# A Padé family of iterations for the matrix sector function and the matrix *p*th root

Krystyna Ziętak

(the joint work with Beata Laszkiewicz)

Institute of Mathematics and Computer Science
Wrocław University of Technology

Gdańsk, GAMM 2009

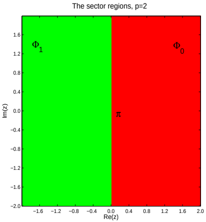
# Outline

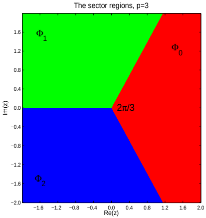
# Sector regions

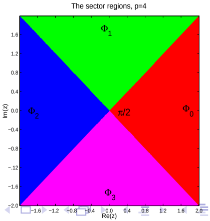
$$\Phi_k = \left\{ z \in \mathbb{C} : \frac{2k\pi}{p} - \frac{\pi}{p} < \arg(z) < \frac{2k\pi}{p} + \frac{\pi}{p} \right\}$$

$k = 0, \ldots, p-1$

p=2 p=3 p=4



# Scalar $p$-sector function

$s_p(\lambda)$ is the nearest *pth* root of unity to $\lambda$

**Representation**

$$s_p(\lambda) = \frac{\lambda}{\sqrt[p]{\lambda^p}}$$

- $\sqrt[p]{a}$ principal *pth* root of $a \notin \mathbb{R}^-$
- $s_p(\lambda)$ is not defined for the *pth* roots of nonpositive real numbers.

# Principal matrix *p*th root

Let nonsingular complex matrix $A$ have no negative eigenvalue. Then there exists unique principal $p$th root of $A$:

$$X = A^{1/p}$$

$$X^p = A, \quad \arg \lambda_j(X) \in \left(-\frac{\pi}{p}, \frac{\pi}{p}\right).$$

N.J. Higham, *Functions of Matrices: Theory and Computation*, SIAM 2008

- $A \in \mathbb{C}^{n \times n}$ nonsingular
- $arg(\lambda_j) \neq 2\pi(q + \frac{1}{2})/p$ for $q \in \{0, \ldots, p-1\}$

**Matrix sector function of $A \in \mathbb{C}^{n \times n}$**

$$S = \mathrm{sect}_p(A) = A\left(\sqrt[p]{A^p}\right)^{-1}$$

Shieh, Tsay, Wang 1984

$S$ is specific *pth* root of $I$: $S^p = I$,
$AS = SA$

## Matrix sector function

$$\mathrm{sect}_p(A) = Z\mathrm{diag}\left(s_p(\lambda_j)I_{r_j}\right)Z^{-1}$$

$$A = Z\mathrm{diag}\left(J_1, J_2, \ldots, J_m\right)Z^{-1},$$

Jordan canonical form

Jordan block $J_k(\lambda)$ of order $r_k$

# Algorithms for matrix sector function

- $$\mathrm{sect}_p(A) = A(A^p)^{-1/p}$$
  $$\mathrm{sect}_p(A) = A\exp(-\log(A^p)/p)$$
- Schur algorithms based on Schur decomposition $A = QRQ^H$
  $$\mathrm{sect}_p(A) = Q\,\mathrm{sect}_p(R)Q^H$$
- Newton's and Halley's iterations
- Padé family of iterations

# Padé family iterations for scalar sector function

$$s_p(\lambda) = \frac{\lambda}{\sqrt[p]{\lambda^p}} = \frac{\lambda}{\sqrt[p]{1-z}}$$

$$z = 1 - \lambda^p$$

## $P_{km}/Q_{km}$ - $[k/m]$ Padé approximant to $\sqrt[p]{1-z}$

$$x_{i+1} = h_{km}(x_i) = x_i \frac{P_{km}(1-x_i^p)}{Q_{km}(1-x_i^p)}, \qquad x_0 = \lambda_j$$

