

# On Chebyshev Polynomials of Matrices

Petr Tichý

joint work with

Vance Faber and Jörg Liesen

Institute of Computer Science,  
Academy of Sciences of the Czech Republic

May 24-28, 2010, Novi Sad  
Applied Linear Algebra (ALA) 2010

# Chebyshev polynomials of a compact set

- Chebyshev polynomials on the interval  $[-1; 1]$  [Chebyshev 1859].
- Generalized by [Georg Faber 1920] to the idea of the Chebyshev polynomials of  $\Omega$ , where  $\Omega$  is a compact set in the complex plane  $\mathbb{C}$ : These polynomials  $T_m^\Omega(z)$  solve the problem

$$\min_{p \in \mathcal{M}_m(\Omega)} \|p(z)\|_\infty$$

where  $\mathcal{M}_m$  is the class of **monic polynomials** of degree  $m$ .

## Example:

$\Omega$  is an interval, a set of discrete points, the unit circle, etc.

# Chebyshev polynomials of normal matrices

Let  $\mathbf{A} \in \mathbb{C}^{n \times n}$  be normal, i.e.,

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*, \quad \mathbf{Q}^*\mathbf{Q} = \mathbf{I}.$$

Let  $\|\cdot\|$  be the spectral norm and consider the problem

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|.$$

Then

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\| = \min_{p \in \mathcal{M}_m} \|p(\mathbf{\Lambda})\| = \min_{p \in \mathcal{M}_m(\Omega)} \|p(z)\|_\infty$$

where  $\Omega = \{\lambda_1, \dots, \lambda_n\}$ .

The problem for  $\mathbf{A}$  is solved by the Chebyshev polynomial of  $\Omega$ .

# Chebyshev polynomials of general matrices

Let  $\mathbf{A} \in \mathbb{C}^{n \times n}$  be a general matrix. We consider the problem

$$\min_{p \in \mathcal{M}_m} \| p(\mathbf{A}) \| .$$

- Introduced in [Greenbaum, Trefethen 1994].
- Unique solution  $T_m^{\mathbf{A}}(z) \in \mathcal{M}_m$  exists if  $m < d(\mathbf{A})$ , [Greenbaum, Trefethen 1994; Liesen, T. 2009].
- $T_m^{\mathbf{A}}(z)$  is called the  $m$ th Chebyshev polynomial of  $\mathbf{A}$ , or the  $m$ th ideal Arnoldi polynomial of  $\mathbf{A}$ .
- Previous work on these polynomials in [Toh PhD thesis 1996], [Toh, Trefethen 1998], [Trefethen, Embree 2005].
- Here: [Faber, Liesen, T. 2010].

# Motivation

[Toh, Trefethen 1998] „Chebyshev polynomials of matrices are never far away from any discussion of convergence of Krylov subspace iterations in numerical linear algebra”.

GMRES and Arnoldi approximation problems:

[Greenbaum, Trefethen 1994]

$$\min_{p \in \pi_m} \|p(\mathbf{A})b\| \quad (\text{GMRES}),$$

$$b \approx \{\mathbf{A}b, \dots, \mathbf{A}^m b\},$$

$$\min_{q \in \mathcal{M}_m} \|q(\mathbf{A})b\| \quad (\text{Arnoldi}),$$

$$\mathbf{A}^m b \approx \{b, \dots, \mathbf{A}^{m-1}b\}.$$

One may remove  $b$  from the discussion and pose the following “ideal” approximation problems:

$$\min_{p \in \pi_m} \|p(\mathbf{A})\| \quad (\text{Ideal GMRES}),$$

$$I \approx \{\mathbf{A}, \dots, \mathbf{A}^m\},$$

$$\min_{q \in \mathcal{M}_m} \|q(\mathbf{A})\| \quad (\text{Ideal Arnoldi}),$$

$$\mathbf{A}^m \approx \{I, \dots, \mathbf{A}^{m-1}\}$$

(Chebyshev polynomial of  $\mathbf{A}$ )

