda})\hat{x} \neq 0$, and suppose that $T(\hat{\lambda}) = T(\hat{\lambda})^H$. Then there exist $\varepsilon > 0$ and $\tau > 0$ such that the one-sided Rayleigh functional $p(\cdot)$ is defined in $K_\varepsilon(\hat{x})$, and it holds that

$$|p(x) - \hat{\lambda}| \leq \frac{8}{3} \left| \frac{\|T(\hat{\lambda})\|}{\hat{x}^H T'(\hat{\lambda})\hat{x}} \right| \tan^2 \xi,$$

where $\xi := \angle(\text{span}\{x\}, \text{span}\{\hat{x}\})$.

**Theorem 4** (Schwetlick, Schreiber 2012)
Let $(\hat{\lambda}, \hat{x})$ be an eigenpair of $T(\cdot)$ with $\|\hat{x}\| = 1$ and $\hat{x}^H T'(\hat{\lambda})\hat{x} \neq 0$. Then there exist $\varepsilon > 0$ and $\tau > 0$ such that the one-sided Rayleigh functional $p(\cdot)$ is defined in $K_\varepsilon(\hat{x})$, and it holds that

$$|p(x) - \hat{\lambda}| \leq \frac{10}{3} \left| \frac{\|T(\hat{\lambda})\|}{\hat{x}^H T'(\hat{\lambda})\hat{x}} \right| \tan \xi,$$

where $\xi := \angle(\text{span}\{x\}, \text{span}\{\hat{x}\})$.
In general, the Rayleigh functional is not easy to evaluate. Applying Newton’s method to the defining equation $y^H T(p(x))x = 0$ one obtains the generalized Rayleigh quotient

$$p_L : B(\hat{\lambda}, \tau) \times K_\varepsilon(\hat{x}) \times K_\varepsilon(\hat{y}) \to B(\hat{\lambda}, \tau), \quad p_L(\hat{\lambda}, x, y) := \lambda - \frac{y^H T(\lambda)x}{y^H T'(\lambda)x}.$$
In general, the Rayleigh functional is not easy to evaluate. Applying Newton’s method to the defining equation $y^H T(p(x))x = 0$ one obtains the generalized Rayleigh quotient

$$p_L : B(\hat{\lambda}, \tau) \times \mathcal{K}_\varepsilon(\hat{x}) \times \mathcal{K}_\varepsilon(\hat{y}) \rightarrow B(\hat{\lambda}, \tau), \ p_L(\hat{\lambda}, x, y) := \lambda - \frac{y^H T(\lambda)x}{y^H T'(\lambda)x}.$$  

**Theorem 5** (Schwetlick, Schreiber 2012)
Under the conditions of Thm 1 the generalized Rayleigh quotient $p_L$ is defined for all $\lambda \in B(\hat{\lambda}, \tau)$ and $(x, y) \in \mathcal{K}_\varepsilon(\hat{x}) \times \mathcal{K}_\varepsilon(\hat{y})$, and it holds that

$$|p_L(\lambda, x, y) - \hat{\lambda}| \leq \frac{4\|T(\hat{\lambda})\|}{|\hat{y}^H T'(\hat{\lambda})\hat{x}|} \tan \xi \tan \eta + \frac{2L}{|\hat{y}^H T'(\hat{\lambda})\hat{x}|} \frac{|\lambda - \hat{\lambda}|^2}{\cos \xi \cos \eta},$$

where $L$ denotes the Lipschitz constant of $T'(\cdot)$. 
Outline

1. Examples
2. Basic Properties
3. Rayleigh Functional
4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method
5. Invariant pairs
For polynomial or rational eigenproblems the simplest approach is to use linearization and to apply standard methods for linear eigenproblems.
Dense nonlinear eigenproblems

For polynomial or rational eigenproblems the simplest approach is to use linearization and to apply standard methods for linear eigenproblems.

For general nonlinear eigenvalue problems, the classical approach is to formulate the eigenvalue problem as a system of nonlinear equations and to use variations of Newton’s method or the inverse iteration method.
For polynomial or rational eigenproblems the simplest approach is to use linearization and to apply standard methods for linear eigenproblems.

For general nonlinear eigenvalue problems, the classical approach is to formulate the eigenvalue problem as a system of nonlinear equations and to use variations of Newton’s method or the inverse iteration method.

Dense solvers only apply to relatively small eigenproblems (maybe dimensions up to 1000, but this depends on the computer in use). For higher dimensions iterative projection methods like Jacobi–Davidson method or the Nonlinear Arnoldi method are more appropriate. Methods of this type are considered in the lecture on fluid–solid vibrations.
Outline

1. Examples
2. Basic Properties
3. Rayleigh Functional
4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method
5. Invariant pairs
Let

\[ P(\lambda) := \sum_{j=0}^{m} \lambda^j A_j \]

be a regular (i.e. \( \text{det} P(\lambda) \neq 0 \)) matrix polynomial.
Let

\[ P(\lambda) := \sum_{j=0}^{m} \lambda^j A_j \]

be a regular (i.e. \( \det P(\lambda) \neq 0 \)) matrix polynomial.

\[ L(\lambda) = \lambda X + Y, \quad X, Y \in \mathbb{C}^{mn \times mn} \]

is a **Linearization** of \( P(\lambda) \) if

\[ E(\lambda)L(\lambda)F(\lambda) = \begin{pmatrix} P(\lambda) & O \\ O & I_{(m-1)n} \end{pmatrix} \]

for some **unimodular** \( E(\lambda) \) and \( F(\lambda) \) (i.e. \( \det E(\lambda) = \pm 1, \det F(\lambda) = \pm 1 \)).
Let

\[ P(\lambda) := \sum_{j=0}^{m} \lambda^j A_j \]

be a regular (i.e. \( \det P(\lambda) \neq 0 \)) matrix polynomial.

\[ L(\lambda) = \lambda X + Y, \quad X, Y \in \mathbb{C}^{mn \times mn} \]

is a Linearization of \( P(\lambda) \) if

\[ E(\lambda)L(\lambda)F(\lambda) = \begin{pmatrix} P(\lambda) & O \\ O & I_{(m-1)n} \end{pmatrix} \]

for some unimodular \( E(\lambda) \) and \( F(\lambda) \) (i.e. \( \det E(\lambda) = \pm 1, \det F(\lambda) = \pm 1 \)).

Obviously, \( \lambda \) is an eigenvalue of \( p(\lambda) \) if and only if \( \lambda \) is an eigenvalue of \( L(\lambda) \).
Companion form of linearization

\[ L(\lambda) = \lambda \begin{pmatrix} A_m & 0 & \cdots & 0 \\ 0 & I_n & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n \end{pmatrix} + \begin{pmatrix} A_{m-1} & A_{m-2} & \cdots & A_0 \\ -I_n & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -I_n & 0 \end{pmatrix}. \]
Companion form of linearization

\[ L(\lambda) = \lambda \begin{pmatrix} A_m & 0 & \cdots & 0 \\ 0 & I_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & I_n \end{pmatrix} + \begin{pmatrix} A_{m-1} & A_{m-2} & \cdots & A_0 \\ -I_n & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -I_n & 0 \end{pmatrix}. \]

\( x \) is a right eigenvector of \( P(\lambda) \) if and only if

\[ z = \begin{pmatrix} \lambda^{m-1}x \\ \lambda^{m-2}x \\ \vdots \\ \lambda x \\ x \end{pmatrix} \]

is a right eigenvector of \( L(\lambda) \).
Companion form of linearization

\[
L(\lambda) = \lambda \begin{pmatrix}
A_m & 0 & \cdots & 0 \\
0 & I_n & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & I_n
\end{pmatrix} + \begin{pmatrix}
A_{m-1} & A_{m-2} & \cdots & A_0 \\
-l_n & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & -l_n & 0
\end{pmatrix}.
\]

\(x\) is a right eigenvector of \(P(\lambda)\) if and only if

\[
z = \begin{pmatrix}
\lambda^{m-1}x \\
\lambda^{m-2}x \\
\vdots \\
\lambda x \\
x
\end{pmatrix}
\]
is a right eigenvector of \(L(\lambda)\)

Left eigenvectors: more complicated.
With

$$\Lambda := (\lambda^{m-1}, \lambda^{m-2}, \ldots, \lambda, 1)^T.$$ 

Mackey, Mackey, Mehl & Mehrmann (2006) define

$$\mathbb{L}_1(P) := \{L(\lambda) : L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda), \ v \in \mathbb{C}^m\},$$

$$\mathbb{L}_2(P) := \{L(\lambda) : (\Lambda^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda), \ w \in \mathbb{C}^m\}.$$
**\( L_1 \) and \( L_2 \) linearizations**

With

\[
\Lambda := (\lambda^{m-1}, \lambda^{m-2}, \ldots, \lambda, 1)^T.
\]

Mackey, Mackey, Mehl & Mehrmann (2006) define

\[
L_1(P) := \{ L(\lambda) : L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda), \ v \in \mathbb{C}^m \},
\]

\[
L_2(P) := \{ L(\lambda) : (\Lambda^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda), \ w \in \mathbb{C}^m \}.
\]

They show that

- \( L_1 \) and \( L_2 \) are vector spaces of dimension \( m(m - 1)n^2 + m \)
- Almost all pencils in \( L_1 \) and \( L_2 \) are linearizations of \( P(\lambda) \).
With

\( \Lambda := (\lambda^{m-1}, \lambda^{m-2}, \ldots, \lambda, 1)^T. \)

Mackey, Mackey, Mehl & Mehrmann (2006) define

\[
\mathbb{L}_1(P) := \{ L(\lambda) : L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda), \ v \in \mathbb{C}^m \},
\]

\[
\mathbb{L}_2(P) := \{ L(\lambda) : (\Lambda^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda), \ w \in \mathbb{C}^m \}.
\]

They show that

- \( \mathbb{L}_1 \) and \( \mathbb{L}_2 \) are vector spaces of dimension \( m(m-1)n^2 + m \)
- Almost all pencils in \( \mathbb{L}_1 \) and \( \mathbb{L}_2 \) are linearizations of \( P(\lambda) \).