# Padé family for matrix sector function

$$X_{i+1} = h_{km}(X_i), \quad X_0 = A$$

If for every $\lambda_j(A)$ scalar Padé iterations converge to $\mathrm{sect}_p(\lambda_j)$ then matrix Padé iterations are convergent to $\mathrm{sect}_p(A)$.

pure rational iterations

function $h_{km}$ does not depend on $A$

Iannazzo 2007

$$h_{01}(z) = \frac{pz}{z^p + (p-1)}$$

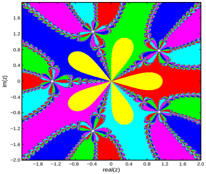
$$h_{10}(z) = \frac{z}{p}[-z^p + (1+p)], \quad h_{11}(z) = z\frac{(p-1)z^p + (p+1)}{(p+1)z^p + (p-1)}$$

$$h_{12}(z) = \frac{2pz[(2p-1)z^p + (p+1)]}{(p+1)z^{2p} + (4p^2+2p-2)z^p + (2p^2-3p+1)}$$

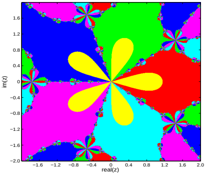
$$h_{22}(z) = \frac{z[(2p^2-3p+1)z^{2p} + (8p^2-2)z^p + (2p^2+3p+1)]}{(2p^2+3p+1)z^{2p} + (8p^2-2)z^p + (2p^2-3p+1)}$$

# Region of convergence of Padé iterations $[m-1/m]$

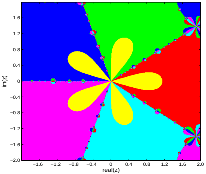
$[0/1]$




$[1/2]$




$[3/4]$




"‘Yellow flower" $\mathbb{L}_p^{(\mathrm{Pade}} = \{z \in \mathbb{C}: \ |1 - z^p| < 1\}$

# Pade $[k/m]$ for sign $(p = 2)$

**Kenney, Laub (1991) - local convergence**

For $k \geq m-1$, if $|1-x_0^2| < 1$ then

$$|1-x_n^2| \leq |1-x_0^2|^{(k+m+1)^n}$$

and

$$\lim_{n\to\infty} x_n = sign(x_0)$$

$$x_0 \in C$$

# Padé $[k/m]$ for sector

## First conjecture for Padé iterations for sector

For $k \geq m-1$, if

$$x_{n+1} = x_n \frac{P_{km}(1-x_n^p)}{Q_k m(1-x_n^p)}$$

$$|1-x_0^p| < 1$$

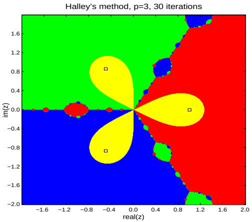
then

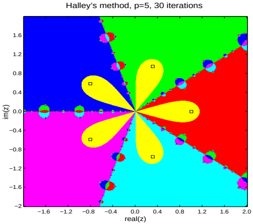
$$|1-x_n^p| \leq |1-x_0^p|^{(k+m+1)^n}$$

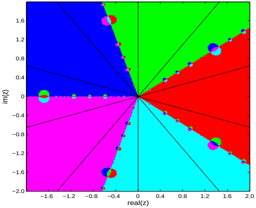
$$\lim_{n\to\infty} x_n \to s_p(x_0)$$

# Halley - Pade [1/1] for sector

$$X_{i+1} = X_i \frac{(p-1)X_i^p + (p+1)I}{(p+1)X_i^p + (p-1)I}, \qquad X_0 = A$$




If all eigenvalues of $A$ lie in

$$\mathbb{B}_p^{(\mathrm{Hall})} = \bigcup_{k=0}^{p-1} \left\{ z \in \mathbb{C} : \frac{2k\pi}{p} - \frac{\pi}{2p} < \arg(z) < \frac{2k\pi}{p} + \frac{\pi}{2p} \right\}$$

then Halley is convergent to sector.