# Motivation Example

Let  $\lambda \in \mathbb{C}$ . Consider an  $n$  by  $n$  Jordan block

$$\mathbf{J}_\lambda = \begin{bmatrix} \lambda & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix} \in \mathbb{C}^{n \times n}.$$

**Question:** How do the ideal GMRES and Chebyshev polynomials of  $\mathbf{J}_\lambda$  look like?

- Ideal GMRES polynomial of  $\mathbf{J}_\lambda$  - a very difficult problem [T., Liesen, Faber 2007].
- Chebyshev polynomial of  $\mathbf{J}_\lambda$  [Liesen, T. 2009]:

$$T_m^{\mathbf{J}_\lambda}(z) = (z - \lambda)^m.$$

# Outline

- 1 General results
- 2 Matrices and sets in the complex plane

## Theorem

[Faber, Liesen, T. 2010]

For  $\mathbf{A} \in \mathbb{C}^{n \times n}$  and  $\alpha \in \mathbb{C}$  the following hold:

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A} + \alpha \mathbf{I})\| = \min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|,$$

$$\min_{p \in \mathcal{M}_m} \|p(\alpha \mathbf{A})\| = |\alpha|^m \min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|.$$

- Shift invariance: Not surprising, because the polynomials are normalized at infinity.
- Paper contains explicit relations between the coefficients of  $T_m^{\mathbf{A}}(z)$ ,  $T_m^{\mathbf{A} + \alpha \mathbf{I}}(z)$ , and  $T_m^{\alpha \mathbf{A}}(z)$ .



## Example - shift of a matrix

Let  $a, b \in \mathbb{R}$  be given. Consider the block-diagonal matrix  $\mathbf{A}$  with two  $n \times n$  Jordan blocks,

$$\mathbf{A} \equiv \begin{bmatrix} \mathbf{J}_a & 0 \\ 0 & \mathbf{J}_b \end{bmatrix}.$$

Define

$$\alpha \equiv \frac{a+b}{2}.$$

Then

$$\mathbf{A} - \alpha\mathbf{I} = \begin{bmatrix} \mathbf{J}_\lambda & 0 \\ 0 & \mathbf{J}_{-\lambda} \end{bmatrix} \quad \text{where} \quad \lambda \equiv \frac{a-b}{2},$$

and the previous theorem implies

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\| = \min_{p \in \mathcal{M}_m} \|p(\mathbf{A} - \alpha\mathbf{I})\|.$$

## Symmetry with respect to the origin

The Chebyshev polynomials of real intervals that are symmetric with respect to the origin are alternating between *even* and *odd*, i.e.

$$T_m^{[-a,a]}(z) = (-1)^m T_m^{[-a,a]}(-z).$$

Analogous result for Chebyshev polynomials of  $\mathbf{A}$ ?

### Theorem

[Faber, Liesen, T. 2010]

Let  $\mathbf{A} \in \mathbb{C}^{n \times n}$  and a positive integer  $m < d(\mathbf{A})$  be given. If there exists a unitary matrix  $\mathbf{P}$  such that either

$$\mathbf{P}^* \mathbf{A} \mathbf{P} = -\mathbf{A} \quad \text{or} \quad \mathbf{P}^* \mathbf{A} \mathbf{P} = -\mathbf{A}^T,$$

then

$$T_m^{\mathbf{A}}(z) = (-1)^m T_m^{\mathbf{A}}(-z).$$

## Example

$$\mathbf{A} = \begin{bmatrix} \mathbf{J}_\lambda & \\ & \mathbf{J}_{-\lambda} \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} & \mathbf{I}^\pm \\ \mathbf{I}^\pm & \end{bmatrix},$$

where  $\mathbf{I}^\pm = \text{diag}(1, -1, 1, \dots, (-1)^{n-1})$ . Then

$$\mathbf{J}_{-\lambda} = -\mathbf{I}^\pm \mathbf{J}_\lambda \mathbf{I}^\pm \Rightarrow \mathbf{P}^* \mathbf{A} \mathbf{P} = -\mathbf{A}.$$