**Quadratic case (m = 2):** \( L(\lambda) = \lambda X + Y \in \mathbb{L}_1(P) \) iff

\[
\begin{pmatrix}
  v_1 A_2 & v_1 A_1 & v_1 A_0 \\
  v_2 A_2 & v_2 A_1 & v_2 A_0
\end{pmatrix} = \begin{pmatrix}
  X_{11} & X_{12} + Y_{11} & Y_{12} \\
  X_{21} & X_{22} + Y_{21} & Y_{22}
\end{pmatrix}
\]
Recall

\[ \mathbb{L}_1(P) := \{ \mathbb{L}(\lambda) : \mathbb{L}(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda), \ v \in \mathbb{C}^m \}, \]
Recall
\[ L_1(P) := \{ L(\lambda) : L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda), \ v \in \mathbb{C}^m \}, \]

Note
\[ L(\lambda)(\Lambda \otimes x) = L(\lambda)(\Lambda \otimes I_n)(1 \otimes x) = (v \otimes P(\lambda))(1 \otimes x) = v \otimes P(\lambda)x. \]
Recall

\[ \mathbb{L}_1(P) := \{ L(\lambda) : L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda), \ v \in \mathbb{C}^m \}, \]

Note

\[ L(\lambda)(\Lambda \otimes x) = L(\lambda)(\Lambda \otimes I_n)(1 \otimes x) = (v \otimes P(\lambda))(1 \otimes x) = v \otimes P(\lambda)x. \]

So \((x, \lambda)\) is an eigenpair of \(P(\lambda)\) iff \((\Lambda \otimes x, \lambda)\) is an eigenpair of \(L(\lambda)\).

- Right eigenvectors of \(P\) can be recovered from right eigenvectors of linearizations in \(\mathbb{L}_1\).
- Left eigenvectors of \(P\) can be recovered from right eigenvectors of linearizations in \(\mathbb{L}_2\).
Mackey, Mackey, Mehl & Mehrmann (2006) define

$$DL(P) = L_1(P) \cap L_2(P).$$
Mackey, Mackey, Mehl & Mehrmann (2006) define

$$DL(P) = L_1(P) \cap L_2(P).$$

They show that

- $L \in DL(P)$ iff $w = v$ in the definitions of $L_1$ and $L_2$
- $DL(P)$ is a vector space of dimension $m$
- Almost all pencils in $DL(P)$ are linearizations of $P$
Mackey, Mackey, Mehl & Mehrmann (2006) define

\[ \mathbb{DL}(P) = L_1(P) \cap L_2(P). \]

They show that

- \( L \in \mathbb{DL}(P) \) iff \( w = v \) in the definitions of \( L_1 \) and \( L_2 \)
- \( \mathbb{DL}(P) \) is a vector space of dimension \( m \)
- Almost all pencils in \( \mathbb{DL}(P) \) are linearizations of \( P \)

For \( Q(\lambda) := \lambda^2 A + \lambda B + C \), \( L(\lambda) \in \mathbb{DL}(Q) \) iff

\[
L(\lambda) = \lambda \begin{pmatrix} v_1 A & V_2 A \\ v_2 A & v_2 B - v_1 C \end{pmatrix} + \begin{pmatrix} v_1 B - v_2 A & v_1 C \\ v_1 C & v_2 C \end{pmatrix}, \quad v \in \mathbb{C}.
\]
Linearization Theorem (4M, 2006)

Let $P(\lambda)$ be a regular matrix polynomial, and let $L(\lambda) \in \mathbb{L}_1(P)$. Then the following statements are equivalent:

- $L(\lambda)$ is a linearization of $P(\lambda)$
- $L(\lambda)$ is a regular pencil
- $L(\lambda)$ is a strong linearization of $P(\lambda)$
Linearization Theorem (4M, 2006)

Let $P(\lambda)$ be a regular matrix polynomial, and let $L(\lambda) \in \mathbb{L}_1(P)$. Then the following statements are equivalent:

- $L(\lambda)$ is a linearization of $P(\lambda)$
- $L(\lambda)$ is a regular pencil
- $L(\lambda)$ is a strong linearization of $P(\lambda)$

$L(\lambda) = \lambda X + Y$ is a **strong linearization** of $P(\lambda)$, if it is a linearization of $P(\lambda)$, and $X + \lambda Y$ is a linearization of

$$
\text{rev } P(\lambda) := \lambda^m P(1/\lambda) = \sum_{j=0}^{m} \lambda^j A_{m-j}.
$$
Linearization Theorem (4M, 2006)

Let $P(\lambda)$ be a regular matrix polynomial, and let $L(\lambda) \in \mathbb{L}_1(P)$. Then the following statements are equivalent:

- $L(\lambda)$ is a linearization of $P(\lambda)$
- $L(\lambda)$ is a regular pencil
- $L(\lambda)$ is a strong linearization of $P(\lambda)$

$L(\lambda) = \lambda X + Y$ is a strong linearization of $P(\lambda)$, if it is a linearization of $P(\lambda)$, and $X + \lambda Y$ is a linearization of

$$\text{rev } P(\lambda) := \lambda^m P(1/\lambda) = \sum_{j=0}^{m} \lambda^j A_{m-j}.$$ 

The Jordan structure of all finite eigenvalues of $P(\lambda)$ can be recovered from any linearization of $P$. For strong linearizations this is also the case for the eigenvalue $\infty$, i.e. the eigenvalue 0 of rev $P(\lambda)$. 
Outline

1. Examples
2. Basic Properties
3. Rayleigh Functional
4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method
5. Invariant pairs
For the characteristic equation

\[ \det T(\lambda) = 0, \]

it was suggested by Kublanovskaya (1969,1970) to use a QR-decomposition with column pivoting \( T(\lambda)P(\lambda) = Q(\lambda)R(\lambda) \), where \( P(\lambda) \) is a permutation matrix which is chosen such that the diagonal elements \( r_{jj}(\lambda) \) of \( R(\lambda) \) are decreasing in magnitude, i.e. \( |r_{11}(\lambda)| \geq |r_{22}(\lambda)| \geq \cdots \geq |r_{nn}(\lambda)| \). Then \( \lambda \) is an eigenvalue if and only if \( r_{nn}(\lambda) = 0 \).
For the characteristic equation

$$\det T(\lambda) = 0,$$

it was suggested by Kublanovskaya (1969, 1970) to use a QR-decomposition with column pivoting $T(\lambda)P(\lambda) = Q(\lambda)R(\lambda)$, where $P(\lambda)$ is a permutation matrix which is chosen such that the diagonal elements $r_{jj}(\lambda)$ of $R(\lambda)$ are decreasing in magnitude, i.e. $|r_{11}(\lambda)| \geq |r_{22}(\lambda)| \geq \cdots \geq |r_{nn}(\lambda)|$. Then $\lambda$ is an eigenvalue if and only if $r_{nn}(\lambda) = 0$.

Applying Newton’s method to this equation, one obtains the iteration

$$\lambda_{k+1} = \lambda_k - \frac{1}{e_n^H Q(\lambda_k)^H T'(\lambda_k) P(\lambda_k) R(\lambda_k)^{-1} e_n}$$

for approximations to an eigenvalue.
For the characteristic equation

$$\det T(\lambda) = 0,$$

it was suggested by Kublanovskaya (1969, 1970) to use a $QR$-decomposition with column pivoting $T(\lambda)P(\lambda) = Q(\lambda)R(\lambda)$, where $P(\lambda)$ is a permutation matrix which is chosen such that the diagonal elements $r_{jj}(\lambda)$ of $R(\lambda)$ are decreasing in magnitude, i.e. $|r_{11}(\lambda)| \geq |r_{22}(\lambda)| \geq \cdots \geq |r_{nn}(\lambda)|$. Then $\lambda$ is an eigenvalue if and only if $r_{nn}(\lambda) = 0$.

Applying Newton’s method to this equation, one obtains the iteration

$$\lambda_{k+1} = \lambda_k - \frac{1}{e_n^H Q(\lambda_k)^H T'(\lambda_k) P(\lambda_k) R(\lambda_k)^{-1} e_n}$$

for approximations to an eigenvalue.

Approximations to left and right eigenvectors can be obtained from

$$y_k = Q(\lambda_k)e_n \quad \text{and} \quad x_k = P(\lambda_k)R(\lambda_k)^{-1}e_n.$$
An improved version of Kublanovskaya’s method was suggested by Jain, Singhal & Huseyin (1983), and also quadratic convergence was shown.
An improved version of Kublanovskaya’s method was suggested by Jain, Singhal & Huseyin (1983), and also quadratic convergence was shown.