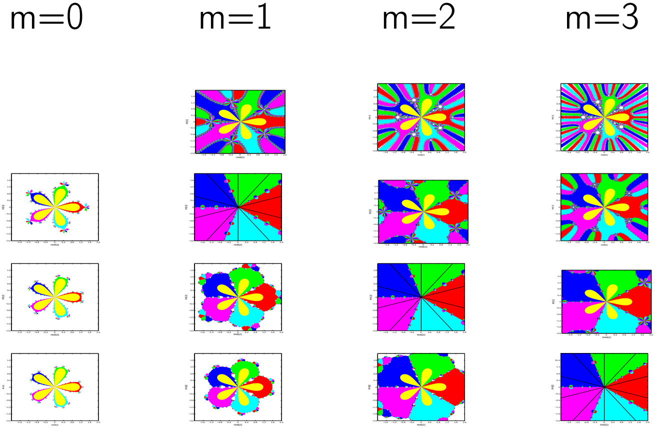
# Principal Padé $[m/m]$ iterations for matrix sector

**Second conjecture for principal Padé iterations for sector**

If all eigenvalues of $A$ lie in $\mathbb{B}_p^{(\mathrm{Hall})}$ then principal Pade $[m/m]$ iterations are convergent to sector.

"Yellow flower" $\mathbb{L}_p^{(\mathrm{Pade})}$ is in $\mathbb{B}_p^{(\mathrm{Hall})}$.

# *Regions of convergence determined experimentally for Padé $[k/m]$ and $p = 5$*



# Particular cases

- The first conjecture with "flowers" is true for Padé $[k/0]$ (it follows from the result of Lakic for the $p$th root)
- The first conjecture is true for Padé $[0/1]$.
- The second conjecture is true for Halley $[1/1]$.

## Structure-preserving matrix iterations

$$X_{k+1} = f(X_k) \in \mathbb{G} \quad \text{if} \quad X_0 = A \in \mathbb{G}$$

## Automorphism groups $\mathbb{G}$

$$A^\star = M^{-1} A^T M, \quad \text{or} \quad A^\star = M^{-1} A^* M, \quad \textbf{adjoint}$$

$$\mathbb{G} = \{A : A^\star = A^{-1}\}$$

## bilinear (real or complex) or sesquilinear forms

$$< x, y >= x^T M y, \quad < x, y >= x^* M y$$

## Examples of automorphism groups

- $M = I$, $A$ real, $\mathbb{G}$ – real orthogonals
- $M = I$, $A$ complex, $\mathbb{G}$ – unitaries
- $M = I$, $A$ complex, $\mathbb{G}$ – complex orthogonals
- $M = J$, $A$ real, $\mathbb{G}$ – real sympletics
- sympletics, perpletics, pseudo-unitaries,...

# Structure-preserving rational matrix iterations

**Higham, Mackey, Tisseur 2004**

If $f(x) = X^r w(X^t) \mathrm{rev}\, w(X^t)^{-1}$ and $X \in \mathbb{G}$ then $f(X) \in \mathbb{G}$.

$$w(x^t) = a_0 + a_1 x^t + a_2 x^{2t} + \cdots + a_k x^{kt}$$
$$\mathrm{rev}\, w(x^t) = a_k + a_{k-1} x^t + a_{k-2} x^{2t} + \cdots + a_0 x^{kt}$$

- Higham Mackey, Tisseur – principal Padé iterations for matrix sign are structure-preserving
- Iannazzo – some methods from König's family of iterations for matrix sector are structure-preserving

