Structured Factorizations: Theory and Computation

Niloufer Mackey
Department of Mathematics
Western Michigan University
Kalamazoo, U.S.A.

Joint work with
Nick Higham, D. Steven Mackey, and Françoise Tisseur

Householder Symposium
Champion, PA. 27 May 2005
The main theoretical question

If we apply one of the standard factorizations to a matrix that is already structured, to what extent do the factors have additional structure related to the original matrix?

- Square Roots
- Matrix Sign Decomposition
- Polar Decomposition
- eigendecomposition
- SVD
What kind of structure do our matrices have?

They belong to the

- automorphism group
- or Lie algebra
- or Jordan algebra

of a scalar product, that is, a nondegenerate bilinear or sesquilinear form on $\mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}, \mathbb{C}$).

Basic example: $\langle x, y \rangle = x^T y$ defined on $\mathbb{R}^n$. Then

- automorphism group = orthogonals
- Lie algebra = skew-symmetrics
- Jordan algebra = symmetrics
Scalar Products on $K^n$ ($K = \mathbb{R}, \mathbb{C}$)

- A map $K^n \times K^n \to K$, $x, y \mapsto \langle x, y \rangle$

- Bilinear: (real or complex)
  1. $\langle x + y, z + w \rangle = \langle x, z \rangle + \langle y, z \rangle + \langle x, w \rangle + \langle y, w \rangle$
  2. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle = \langle x, \alpha y \rangle$

- Sesquilinear: ($K = \mathbb{C}$)
  2’. $\langle \bar{\alpha} x, y \rangle = \alpha \langle x, y \rangle = \langle x, \alpha y \rangle$

- Matrix Representation
  - Bilinear: $\langle x, y \rangle = x^T M y$
  - Sesquilinear: $\langle x, y \rangle = x^* M y$

- Nondegenerate: $M$ is nonsingular

A scalar product does not have to be symmetric, skew-symmetric, Hermitian, or positive definite.
Adjoint

- For each matrix $A$, its adjoint wrt a scalar product $\langle \cdot, \cdot \rangle$ is the unique matrix $A^*$ s.t.

  $$\langle Ax, y \rangle = \langle x, A^* y \rangle \quad \forall x, y \in \mathbb{K}^n$$

- If $M$ is the matrix of $\langle \cdot, \cdot \rangle$, then

  $$A^* = \begin{cases} M^{-1} A^T M & \text{for bilinear forms,} \\ M^{-1} A^* M & \text{for sesquilinear forms.} \end{cases}$$
Structured Classes Associated with a Scalar Product

- \( \langle Gx, Gy \rangle = \langle x, y \rangle, \ \forall x, y \in \mathbb{K}^n \)
  Automorphisms or Isometries of \( \langle \cdot, \cdot \rangle \).

- \( \langle Kx, y \rangle = -\langle x, Ky \rangle, \ \forall x, y \in \mathbb{K}^n \)
  Skew-adjoint with respect to \( \langle \cdot, \cdot \rangle \).

- \( \langle Sx, y \rangle = \langle x, Sy \rangle, \ \forall x, y \in \mathbb{K}^n \)
  Self-adjoint with respect to \( \langle \cdot, \cdot \rangle \).

In terms of adjoint

\[ G = \{ A \in \mathbb{K}^{n \times n} : A^* = A^{-1} \} , \]
\[ L = \{ A \in \mathbb{K}^{n \times n} : A^* = -A \} , \]
\[ J = \{ A \in \mathbb{K}^{n \times n} : A^* = A \} . \]
### Familiar Classes

\[
J = \begin{bmatrix}
0 & I_n \\
-I_n & 0
\end{bmatrix}, \quad \Sigma_{p,q} = \begin{bmatrix}
I_p & 0 \\
0 & -I_q
\end{bmatrix}^{n \times n}
\]