A similar approach was presented by Yang (1983), via a representation of Newton’s method using the $LU$-factorization of $T(\lambda)$.
An improved version of Kublanovskaya’s method was suggested by Jain, Singhal & Huseyin (1983), and also quadratic convergence was shown. A similar approach was presented by Yang (1983), via a representation of Newton’s method using the $LU$-factorization of $T(\lambda)$.

Other variations of this method can be found in the books of Zurmühl & Falk (1984, 1985).
An improved version of Kublanovskaya’s method was suggested by Jain, Singhal & Huseyin (1983), and also quadratic convergence was shown.

A similar approach was presented by Yang (1983), via a representation of Newton’s method using the $LU$-factorization of $T(\lambda)$.

Other variations of this method can be found in the books of Zurmühl & Falk (1984, 1985).

However, this relatively simple idea is not efficient, since it computes eigenvalues one at a time and needs several $O(n^3)$ factorizations per eigenvalue. It is, however, useful in the context of iterative refinement of computed eigenvalues and eigenvectors.
Outline

1. Examples
2. Basic Properties
3. Rayleigh Functional
4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method
5. Invariant pairs
Let $\lambda_k$ be an approximation to an eigenvalue of $T(\lambda)x = 0$. Linearizing $T(\lambda_k - \theta)x = 0$ yields
\[ T(\lambda_k)x = \theta T'(\lambda_k)x. \]

This suggests the method
Successive linear approximations

Let $\lambda_k$ be an approximation to an eigenvalue of $T(\lambda)x = 0$. Linearizing $T(\lambda_k - \theta)x = 0$ yields

$$T(\lambda_k)x = \theta T'(\lambda_k)x.$$ 

This suggests the method

1: start with an approximation $\lambda_1 \in D$ to an eigenvalue of $T(\lambda)x = 0$
2: for $k = 1, 2, \ldots$ until convergence do
3: solve the linear eigenproblem $T(\lambda_k)x = \theta T'(\lambda_k)x$
4: choose suitable eigenvalue $\theta$ (usually smallest in modulus)
5: set $\lambda_{k+1} = \lambda_k - \theta$
6: end for
THEOREM
Let $T(\lambda)$ be twice continuously differentiable, and let $\hat{\lambda}$ be an eigenvalue of $T(\lambda)x = 0$ such that $T'(\hat{\lambda})$ is nonsingular and 0 is an algebraically simple eigenvalue of $T'(\hat{\lambda})^{-1} T(\hat{\lambda})$. Then the method of successive linear problems converges quadratically to $\hat{\lambda}$. 

Sketch of proof: For $\Phi(x, \theta, \lambda) := (T(\lambda)x - \theta T'(\lambda)x - Hx - 1)$, it holds $\Phi(\hat{x}, 0, \hat{\lambda}) = 0$, and by the implicit function theorem $\Phi(x, \theta, \lambda) = 0$ defines differentiable functions $x: U(\hat{\lambda}) \to \mathbb{C}^n$ and $\theta: U(\hat{\lambda}) \to \mathbb{C}$ on a neighborhood of $\hat{\lambda}$ such that $\Phi(x(\lambda), \theta(\lambda), \lambda) = 0$. 

$\theta'(\hat{\lambda}) = 1$ yields the quadratic convergence of the method $\lambda_{k+1} = \lambda_k - \theta(\lambda_k)$. 


THEOREM
Let $T(\lambda)$ be twice continuously differentiable, and let $\hat{\lambda}$ be an eigenvalue of $T(\lambda)x = 0$ such that $T'(\hat{\lambda})$ is nonsingular and 0 is an algebraically simple eigenvalue of $T'(\hat{\lambda})^{-1} T(\hat{\lambda})$. Then the method of successive linear problems converges quadratically to $\hat{\lambda}$.

Sketch of proof: For

$$\Phi(x, \theta, \lambda) := \begin{pmatrix} T(\lambda)x - \theta T'(\lambda)x \\ \ell^H x - 1 \end{pmatrix}$$

it holds $\Phi(\hat{x}, 0, \hat{\lambda}) = 0$, and by the implicit function theorem $\Phi(x, \theta, \lambda) = 0$ defines differentiable functions $x : U(\hat{\lambda}) \to \mathbb{C}^n$ and $\theta : U(\hat{\lambda}) \to \mathbb{C}$ on a neighborhood of $\hat{\lambda}$ such that $\Phi(x(\lambda), \theta(\lambda), \lambda) = 0$. 
**THEOREM**

Let \( T(\lambda) \) be twice continuously differentiable, and let \( \hat{\lambda} \) be an eigenvalue of \( T(\lambda)x = 0 \) such that \( T'(\hat{\lambda}) \) is nonsingular and 0 is an algebraically simple eigenvalue of \( T'(\hat{\lambda})^{-1} T(\hat{\lambda}) \). Then the method of successive linear problems converges quadratically to \( \hat{\lambda} \).

**Sketch of proof:** For

\[
\Phi(x, \theta, \lambda) := \begin{pmatrix} T(\lambda)x - \theta T'(\lambda)x \\ \ell^H x - 1 \end{pmatrix}
\]

it holds \( \Phi(\hat{x}, 0, \hat{\lambda}) = 0 \), and by the implicit function theorem \( \Phi(x, \theta, \lambda) = 0 \) defines differentiable functions \( x : U(\hat{\lambda}) \to \mathbb{C}^n \) and \( \theta : U(\hat{\lambda}) \to \mathbb{C} \) on a neighborhood of \( \hat{\lambda} \) such that \( \Phi(x(\lambda), \theta(\lambda), \lambda) = 0 \).

\( \theta'(\hat{\lambda}) = 1 \) yields the quadratic convergence of the method \( \lambda_{k+1} = \lambda_k - \theta(\lambda_k) \).
Inverse iteration

Applying Newton’s method to the nonlinear system

\[ F \left( \begin{array}{c} x \\ \lambda \end{array} \right) = \left( \begin{array}{c} T(\lambda)x \\ \ell^H x - 1 \end{array} \right), \]

(\ell \text{ suitable scaling vector}) yields

\[
\left( \begin{array}{cc}
T(\lambda_k) & T'(\lambda_k)x_k \\
\ell^H & 0
\end{array} \right) \left( \begin{array}{c}
x_{k+1} - x_k \\
\lambda_{k+1} - \lambda_k
\end{array} \right) = - \left( \begin{array}{c}
T(\lambda_k)x_k \\
\ell^H x_k - 1
\end{array} \right).
\]
Methods for dense nonlinear eigenvalue problems

Inverse iteration

Applying Newton’s method to the nonlinear system

\[ F \left( \begin{array}{c} x \\ \lambda \end{array} \right) = \left( \begin{array}{c} T(\lambda)x \\ \ell^Hx - 1 \end{array} \right), \]

(\ell \text{ suitable scaling vector}) yields

\[ \left( \begin{array}{cc} T(\lambda_k) & T'(\lambda_k)x_k \\ \ell^H & 0 \end{array} \right) \left( \begin{array}{c} x_{k+1} - x_k \\ \lambda_{k+1} - \lambda_k \end{array} \right) = - \left( \begin{array}{c} T(\lambda_k)x_k \\ \ell^Hx_k - 1 \end{array} \right). \]

Assuming that \( x_k \) is already scaled such that \( \ell^Hx_k = 1 \) the second equation reads \( \ell^Hx_{k+1} = \ell^Hx_k = 1 \), and the first one

\[ x_{k+1} = -(\lambda_{k+1} - \lambda_k)T(\lambda_k)^{-1}T'(\lambda_k)x_k. \]
Inverse iteration

Applying Newton’s method to the nonlinear system

\[
F \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} T(\lambda)x \\ \ell^H x - 1 \end{pmatrix},
\]

(\ell \text{ suitable scaling vector}) yields

\[
\begin{pmatrix} T(\lambda_k) & T'(\lambda_k)x_k \\ \ell^H & 0 \end{pmatrix} \begin{pmatrix} x_{k+1} - x_k \\ \lambda_{k+1} - \lambda_k \end{pmatrix} = - \begin{pmatrix} T(\lambda_k)x_k \\ \ell^H x_k - 1 \end{pmatrix}.
\]

Assuming that \( x_k \) is already scaled such that \( \ell^H x_k = 1 \) the second equation reads \( \ell^H x_{k+1} = \ell^H x_k = 1 \), and the first one

\[
x_{k+1} = -(\lambda_{k+1} - \lambda_k) T(\lambda_k)^{-1} T'(\lambda_k)x_k.
\]

Multiplying by \( \ell \) yields

\[
\lambda_{k+1} = \lambda_k - \frac{\ell^H x_k}{\ell^H T(\lambda_k)^{-1} T'(\lambda_k)x_k}.
\]
Inverse iteration ct.

1: start with an approximation $\lambda_1 \in D$ to an eigenvalue of $T(\lambda)x = 0$ and an approximation $x_1$ to an appropriate eigenvector
2: \textbf{for} $k = 1, 2, \ldots$ until convergence \textbf{do}
3: solve $T(\lambda_k)x_{k+1} = T'(\lambda_k)x_k$ for $x_{k+1}$
4: set $\lambda_{k+1} = \lambda_k - \ell^H x_k / \ell^H x_{k+1}$
5: normalize $x_{k+1} \leftarrow x_{k+1} / \|x_{k+1}\|$
6: \textbf{end for}
Inverse iteration ct.