Moreover

$$T_m^{\mathbf{A}}(\mathbf{A}) = \begin{bmatrix} T_m^{\mathbf{A}}(\mathbf{J}_\lambda) & \\ & T_m^{\mathbf{A}}(\mathbf{J}_{-\lambda}) \end{bmatrix},$$

and

$$\| T_m^{\mathbf{A}}(\mathbf{J}_{-\lambda}) \| = \| \mathbf{I}^\pm T_m^{\mathbf{A}}(-\mathbf{J}_\lambda) \mathbf{I}^\pm \| = \| T_m^{\mathbf{A}}(\mathbf{J}_\lambda) \|,$$

i.e., the Chebyshev polynomial of  $\mathbf{A}$  attains the same norm on each of the two diagonal blocks.

# An alternation theorem

- Chebyshev polynomials for compact sets are characterized by alternation properties.
- Example:  $T_m(z)$  for  $[a, b] \subset \mathbb{R}$  has at least  $m + 1$  alternations.

## An Alternation Theorem for Matrices

[Faber, Liesen, T. 2010]

Consider a block-diagonal matrix  $\mathbf{A} = \text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_h)$  where  $d(\mathbf{A}_j) \leq k$ ,  $j = 1, \dots, h$ . Then the matrix

$$T_{k \cdot \ell}^{\mathbf{A}}(\mathbf{A}) = \text{diag}(\mathbf{B}_1, \dots, \mathbf{B}_h) \quad \ell = 1, 2, \dots,$$

has at least  $\ell + 1$  diagonal blocks  $\mathbf{B}_j$  such that

$$\|\mathbf{B}_j\| = \|T_{k \cdot \ell}^{\mathbf{A}}(\mathbf{A})\|.$$

**Example:** If  $\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{C}^{n \times n}$ , then  $T_m^{\mathbf{A}}(\mathbf{A})$  has at least  $m + 1$  diagonal entries with the same maximal absolute value.

# Example

$$\mathbf{A} = \text{diag}(\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4)$$

where each  $\mathbf{A}_j = \mathbf{J}_{\lambda_j}$  is a  $3 \times 3$  Jordan block. The four eigenvalues are  $-3, -0.5, 0.5, 0.75$ , and  $k = d(\mathbf{A}_j) = 3$ .

$m$	$\ T_m^{\mathbf{A}}(\mathbf{A}_1)\ $	$\ T_m^{\mathbf{A}}(\mathbf{A}_2)\ $	$\ T_m^{\mathbf{A}}(\mathbf{A}_3)\ $	$\ T_m^{\mathbf{A}}(\mathbf{A}_4)\ $
1	<u>2.6396</u>	1.4620	2.3970	<u>2.6396</u>
2	<u>4.1555</u>	<u>4.1555</u>	3.6828	<u>4.1555</u>
3	<u>9.0629</u>	5.6303	7.6858	<u>9.0629</u>
4	<u>14.0251</u>	<u>14.0251</u>	11.8397	<u>14.0251</u>
5	<u>22.3872</u>	20.7801	17.6382	<u>22.3872</u>
6	<u>22.6857</u>	<u>22.6857</u>	20.3948	<u>22.6857</u>
7	<u>26.3190</u>	<u>26.3190</u>	<u>26.3190</u>	<u>26.3190</u>

# Chebyshev polynomials $T_m^\Omega(z)$ of compact sets $\Omega \subset \mathbb{C}$

... unique polynomials that solve the problem

$$\min_{p \in \mathcal{M}_m} \max_{z \in \Omega} |p(z)|.$$

## Chebyshev polynomials of $\Omega$ and $\Psi$

[Kamo, Borodin 1994]