# Structure-preserving by principal Padé iterations for sector

$$h_{mm}(x) = x\frac{\sum_{j=0}^{m} b_j^{(m)} x^{pj}}{\sum_{j=0}^{m} c_j^{(m)} x^{pj}}$$

$$b_j^{(m)} = (-1)^j \sum_{k=j}^{m} \binom{k}{j} \frac{\left(\frac{1}{p}\right)_k \left(\frac{1}{p} - m\right)_m (k-2m)_m}{k!(-2m)_m \left(k + \frac{1}{p} - m\right)_m}$$

$$(\alpha)_j = \alpha(\alpha+1)\dots(\alpha+j-1)$$

**by means of Zeilberger algorithm**

$$b_j^{(m)} = \binom{m}{j} \frac{m!}{(2m)!p^m} \prod_{k=m-j+1} (kp-1) \prod_{k=j+1} (kp+1),$$

# Generally convergent methods for polynomials

## Definition

Rational iterative root-finding algorithm is said *generally convergent* if it converges to a root for almost every initial guess and for almost every polynomial.

- Newton's method is generally convergent for quadratic polynomials.
- There does not exist a generally convergent algorithm for polynomials of degree greater than 3.

$$w(x) = x^3 + (c-1)x - c, \quad \text{different roots}$$

### C. McMullen 1987

Every generally convergent algorithm for cubic polynomials is obtained by rational $f$ such that

- $f$ is convergent for $x^3 - 1$
- centralizer of $f$ contains Möbius transformations that permute cube roots of unity

Then the generally convergent algorithm is given by

$$M_c \circ f \circ M_c^{-1}$$

$M_c$ Möbius transformation carrying cube roots of unity to roots of $w(x)$

- Centralizer of $a \in G$ is set of elements of group $G$ which commute with $a$
- Möbius transformation

$$\frac{ax+b}{cx+d}, \qquad ad-bc \neq 0$$

- roots of $w(x) = x^3 + (c-1)x - c$:

$$1, \quad \frac{1}{2}\left(-1-\sqrt{1-4c}\right), \quad \frac{1}{2}\left(-1+\sqrt{1-4c}\right)$$

## J.M. Hawkins 2002

Characterization rational iterations $f$ for $x^3 - 1$ which generate generally convergent algorithms for cubic polynomials – they have to be structure-preserving!

- Generally convergent algorithms for cubic polynomial, proposed by Hawkins, are generated by Padé iterations [2/2] and [3/3] for sector with $p = 3$.
- We can use Padé iterations for sector with $p = 3$ to construct algorithm of arbitrary high order of convergence.

## Padé $[k/m]$ iterations for matrix $p$th root

$$X_{i+1} = X_i P_{km}(I - A^{-1}X_i^p) Q_{km}(I - A^{-1}X_i^p)^{-1}, \quad X_0 = I$$

where $P_{km}(z)/Q_{km}(z)$ is $[k/m]$ Pade for $f(z) = (1-z)^{-1/p}$

## Coupled Padé iteration for the $p$th root of $A$

$$X_{i+1} = X_i h(Y_i), \quad Y_{i+1} = Y_i h(Y_i)^p, \quad X_0 = I,\, Y_0 = A^{-1}$$

$X_i$ tends to $A^{1/p}$, $Y_i$ tends to $I$,
where $h(t) = P_{km}(1-t)/Q_{km}(1-t)$

[1/1] coupled iterations proposed by Iannazzo

# References


- Higham, Mackey, Mackey, Tisseur, Computing the polar decomposition ... in matrix groups, *SIMAX*, 25 (2004).
- Iannazzo, *Numerical solution of certain nonlinear matrix equations*, PhD, Pisa 2007.
- Iannazzo, A family of rational iterations ..., *SIMAX* 30 (2008).
- Kenney, Laub, *Rational iterative methods for the matrix sign function*, *SIMAX* 12 (1991).
- Koç, Bakkaloğlu, Halley's method for the matrix sector function, *IEEE Trans. on Automatic Control* 40 (1995).
- Laszkiewicz, Ziętak, A Padé family of iterations for the matrix sector function and the matrix *p*th root, submitted.
- Shieh, Tsay,Wang, Matrix sector function ..., *IEE Proc.* 131 (1984).

# Thank you for your attention!

Wrocław University of Technology