1: start with an approximation $\lambda_1 \in D$ to an eigenvalue of $T(\lambda)x = 0$ and an approximation $x_1$ to an appropriate eigenvector
2: for $k = 1, 2, \ldots$ until convergence do
3: solve $T(\lambda_k)x_{k+1} = T'(\lambda_k)x_k$ for $x_{k+1}$
4: set $\lambda_{k+1} = \lambda_k - \ell^H x_k / \ell^H x_{k+1}$
5: normalize $x_{k+1} \leftarrow x_{k+1} / \|x_{k+1}\|$
6: end for

THEOREM
Let $\hat{\lambda}$ be an eigenvalue of $T(\cdot)$ such that $\mu = 0$ is an algebraically simple eigenvalue of $T'(\hat{\lambda})^{-1} T(\hat{\lambda}) y = \mu y$, and let $\hat{x}$ be a corresponding eigenvector with $\ell^H \hat{x} = 1$. Then the inverse iteration converges locally and quadratically to $(\hat{x}, \hat{\lambda})$. 
Inverse iteration ct.

1. start with an approximation $\lambda_1 \in D$ to an eigenvalue of $T(\lambda)x = 0$ and an approximation $x_1$ to an appropriate eigenvector
2. for $k = 1, 2, \ldots$ until convergence do
3. solve $T(\lambda_k)x_{k+1} = T'(\lambda_k)x_k$ for $x_{k+1}$
4. set $\lambda_{k+1} = \lambda_k - \ell^H x_k / \ell^H x_{k+1}$
5. normalize $x_{k+1} \leftarrow x_{k+1} / \|x_{k+1}\|$  
6. end for

THEOREM
Let $\hat{\lambda}$ be an eigenvalue of $T(\cdot)$ such that $\mu = 0$ is an algebraically simple eigenvalue of $T'(\hat{\lambda})^{-1} T(\hat{\lambda}) y = \mu y$, and let $\hat{x}$ be a corresponding eigenvector with $\ell^H \hat{x} = 1$. Then the inverse iteration converges locally and quadratically to $(\hat{x}, \hat{\lambda})$.

Sketch of proof: Only have to show that the only solution of

$$F'(\hat{x}, \hat{\lambda}) \begin{pmatrix} y \\ \mu \end{pmatrix} = \begin{pmatrix} T(\hat{\lambda}) & T'(\hat{\lambda}) \hat{x} \\ \ell^H & 0 \end{pmatrix} \begin{pmatrix} y \\ \mu \end{pmatrix} = 0$$

is the trivial one.
Rayleigh functional iteration

For linear Hermitian eigenproblems one get’s cubic convergence if the eigenparameter is updated using the Rayleigh quotient. The generalization to nonlinear eigenproblems is the **Rayleigh functional iteration**.
Rayleigh functional iteration

For linear Hermitian eigenproblems one gets cubic convergence if the eigenparameter is updated using the Rayleigh quotient. The generalization to nonlinear eigenproblems is the Rayleigh functional iteration

1: start with an approximation $\lambda_1 \in D$ to an eigenvalue of $T(\lambda)x = 0$ and an approximation $x_1$ to an appropriate eigenvector
2: for $k = 1, 2, \ldots$ until convergence do
3: solve $T(\lambda_k)x_{k+1} = T'(\lambda_k)x_k$ for $x_{k+1}$
4: normalize $x_{k+1} \leftarrow x_{k+1}/\|x_{k+1}\|$ 
5: set $\lambda_{k+1} = p(x_{k+1})$
6: end for
Rayleigh functional iteration

For linear Hermitian eigenproblems one get’s cubic convergence if the eigenparameter is updated using the Rayleigh quotient. The generalization to nonlinear eigenproblems is the Rayleigh functional iteration

1: start with an approximation $\lambda_1 \in D$ to an eigenvalue of $T(\lambda)x = 0$ and an approximation $x_1$ to an appropriate eigenvector
2: for $k = 1, 2, \ldots$ until convergence do
3: solve $T(\lambda_k)x_{k+1} = T'(\lambda_k)x_k$ for $x_{k+1}$
4: normalize $x_{k+1} \leftarrow x_{k+1}/\|x_{k+1}\|$  
5: set $\lambda_{k+1} = p(x_{k+1})$
6: end for

THEOREM (Rothe, V. 1990)
Let $T(\lambda) = (T(\lambda))^H$ for every $\lambda \in J$, and let $\hat{\lambda} \in J$ be an eigenvalue of $T(\cdot)$ such that $\mu = 0$ is an algebraically simple eigenvalue of $T'(\hat{\lambda})^{-1}T(\hat{\lambda})y = \mu y$, and let $\hat{x}$ be a corresponding eigenvector. Then the Rayleigh functional iteration converges locally and cubically to $(\hat{x}, \hat{\lambda})$. 
For the non-Hermitian nonlinear eigenproblem Ruhe (1973) suggested to use

\[ \ell^k = T(\lambda_k)^H y^k \]

for the normalization, where \( y^k \) is an approximation to a left eigenvector.
For the non-Hermitian nonlinear eigenproblem Ruhe (1973) suggested to use 
\[ \ell_k = T(\lambda_k)^H y_k \] 
for the normalization, where \( y^k \) is an approximation to a left eigenvector.

Then the update for \( \lambda \) becomes

\[ \lambda_{k+1} = \lambda_k - \frac{(y^k)^H T(\lambda_k) x^k}{(y^k)^H T'(\lambda_k) x^k}, \]

which is the Rayleigh functional for general nonlinear eigenproblems proposed in Lancaster (1966), and which can be interpreted as one Newton step for solving the equation 
\[ f_k(\lambda) := (y^k)^H T(\lambda) x^k = 0. \]
Two-sided Rayleigh quotient iteration

For a highly nonnormal matrix $A$ it is often advantageous to replace the Rayleigh quotient by Ostrowski’s two-sided Rayleigh quotient (1959)

$$\theta(u, v) = \frac{v^H Au}{v^H u}$$

where $v$ and $u$ denotes an approximate left and right eigenvector of $A$, respectively, and to improve these approximations by solving simultaneously the two linear systems

$$(A - \theta I)\tilde{u} = u \quad \text{and} \quad (A^H - \bar{\theta} I)\tilde{v} = v.$$
Two-sided Rayleigh quotient iteration

For a highly nonnormal matrix $A$ it is often advantageous to replace the Rayleigh quotient by Ostrowski’s two-sided Rayleigh quotient (1959)

$$\theta(u, v) = \frac{v^H A u}{v^H u}$$

where $v$ and $u$ denotes an approximate left and right eigenvector of $A$, respectively, and to improve these approximations by solving simultaneously the two linear systems

$$(A - \theta I)\tilde{u} = u \quad \text{and} \quad (A^H - \bar{\theta} I)\tilde{v} = v.$$ 

Parlett (1979) proved that the two-sided Rayleigh quotient iteration converges cubically to simple eigenvalues.
Two-sided Rayleigh functional iteration

**Require:** Initial triple \((\lambda_0, u_0, v_0)\) with \(u_0^H u_0 = v_0^H v_0 = 1\)

1. **for** \(k = 0, 1, 2, \ldots\) **until** convergence **do**
2. **solve** \(T(\lambda_k) x_{k+1} = T'(\lambda_k) u_k\), set \(u_{k+1} = x_{k+1}/\|x_{k+1}\|\)
3. **solve** \(T(\lambda_k)^H y_{k+1} = T'(\lambda_k)^H v_k\), set \(v_{k+1} = y_{k+1}/\|y_{k+1}\|\)
4. **solve** \(v_{k+1}^H T(\lambda_{k+1}) u_{k+1} = 0\) for \(\lambda_{k+1}\)
5. **end for**
Two-sided Rayleigh functional iteration

**Require:** Initial triple \((\lambda_0, u_0, v_0)\) with \(u_0^H u_0 = v_0^H v_0 = 1\)

1. for \(k = 0, 1, 2, \ldots\) until convergence do
2. solve \(T(\lambda_k)x_{k+1} = T'(\lambda_k)u_k\), set \(u_{k+1} = x_{k+1}/\|x_{k+1}\|\)
3. solve \(T(\lambda_k)^H y_{k+1} = T'(\lambda_k)^H v_k\), set \(v_{k+1} = y_{k+1}/\|y_{k+1}\|\)
4. solve \(v_{k+1}^H T(\lambda_k+1 u_{k+1} = 0\) for \(\lambda_{k+1}\)
5. end for

**THEOREM (Schreiber 2008)**
Let \(\hat{\lambda}\) be an algebraically simple eigenvalue of \(T(\cdot)\) with left and right eigenvector \(\hat{x}\) and \(\hat{y}\), respectively, and let \(T(\lambda)\) be holomorphic on a neighborhood of \(\hat{\lambda}\). Then with \(\hat{x}_k := \hat{x}/u_k^H \hat{x}\) and \(\hat{y}_k := \hat{y}/v_k^H \hat{y}\) it holds for sufficiently good initial approximations \((\lambda_0, u_0, v_0)\) that

\[
\begin{align*}
|\lambda_{k+1} - \hat{\lambda}| & \leq C_0 |\lambda_k - \hat{\lambda}|^2 \|u_k - \hat{x}_k\| \|v_k - \hat{y}_k\| \\
\|u_{k+1} - \hat{x}_k\| & \leq C_1 \|u_k - \hat{x}_k\|^2 \|v_k - \hat{y}_k\| \\
\|v_{k+1} - \hat{y}_k\| & \leq C_1 \|u_k - \hat{x}_k\| \|v_k - \hat{y}_k\|^2
\end{align*}
\]
Osborne (1978) considers Newton’s method for the complex function $\beta(\lambda)$ defined by

$$T(\lambda)u = \beta(\lambda)x, \quad v^H u = \kappa,$$

where $\kappa$ is a given constant, and $x$ and $v$ are given vectors.
Osborne (1978) considers Newton’s method for the complex function $\beta(\lambda)$ defined by

$$T(\lambda)u = \beta(\lambda)x, \quad v^H u = \kappa,$$

where $\kappa$ is a given constant, and $x$ and $v$ are given vectors.