Let  $T_k^\Omega$  be the  $k$ th Chebyshev polynomial of the infinite compact set  $\Omega \subset \mathbb{C}$ , let  $p(z)$  be a monic polynomial of degree  $\ell$ , and let

$$\Psi \equiv p^{-1}(\Omega) = \{z \in \mathbb{C} : p(z) \in \Omega\}$$

be the pre-image of  $\Omega$  under the polynomial map  $p$ . Then

$$T_{k \cdot \ell}^\Psi(z) = T_k^\Omega(p(z)).$$

# Special bidiagonal matrices

Given are  $\lambda_1, \dots, \lambda_\ell \in \mathbb{C}$  and  $n \geq 1$ . Consider

$$\mathbf{A} = \begin{bmatrix} \mathbf{D} & \mathbf{E} & & \\ & \mathbf{D} & \ddots & \\ & & \ddots & \mathbf{E} \\ & & & \mathbf{D} \end{bmatrix} \in \mathbb{C}^{\ell \cdot h \times \ell \cdot h},$$

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 1 & & \\ & \lambda_2 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_\ell \end{bmatrix} \in \mathbb{C}^{\ell \times \ell}, \quad \mathbf{E} = \begin{bmatrix} & & & \\ & & & \\ & & & \\ 1 & & & \end{bmatrix} \in \mathbb{R}^{\ell \times \ell},$$

[Reichel, Trefethen 1992] related the pseudospectra of  $\mathbf{A}$  to their symbol  $f_{\mathbf{A}}(z) = \mathbf{D} + z\mathbf{E}$ .

# Chebyshev polynomials for lemniscates

$$\mathbf{A} = \begin{bmatrix} \mathbf{D} & \mathbf{E} & & \\ & \mathbf{D} & \ddots & \\ & & \ddots & \mathbf{E} \\ & & & \mathbf{D} \end{bmatrix} \in \mathbb{C}^{\ell \cdot h \times \ell \cdot h}.$$

- Let  $p(z) = (z - \lambda_1) \cdots (z - \lambda_\ell)$ .
- The lemniscatic region  $\mathcal{L}(p) \equiv \{z \in \mathbb{C} : |p(z)| \leq 1\}$ .
- $\Psi \equiv \mathcal{L}(p)$ ,  $\Omega \equiv$  the unit circle.

Chebyshev polynomials of  $\mathbf{A}$  and of  $\mathcal{L}(p)$  [Faber, Liesen, T. 2010]

$$T_{k \cdot \ell}^{\mathcal{L}(p)}(z) = (p(z))^k = T_{k \cdot \ell}^{\mathbf{A}}(z), \quad k = 1, 2, \dots, h - 1.$$

Moreover,

$$\max_{z \in \mathcal{L}(p)} |T_{k \cdot \ell}^{\mathcal{L}(p)}(z)| = \|T_{k \cdot \ell}^{\mathbf{A}}(\mathbf{A})\|.$$



# Summary

- We considered Chebyshev polynomials of matrices and showed general properties (shifts and scaling, alternation).
- We can relate Chebyshev polynomials for lemniscatic regions to those for certain block-Toeplitz matrices.

**Open question:** Is it possible to translate the problem

$$\min_{p \in \mathcal{M}_m} \|p(\mathbf{A})\|$$

into the problem

$$\min_{p \in \mathcal{M}_m} \max_{z \in \Omega} |p(z)|$$

where  $\Omega$  is a set in the complex plane associated with  $\mathbf{A}$ ?

## Related papers

- V. Faber, J. Liesen and P. Tichý,  
[On Chebyshev polynomials of matrices, accepted for publication in SIMAX (2010).]
- K-C. Toh, N. L. Trefethen,  
[The Chebyshev polynomials of a matrix, SIMAX 20 (1999), no. 2, 400–419]
- A. Greenbaum and N. L. Trefethen,  
[GMRES/CR and Arnoldi/Lanczos as matrix approximation problems, SISC 15 (1994), no. 2, 359–368]

More details can be found at

<http://www.cs.cas.cz/tichy>  
<http://www.math.tu-berlin.de/~liesen>

Thank you for your attention!