This approach generalizes the method of Kublanovskaya, inverse iteration, and a method proposed in Osborne & Michelson (1964).
Osborne (1978) considers Newton’s method for the complex function $\beta(\lambda)$ defined by

$$T(\lambda)u = \beta(\lambda)x, \quad v^H u = \kappa,$$

where $\kappa$ is a given constant, and $x$ and $v$ are given vectors.

This approach generalizes the method of Kublanovskaya, inverse iteration, and a method proposed in Osborne & Michelson (1964).

It was proved that the rate of convergence is quadratic, and that cubic convergence can be obtained if not only $\lambda$, but also $x$ and/or $v$ are updated appropriately, thus unifying the results in Anselone & Rall (1968), Kublanovskaya (1970), Lancaster (2002), Osborne (1964), and Osborne & Michelson (1964).
The disadvantage of inverse iteration with respect to efficiency is the large number of factorizations that are needed for each of the eigenvalues.
The disadvantage of inverse iteration with respect to efficiency is the large number of factorizations that are needed for each of the eigenvalues.

The obvious idea then is to use a version of a simplified Newton method, where the shift $\sigma$ is kept fixed during the iteration, i.e. to use

$$x_{k+1} = T(\sigma)^{-1} T'(\lambda_k)x_k$$

for some fixed shift $\sigma$. 

[Note: the equations should be rendered in a mathematical typesetting tool to be properly readable and formatted.]
The disadvantage of inverse iteration with respect to efficiency is the large number of factorizations that are needed for each of the eigenvalues.

The obvious idea then is to use a version of a simplified Newton method, where the shift $\sigma$ is kept fixed during the iteration, i.e. to use

$$x_{k+1} = T(\sigma)^{-1} T'(\lambda_k)x_k$$

for some fixed shift $\sigma$.

However, in general this method does not converge in the nonlinear case.
The disadvantage of inverse iteration with respect to efficiency is the large number of factorizations that are needed for each of the eigenvalues.

The obvious idea then is to use a version of a simplified Newton method, where the shift $\sigma$ is kept fixed during the iteration, i.e. to use

$$x_{k+1} = T(\sigma)^{-1} T'(\lambda_k)x_k$$

for some fixed shift $\sigma$.

However, in general this method does not converge in the nonlinear case. The iteration converges to an eigenpair of a linear problem

$$T(\sigma)x = \gamma T'(\tilde{\lambda})x,$$

from which one cannot recover an eigenpair of the nonlinear problem $T(\lambda)x = 0$. 
If $T(\lambda)$ is twice continuously differentiable,

\[
x^k - x^{k+1} = x^k + (\lambda_{k+1} - \lambda_k) T(\lambda_k)^{-1} T'(\lambda_k)x^k
\]

\[
= T(\lambda_k)^{-1}[T(\lambda_k) + (\lambda_{k+1} - \lambda_k) T'(\lambda_k)]x^k
\]

\[
= T(\lambda_k)^{-1} T(\lambda_{k+1})x^k + O((\lambda_{k+1} - \lambda_k)^2)
\]
Residual inverse iteration

If \( T(\lambda) \) is twice continuously differentiable,

\[
x^{k+1} - x^k = x^k + (\lambda_{k+1} - \lambda_k) T(\lambda_k)^{-1} T'(\lambda_k)x^k \\
= T(\lambda_k)^{-1}(T(\lambda_k) + (\lambda_{k+1} - \lambda_k) T'(\lambda_k))x^k \\
= T(\lambda_k)^{-1} T(\lambda_{k+1})x^k + O((\lambda_{k+1} - \lambda_k)^2)
\]

Neglecting second order terms yields the update

\[
x^{k+1} = x^k - T(\lambda_k)^{-1} T(\lambda_{k+1})x^k,
\]

replacing \( \lambda_k \) by a fixed shift \( \sigma \) yields an update

\[
x^{k+1} = x^k - T(\sigma)^{-1} T(\lambda_{k+1})x^k,
\]

without misconvergence.
Residual inverse iteration ct.

1: start with an approximation \( x_1 \in D \) to an eigenvector of \( T(\lambda)x = 0 \)
2: for \( k = 1, 2, \ldots \) until convergence do
3: solve \( \ell^HT(\sigma)^{-1}T(\lambda_{k+1})x^k = 0 \) for \( \lambda_{k+1} \)
   (or \( (x^k)^HT(\lambda_{k+1})x^k = 0 \) if \( T(\cdot) \) is Hermitian)
4: compute the residual \( r^k = T(\lambda_{k+1})x^k \)
5: solve \( T(\sigma)d^k = r^k \)
6: set \( x^{k+1} = x^k - d^k, x^{k+1} = x^{k+1}/\|x^{k+1}\| \)
7: end for
Residual inverse iteration ct.

1: start with an approximation \( x_1 \in D \) to an eigenvector of \( T(\lambda)x = 0 \)
2: for \( k = 1, 2, \ldots \) until convergence do
3: solve \( \ell^H T(\sigma)^{-1} T(\lambda_{k+1})x^k = 0 \) for \( \lambda_{k+1} \)
   (or \( (x^k)^H T(\lambda_{k+1})x^k = 0 \) if \( T(\cdot) \) is Hermitian)
4: compute the residual \( r^k = T(\lambda_{k+1})x^k \)
5: solve \( T(\sigma)d^k = r^k \)
6: set \( x^{k+1} = x^k - d^k, \quad x^{k+1} = x^{k+1}/\|x^{k+1}\| \)
7: end for

THEOREM (Neumaier 1985)
Let \( T(\lambda) \) be twice continuously differentiable. Assume that \( \hat{\lambda} \) is a simple eigenvalue of \( T(\lambda)x = 0 \), and let \( \hat{x} \) be a corresponding eigenvector normalized by \( \|\hat{x}\| = 1 \). Then the residual inverse iteration converges for all \( \sigma \) sufficiently close to \( \hat{\lambda} \), and it holds

\[
\frac{\|x^{k+1} - \hat{x}\|}{\|x^k - \hat{x}\|} = O(|\sigma - \hat{\lambda}|), \quad \text{and} \quad |\lambda_{k+1} - \hat{\lambda}| = O(\|x^k - \hat{x}\|^t).
\]

Here, \( t = 1 \) in the general case, and \( t = 2 \) if \( T(\cdot) \) is Hermitian.
Two-sided residual inverse iteration

Require: \((\lambda_0, u_0, v_0)\) and normalization vectors \(w_x, w_y\) with \(w_x^H u_0 = w_y^H v_0 = 1\).

1: for \(k = 0, 1, 2, \ldots\) until convergence do
2: solve \(v_k^H T(\lambda_{k+1}) u^k = 0\) for \(\lambda_{k+1}\)
3: compute residuals \(r_x^k = T(\lambda_{k+1}) u_k, r_y^k = T(\lambda_{k+1})^H v_k\)
4: solve \(T(\sigma) s_k = r_x^k\) for \(s_k\)
5: solve \(T(\sigma)^H t_k = r_y^k\) for \(t_k\)
6: set \(x_{k+1} = u_k - s^k, u_{k+1} = x_{k+1}/\|x_{k+1}\|\)
7: set \(y_{k+1} = v_k - t^k, v_{k+1} = y_{k+1}/\|y_{k+1}\|\)
8: end for
Two-sided residual inverse iteration

**Require:** $(\lambda_0, u_0, v_0)$ and normalization vectors $w_x, w_y$ with $w_x^H u_0 = w_y^H v_0 = 1$.

1. for $k = 0, 1, 2, \ldots$ until convergence do
   2. solve $v_k^H T(\lambda_{k+1}) u^k = 0$ for $\lambda_{k+1}$
   3. compute residuals $r^x_k = T(\lambda_{k+1}) u_k$, $r^y_k = T(\lambda_{k+1})^H v_k$
   4. solve $T(\sigma) s_k = r^x_k$ for $s_k$
   5. solve $T(\sigma)^H t_k = r^y_k$ for $t_k$
   6. set $x_{k+1} = u_k - s^k_k$, $u_{k+1} = x_{k+1}/\|x_{k+1}\|$ 
   7. set $y_{k+1} = v_k - t^k_k$, $v_{k+1} = y_{k+1}/\|y_{k+1}\|$
   8. end for

**THEOREM (Schreiber 2008)**

Let $\hat{\lambda}$ be an algebraically simple eigenvalue of $T(\cdot)$ with left and right eigenvector $\hat{x}$ and $\hat{y}$ scaled by $w_x^H \hat{x} = 1$, $w_y^H \hat{y} = 1$, respectively, and let $T(\lambda)$ be holomorphic on a neighborhood of $\hat{\lambda}$. Then it holds for sufficiently good initial approximations $(\lambda_0, u_0, v_0)$ that

$$\frac{|\lambda_{k+1} - \hat{\lambda}|}{\|u_{k+1} - \hat{x}\|/\|u_k - \hat{x}\|} = O(|\sigma - \hat{\lambda}|)$$

$$\frac{|(\lambda_{k+1} - \hat{\lambda})^{-1}\|u_k - \hat{x}\|\|v_k - \hat{y}\|}{\|u_k - \hat{x}\|/\|u_k - \hat{x}\|} = O(|\sigma - \hat{\lambda}|)$$

$$\frac{|\lambda_{k+1} - \hat{\lambda}|}{\|v_{k+1} - \hat{y}\|/\|v_k - \hat{y}\|} = O(|\sigma - \hat{\lambda}|)$$
If the shift is updated as well one gets

**Require:** \((\lambda_0, u_0, v_0)\) and normalization vectors \(w_x, w_y\) with \(w_x^H u_0 = w_y^H v_0 = 1\).

1. **for** \(k = 0, 1, 2, \ldots\) **until** convergence **do**
2. solve \(v_k^H T(\lambda_{k+1}) u^k = 0\) for \(\lambda_{k+1}\)
3. compute residuals \(r^x_k = T(\lambda_{k+1}) u_k, r^y_k = T(\lambda_{k+1})^H v_k\)
4. solve \(T(\lambda_{k+1}) s_k = r^x_k\) for \(s_k\)
5. solve \(T(\lambda_{k+1})^H t_k = r^y_k\) for \(t_k\)
6. set \(x_{k+1} = u_k - s_k, u_{k+1} = x_{k+1}/\|x_{k+1}\|\)
7. set \(y_{k+1} = v_k - t_k, v_{k+1} = y_{k+1}/\|y_{k+1}\|\)
8. **end for**

**THEOREM** (Schreiber 2008)
Under the conditions of the last **THEOREM** the R-order of convergence of the two-sided residual inverse iteration is at least two, and it holds that

\[
\|\lambda_{k+1} - \hat{\lambda}\| = O\left(\|u_k - \hat{x}\|\|v_k - \hat{y}\|\right)
\]

\[
\|u_{k+1} - \hat{x}\| = O\left(\|u_k - \hat{x}\|\|u_k - 1 - \hat{x}\|\|v_k - 1 - \hat{y}\|\right)
\]

\[
\|v_{k+1} - \hat{y}\| = O\left(\|v_k - \hat{y}\|\|u_k - 1 - \hat{x}\|\|v_k - 1 - \hat{y}\|\right)
\]
Two-sided residual inverse iteration ct.

If the shift is updated as well one gets

**Require:** $(\lambda_0, u_0, v_0)$ and normalization vectors $w_x, w_y$ with $w_x^H u_0 = w_y^H v_0 = 1$.

1. **for** $k = 0, 1, 2, \ldots$ **until** convergence **do**
   2. solve $v_k^H T(\lambda_{k+1}) u^k = 0$ for $\lambda_{k+1}$
   3. compute residuals $r^x_k = T(\lambda_{k+1}) u^k$, $r^y_k = T(\lambda_{k+1})^H v_k$
   4. solve $T(\lambda_{k+1}) s_k = r^x_k$ for $s_k$
   5. solve $T(\lambda_{k+1})^H t_k = r^y_k$ for $t_k$
   6. set $x_{k+1} = u_k - s^k$, $u_{k+1} = x_{k+1}/\|x_{k+1}\|$ 
   7. set $y_{k+1} = v_k - t^k$, $v_{k+1} = y_{k+1}/\|y_{k+1}\|$
   8. **end for**

**THEOREM** (Schreiber 2008)

Under the conditions of the last **THEOREM** the R-order of convergence of the two-sided residual inverse iteration is at least two, and it holds that

\[
|\lambda_{k+1} - \hat{\lambda}| = O(\|u_k - \hat{x}\|\|v_k - \hat{y}\|)
\]

\[
\|u_{k+1} - \hat{x}\| = O(|\|u_k - \hat{x}\|\|u_{k-1} - \hat{x}\|\|v_{k-1} - \hat{y}\|)
\]

\[
\|v_{k+1} - \hat{y}\| = O(|\|v_k - \hat{y}\|\|v_{k-1} - \hat{y}\|\|u_{k-1} - \hat{x}\|)
\]
Outline

1. Examples
2. Basic Properties
3. Rayleigh Functional
4. Methods for dense nonlinear eigenvalue problems
   - Linearization
   - Methods based on characteristic equation
   - Newton’s method
5. Invariant pairs
One of the most fundamental differences from the linear case is that distinct eigenvalues may have linear dependent eigenvectors or even share the same eigenvalue.
One of the most fundamental differences from the linear case is that distinct eigenvalues may have linear dependent eigenvectors or even share the same eigenvalue.

This is a severe hindrance in the development of general numerical schemes for computing several eigenvalues of a nonlinear eigenvalue problem, either simultaneously or subsequently. The usual deflation techniques for linear eigenproblems do not work.
One of the most fundamental differences from the linear case is that distinct eigenvalues may have linear dependent eigenvectors or even share the same eigenvalue.

This is a severe hindrance in the development of general numerical schemes for computing several eigenvalues of a nonlinear eigenvalue problem, either simultaneously or subsequently. The usual deflation techniques for linear eigenproblems do not work.

In the following we show that the concept of invariant pairs offers a way of representing an eigenvalue problem that is insensitive to this phenomenon.
One of the most fundamental differences from the linear case is that distinct eigenvalues may have linear dependent eigenvectors or even share the same eigenvalue.

This is a severe hindrance in the development of general numerical schemes for computing several eigenvalues of a nonlinear eigenvalue problem, either simultaneously or subsequently. The usual deflation techniques for linear eigenproblems do not work.

In the following we show that the concept of invariant pairs offers a way of representing an eigenvalue problem that is insensitive to this phenomenon.

It is convenient to assume that the eigenvalue problem assumes the following form

\[ T(\lambda)x = \sum_{j=1}^{m} f_j(\lambda)A_jx = 0 \]

for holomorphic functions \( f_j : \mathbb{C} \ni \Omega \rightarrow \mathbb{C} \) and constant matrices \( A_j \in \mathbb{C}^{n\times n} \).
**Definition** Let the eigenvalues of \( S \in \mathbb{C}^{k \times k} \) be contained in \( \Omega \subset \mathbb{C} \) and let \( X \in \mathbb{C}^{n \times k} \). Then \((X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}\) is called an invariant pair of the nonlinear eigenvalue problem (1) if

\[
\sum_{j=1}^{m} A_j X f_j(S) = 0. \tag{2}
\]
**Definition** Let the eigenvalues of \( S \in \mathbb{C}^{k \times k} \) be contained in \( \Omega \subset \mathbb{C} \) and let \( X \in \mathbb{C}^{n \times k} \). Then \((X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}\) is called an invariant pair of the nonlinear eigenvalue problem (1) if

\[
\sum_{j=1}^{m} A_j X f_j(S) = 0. \tag{2}
\]

**Example 1** For the quadratic eigenvalue problem \( T(\lambda)x = 0 \) with

\[
T(\lambda) := \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} + \lambda \begin{bmatrix} 7 & -5 \\ 10 & -8 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3}
\]

with distinct eigenvalues \( \lambda = 1 \) and \( \lambda = 2 \) and shared eigenvector \( x := \begin{bmatrix} 1 \\ 2 \end{bmatrix} \)

\[
X = \begin{bmatrix} x & x \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}
\]

is an invariant pair.
To avoid the trivial pairs, such as $X = 0$, an additional property has to be imposed.
To avoid the trivial pairs, such as $X = 0$, an additional property has to be imposed.

We have already seen that requiring $X$ to have full column rank is not reasonable in the context of nonlinear eigenvalue problems.
To avoid the trivial pairs, such as $X = 0$, an additional property has to be imposed.

We have already seen that requiring $X$ to have full column rank is not reasonable in the context of nonlinear eigenvalue problems.

**Definition** A pair $(X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}$ is called **minimal** if there is $\ell \in \mathbb{N}$ such that the matrix

$$V_\ell(X, S) := \begin{bmatrix} X \\ XS \\ \vdots \\ XS^{\ell-1} \end{bmatrix}$$

has rank $k$. 
To avoid the trivial pairs, such as $X = 0$, an additional property has to be imposed.

We have already seen that requiring $X$ to have full column rank is not reasonable in the context of nonlinear eigenvalue problems.

**Definition** A pair $(X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}$ is called **minimal** if there is $\ell \in \mathbb{N}$ such that the matrix

$$V_\ell(X, S) := \begin{bmatrix} X \\ XS \\ \vdots \\ XS^{\ell-1} \end{bmatrix}$$

has rank $k$.

The smallest such $\ell$ is called the **minimality index** of $(X, S)$.
To avoid the trivial pairs, such as $X = 0$, an additional property has to be imposed.

We have already seen that requiring $X$ to have full column rank is not reasonable in the context of nonlinear eigenvalue problems.

**Definition** A pair $(X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}$ is called **minimal** if there is $\ell \in \mathbb{N}$ such that the matrix

$$V_\ell(X, S) := \begin{bmatrix} X \\ XS \\ \vdots \\ XS^{\ell-1} \end{bmatrix}$$

has rank $k$.

The smallest such $\ell$ is called the **minimality index** of $(X, S)$.

An invariant pair $(X, S)$ is called **simple**, if $(X, S)$ is minimal and the algebraic multiplicities of the eigenvalues of $S$ are identical to the ones of the corresponding eigenvalues of $T(\cdot)$. 
For the quadratic eigenvalue problem (3)

\[ T(\lambda) := \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} + \lambda \begin{bmatrix} 7 & -5 \\ 10 & -8 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

it holds

\[ \det T(\lambda) = \lambda^4 - \lambda^3 - 3\lambda^2 + \lambda + 2. \]
For the quadratic eigenvalue problem (3)

\[ T(\lambda) := \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} + \lambda \begin{bmatrix} 7 & -5 \\ 10 & -8 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

it holds

\[ \det T(\lambda) = \lambda^4 - \lambda^3 - 3\lambda^2 + \lambda + 2. \]

\[ \hat{\lambda} = -1 \text{ is an eigenvalue with eigenvector } \hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \]
For the quadratic eigenvalue problem (3)

\[ T(\lambda) := \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} + \lambda \begin{bmatrix} 7 & -5 \\ 10 & -8 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

it holds

\[ \det T(\lambda) = \lambda^4 - \lambda^3 - 3\lambda^2 + \lambda + 2. \]

\( \hat{\lambda} = -1 \) is an eigenvalue with eigenvector \( \hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \).

The pair \((\hat{x}, \hat{\lambda})\) is a minimal invariant pair with minimality index 1.
For the quadratic eigenvalue problem (3)

\[ T(\lambda) := \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} + \lambda \begin{bmatrix} 7 & -5 \\ 10 & -8 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]

it holds

\[ \det T(\lambda) = \lambda^4 - \lambda^3 - 3\lambda^2 + \lambda + 2. \]

\[ \hat{\lambda} = -1 \] is an eigenvalue with eigenvector \( \hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \).

The pair \((\hat{x}, \hat{\lambda})\) is a minimal invariant pair with minimality index 1.

It is not simple, because the algebraic multiplicity of \( \hat{\lambda} \) is 2 as an eigenvalue of \( T(\lambda) \) but only 1 as an eigenvalue of \( S = [-1] \).
The pair \((X_1, S_1)\) with \(X_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\) and \(S_1 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}\) with

\[
\begin{bmatrix}
X_1 \\
X_1 S_1
\end{bmatrix} = \begin{bmatrix}
1 & 1 \\
1 & 1 \\
-1 & 0 \\
-1 & 0
\end{bmatrix}
\]

is a minimal invariant pair with minimality index 2, which is simple.
The pair \((X_1, S_1)\) with \(X_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\) and \(S_1 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}\) with

\[
\begin{bmatrix} X_1 \\ X_1 S_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 0 \\ -1 & 0 \end{bmatrix}
\]

is a minimal invariant pair with minimality index 2, which is simple.

The same is true for the pairs \((X_2, S_2)\) with \(X_2 = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}\) and \(S_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\)
The pair \((X_1, S_1)\) with \(X_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\) and \(S_1 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}\) with

\[
\begin{bmatrix} X_1 \\ X_1 S_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 0 \\ -1 & 0 \end{bmatrix}
\]

is a minimal invariant pair with minimality index 2, which is simple.

The same is true for the pairs \((X_2, S_2)\) with \(X_2 = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}\) and \(S_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\)

and for \((X_3, S_3)\) with \(X_3 := [X_1, X_2]\) and \(S_3 := \text{diag}(X_1, X_2)\).
Theorem
Let \((X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}\) be a minimal invariant pair of \(T(\cdot)\). Then the following statements hold:

1. For any invertible matrix \(Z \in \mathbb{C}^{k \times x}\) the pair \((XZ, Z^{-1}SZ)\) is also a minimal invariant pair.
2. The eigenvalues of \(S\) are eigenvalues of \(T(\cdot)\).
3. The minimality index cannot exceed \(k\).
Theorem
Let \((X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}\) be a minimal invariant pair of \(T(\cdot)\). Then the following statements hold:

1. For any invertible matrix \(Z \in \mathbb{C}^{k \times x}\) the pair \((XZ, Z^{-1}SZ)\) is also a minimal invariant pair.
Theorem

Let \((X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}\) be a minimal invariant pair of \(T(\cdot)\). Then the following statements hold:

1. For any invertible matrix \(Z \in \mathbb{C}^{k \times x}\) the pair \((XZ, Z^{-1}SZ)\) is also a minimal invariant pair.

2. The eigenvalues of \(S\) are eigenvalues of \(T(\cdot)\).
Theorem
Let \((X, S) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}\) be a minimal invariant pair of \(T(\cdot)\). Then the following statements hold:

1. For any invertible matrix \(Z \in \mathbb{C}^{k \times x}\) the pair \((XZ, Z^{-1}SZ)\) is also a minimal invariant pair.

2. The eigenvalues of \(S\) are eigenvalues of \(T(\cdot)\).

3. The minimality index can not exceed \(k\).
Simple invariant pairs are well posed objects in the sense of being regular solutions of nonlinear matrix equations.
Simple invariant pairs are well posed objects in the sense of being regular solutions of nonlinear matrix equations.

Consider the nonlinear matrix operator

\[ T : \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}_\Omega \rightarrow \mathbb{C}^{n \times k} \]

\[ (X, S) \mapsto \sum_{j=1}^m A_j X f_j(S) \]  

(4)

where \( \mathbb{C}^{k \times k}_\Omega \) denotes the set of \( k \times k \) matrices with eigenvalues in \( \Omega \).

Then an invariant pair \((X, S)\) satisfies \( T(X, S) = 0 \), but this relation is not sufficient to characterize \((X, S)\).
Simple invariant pairs are well posed objects in the sense of being regular solutions of nonlinear matrix equations.

Consider the nonlinear matrix operator

\[
\mathcal{T} : \begin{cases}
\mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}_{\Omega} \to \mathbb{C}^{n \times k} \\
(X, S) \mapsto \sum_{j=1}^{m} A_j X f_j(S)
\end{cases}
\]

where \(\mathbb{C}^{k \times k}_{\Omega}\) denotes the set of \(k \times k\) matrices with eigenvalues in \(\Omega\).

Then an invariant pair \((X, S)\) satisfies \(\mathcal{T}(X, S) = 0\), but this relation is not sufficient to characterize \((X, S)\).

To define a scaling condition, choose \(\ell\) such that the matrix \(V_\ell(X, S)\) has rank \(k\), and define the partition

\[
W = \begin{bmatrix}
W_0 \\
W_1 \\
\vdots \\
W_{\ell-1}
\end{bmatrix} := V_\ell(X, S) \left( V_\ell(X, S)^H V_\ell(X, S) \right)^{-1} \in \mathbb{C}^{nk \times k}
\]

with \(W_j \in \mathbb{C}^{n \times k}\). Then it holds \(\nabla(X, S) = 0\) for the operator

\[
\nabla : \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}_{\Omega} \to \mathbb{C}^{n \times k}, \quad \nabla(X, S) := W^H V_\ell(X, S) - I_k.
\]
Invariant pairs

Theorem
If \((X, S)\) is a minimal invariant pair for the nonlinear eigenvalue problem \(T(\cdot)\), then \((X, S)\) is simple if and only if the linear matrix operator

\[
L : \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k} \rightarrow \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}, (\Delta X, \Delta S) \mapsto (\mathcal{D}T(\Delta X, \Delta S), \mathcal{D}V(\Delta X, \Delta S))
\]

is invertible, where \(\mathcal{D}T\) and \(\mathcal{D}V\) denotes the Fréchet derivative of \(T\) and \(V\), respectively.
Theorem
If \((X, S)\) is a minimal invariant pair for the nonlinear eigenvalue problem \(T(\cdot)\), then \((X, S)\) is simple if and only if the linear matrix operator
\[
L : \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k} \to \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}, (\Delta X, \Delta S) \mapsto (DT(\Delta X, \Delta S), DV(\Delta X, \Delta S))
\]
is invertible, where \(DT\) and \(DV\) denotes the Fréchet derivative of \(T\) and \(V\), respectively.

This theorem motivates to apply Newton’s method to the system \(T(X, S) = 0, V(X, S) = 0\) which can be written as
\[
(X_{p+1}, S_{p+1}) = (X_p, S_p) - L^{-1}(T(X_p, S_p), V(X_p, S_p))
\]
where \(L = (DT, DV)\) is the Jacobian matrix of \(T(X, S) = 0, V(X, S) = 0\).
**Theorem**

If \((X, S)\) is a minimal invariant pair for the nonlinear eigenvalue problem \(T(\cdot)\), then \((X, S)\) is simple if and only if the linear matrix operator

\[
L : \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k} \to \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}, (\Delta X, \Delta S) \mapsto (\mathcal{D}T(\Delta X, \Delta S), \mathcal{D}V(\Delta X, \Delta S))
\]

is invertible, where \(\mathcal{D}T\) and \(\mathcal{D}V\) denotes the Fréchet derivative of \(T\) and \(V\), respectively.

This theorem motivates to apply Newton’s method to the system \(T(X, S) = 0, V(X, S) = 0\) which can be written as

\[
(X_{p+1}, S_{p+1}) = (X_p, S_p) - L^{-1}(T(X_p, S_p), V(X_p, S_p))
\]

where \(L = (\mathcal{D}T, \mathcal{D}V)\) is the Jacobian matrix of \(T(X, S) = 0, V(X, S) = 0\).

\[
\mathcal{D}T(\Delta X, \Delta S) = T(\Delta X, S) + \sum_{j=1}^{m} A_j X[Df_j(S)](\Delta S),
\]

\[
\mathcal{D}V(\Delta X, \Delta S) = W_0^H \Delta X + \sum_{j=1}^{m} W_j^H (\Delta XS^j + X[D S^j](\Delta S)).
\]
Newton’s method for computing invariant pairs

**Require:** Initial pair \((X_0, S_0) \in \mathbb{C}^{n \times k} \times \mathbb{C}^{k \times k}\) such that 
\[V_\ell(X_0, S_0)^H V_\ell(X_0, S_0) = I_k\]

1: \(p \leftarrow 0, W \leftarrow V_\ell(X_0, S_0)\)
2: **repeat**
3: \(\text{Res} \leftarrow \mathbb{T}(X_p, S_p)\)
4: Solve linear matrix equation \(\mathbb{L}(\Delta X, \Delta S) = (\text{Res}, O)\)
5: \(\tilde{X}_{p+1} \leftarrow X_p - \Delta X, \tilde{S}_{p+1} \leftarrow S_p - \Delta S\)
6: Compute compact QR decomposition \(V_\ell(\tilde{X}_{p+1}, \tilde{S}_{p+1}) = WR\)
7: \(X_{p+1} \leftarrow \tilde{X}_{p+1} R^{-1}, S_{p+1} \leftarrow R\tilde{S}_{p+1} R^{-1}\)
8: **until** convergence
Example (by courtesy of Daniel Kressner)

\[ T(\lambda) = \lambda I - A_0 - A_1 e^{-\tau \lambda} = \lambda I - \begin{bmatrix} -5 & 1 \\ 2 & -6 \end{bmatrix} - \begin{bmatrix} -2 & 1 \\ 4 & -1 \end{bmatrix} e^{-\lambda} \]
Example (by courtesy of Daniel Kressner)

\[ T(\lambda) = \lambda I - A_0 - A_1 e^{-\tau \lambda} = \lambda I - \begin{bmatrix} -5 & 1 \\ 2 & -6 \end{bmatrix} - \begin{bmatrix} -2 & 1 \\ 4 & -1 \end{bmatrix} e^{-\lambda} \]

For the stability analysis of the corresponding delay differential equation

\[ \dot{x}(t) = A_0 x(t) + A_1 x(t - \tau), \]

it is of interest to compute eigenvalues with largest real part.
Example (by courtesy of Daniel Kressner)

\[ T(\lambda) = \lambda I - A_0 - A_1 e^{-\tau \lambda} = \lambda I - \begin{bmatrix} -5 & 1 \\ 2 & -6 \end{bmatrix} - \begin{bmatrix} -2 & 1 \\ 4 & -1 \end{bmatrix} e^{-\lambda} \]

For the stability analysis of the corresponding delay differential equation

\[ \dot{x}(t) = A_0 x(t) + A_1 x(t - \tau), \]

it is of interest to compute eigenvalues with largest real part.

To obtain an initial guess, one approximates \( T(\cdot) \) by a polynomial

\[ T(\lambda) \approx P(\lambda) = \lambda I - A_0 - A_1 \sum_{i=0}^{j} \frac{1}{i!} (-\lambda \tau)^i \]

and compute \( k \) eigenvalues \( \lambda_1, \ldots, \lambda_k \) with largest real part (e.g. by linearization).
Example 1

As initial value one chooses $S_0 = \text{diag}\{\lambda_1, \ldots, \lambda_k\}$ and computes $X_0$ by block-inverse iteration (cf. Kressner 2009).
Example 1

As initial value one chooses $S_0 = \text{diag}\{\lambda_1, \ldots, \lambda_k\}$ and computes $X_0$ by block-inverse iteration (cf. Kressner 2009).

**Require:** Initial matrix $S_0 \in \mathbb{C}^{k \times k}$

**Require:** Random matrix $X_0 \in \mathbb{C}^{n \times k}$

1: repeat
2: Compute solution $Y$ of the linear equation $T(Y, S_0) = x_0$
3: Compute compact QR decomposition $V_\ell(Y, S_0) = WR$
4: Update $X_0 \leftarrow YR^{-1}$, $S_0 \leftarrow RS_0R^{-1}$
5: until convergence
Example 1 ct.

The following figures contain the results obtained with the block–Newton method with $k = 5$ and $\ell = 3$. For the initial approximation $j = 4$ was chosen.
The following figures contain the results obtained with the block–Newton method with $k = 5$ and $\ell = 3$. For the initial approximation $j = 4$ was chosen.

Initially, three eigenvalues are well and two eigenvalues are poorly approximated.
Example 1 ct.

The following figures contain the results obtained with the block–Newton method with $k = 5$ and $\ell = 3$. For the initial approximation $j = 4$ was chosen.

Initially, three eigenvalues are well and two eigenvalues are poorly approximated.

The first Fig. shows the location of eigenvalue approximations initially (plus), after 3 iterations (circle), and at convergence after 12 iterations (cross).

The second Fig. shows the residual norm in the course of the Newton iteration.
The following figures contain the results obtained with the block–Newton method with $k = 5$ and $\ell = 3$. For the initial approximation $j = 4$ was chosen.

Initially, three eigenvalues are well and two eigenvalues are poorly approximated.

The first Fig. shows the location of eigenvalue approximations initially (plus), after 3 iterations (circle), and at convergence after 12 iterations (cross).

The second Fig. shows the residual norm in the course of the Newton iteration.

The condition number of the Jacobian is $9.2e5$ at convergence, which could explain the poor transient behavior of Newton’s method.
Location of eigenvalue approximations initially (plus), after 3 iterations (circle), and at convergence after 12 iterations (cross).
Residual norm in the course of the Newton iteration.
Example 2

**Example** (by courtesy of Daniel Kressner)
The experiments from the last example are repeated for the matrices

\[
A_0 = \begin{bmatrix}
-0.8498 & 0.1479 & 44.37 \\
0.003756 & -0.2805 & -229.2 \\
-0.1754 & 0.02296 & -0.3608 \\
\end{bmatrix},
A_1 = \begin{bmatrix}
0.28 & 0 & 0 \\
0 & -0.28 & 0 \\
0 & 0 & 0 \\
\end{bmatrix},
\]

and \( \tau = 1 \), which has – according to Green & Wagenknecht (2006) – applications in the stability analysis of a semiconductor laser subject to external feedback.
Example (by courtesy of Daniel Kressner)
The experiments from the last example are repeated for the matrices

\[
A_0 = \begin{bmatrix}
-0.8498 & 0.1479 & 44.37 \\
0.003756 & -0.2805 & -229.2 \\
-0.1754 & 0.02296 & -0.3608
\end{bmatrix}, \quad A_1 = \begin{bmatrix}
0.28 & 0 & 0 \\
0 & -0.28 & 0 \\
0 & 0 & 0
\end{bmatrix},
\]

and \( \tau = 1 \), which has – according to Green & Wagenknecht (2006) – applications in the stability analysis of a semiconductor laser subject to external feedback.

This time, \( j = 2 \) in the polynomial approximation, and \( k = 4, \ell = 2 \).
Example (by courtesy of Daniel Kressner)
The experiments from the last example are repeated for the matrices

\[
A_0 = \begin{bmatrix}
-0.8498 & 0.1479 & 44.37 \\
0.003756 & -0.2805 & -229.2 \\
-0.1754 & 0.02296 & -0.3608
\end{bmatrix}, \quad A_1 = \begin{bmatrix}
0.28 & 0 & 0 \\
0 & -0.28 & 0 \\
0 & 0 & 0
\end{bmatrix},
\]

and \( \tau = 1 \), which has – according to Green & Wagenknecht (2006) – applications in the stability analysis of a semiconductor laser subject to external feedback.

This time, \( j = 2 \) in the polynomial approximation, and \( k = 4, \ell = 2 \).

The condition number of the Jacobian is 2.2e6 at convergence.
Invariant pairs

Eigenvalue approximations

Location of eigenvalue approximations initially (plus), after 3 iterations (circle), and at convergence after 12 iterations (cross).
Residual norm in the course of the Newton iteration.
Invariant pairs

References


K. Schreiber: Nonlinear Eigenvalue Problems: Newton-type Methods and Nonlinear Rayleigh Functionals. VDM Verlag Dr. Müller 2008