<table>
<thead>
<tr>
<th>\langle x, y \rangle</th>
<th>\mathbf{G}</th>
<th>\mathbf{J}</th>
<th>\mathbf{L}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\mathbb{R}^n \times \mathbb{R}^n</td>
<td>x^T y</td>
<td>Real orthog</td>
<td>Symm</td>
</tr>
<tr>
<td>\mathbb{C}^n \times \mathbb{C}^n</td>
<td>x^T y</td>
<td>Cplx orthog</td>
<td>Cplx Symm</td>
</tr>
<tr>
<td>\mathbb{R}^n \times \mathbb{R}^n</td>
<td>x^T \Sigma_{p,q} y</td>
<td>Pseudo-orthog</td>
<td>Pseudo symm</td>
</tr>
<tr>
<td>\mathbb{R}^{2n} \times \mathbb{R}^{2n}</td>
<td>x^T J y</td>
<td>Real symplectics</td>
<td>Skew-Hamil</td>
</tr>
<tr>
<td>\mathbb{C}^{2n} \times \mathbb{C}^{2n}</td>
<td>x^T J y</td>
<td>Cplx symplectics</td>
<td>\text{\scriptsize J-skew-symm}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>\langle x, y \rangle</th>
<th>\mathbf{G}</th>
<th>\mathbf{J}</th>
<th>\mathbf{L}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\mathbb{C}^n \times \mathbb{C}^n</td>
<td>x^* y</td>
<td>Unitaries</td>
<td>Herm</td>
</tr>
<tr>
<td>\mathbb{C}^n \times \mathbb{C}^n</td>
<td>x^* \Sigma_{p,q} y</td>
<td>Pseudo-unitaries</td>
<td>Pseudo Herm</td>
</tr>
<tr>
<td>\mathbb{C}^{2n} \times \mathbb{C}^{2n}</td>
<td>x^* J y</td>
<td>Conj symplectics</td>
<td>\text{\scriptsize J-skew-Herm}</td>
</tr>
</tbody>
</table>
**Aim**: Avoid, if possible, working one scalar product at a time.
**Aim**: Avoid, if possible, working one scalar product at a time.

For example, we ask:
“If \( A \in \mathbb{G} \) and \( A = UH \) is its polar decomposition, are \( U, H \) also in \( \mathbb{G} \)?
**Aim**: Avoid, if possible, working one scalar product at a time.

For example, we ask:
“If $A \in \mathbb{G}$ and $A = UH$ is its polar decomposition, are $U$, $H$ also in $\mathbb{G}$?

Though we were not always successful in addressing all scalar products at once, there were some surprisingly general and elegant answers!
Aim: Avoid, if possible, working one scalar product at a time.

For example, we ask:
“If \( A \in G \) and \( A = UH \) is its polar decomposition, are \( U, H \) also in \( G \)?

Though we were not always successful in addressing all scalar products at once, there were some surprisingly general and elegant answers!

http://www.homepages.wmich.edu/~mackey
Some structured factorizations exist in all scalar products.
Structured Principal Square Root

Theorem 1. Suppose $B$ has no eigenvalues on $\mathbb{R}^-$. Then

$$(B^*)^{1/2} = (B^{1/2})^*$$

and

(a) $B \in \mathbb{G} \implies B^{1/2} \in \mathbb{G}$
(b) $B \in \mathbb{J} \implies B^{1/2} \in \mathbb{J}$
(c) $B \in \mathbb{L} \implies B^{1/2}$ is never in $\mathbb{L}$.

Holds in any scalar product space.
Structured Principal Square Root

Theorem 1. Suppose $B$ has no eigenvalues on $\mathbb{R}^{-}$. Then
$(B^*)^{1/2} = (B^{1/2})^*$ and

(a) $B \in G \implies B^{1/2} \in G$
(b) $B \in J \implies B^{1/2} \in J$
(c) $B \in L \implies B^{1/2}$ is never in $L$.

Holds in any scalar product space.

Proof. (a): $B \in G \implies B^* = B^{-1}$

$\implies (B^*)^{1/2} = (B^{-1})^{1/2}$

$\implies (B^{1/2})^* = (B^{1/2})^{-1}$

$\implies B^{1/2} \in G.$
Structured Matrix Sign Decomposition

Defn: Suppose $A = Z \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} Z^{-1}$ has no pure imaginary eigenvalues, and $\Lambda(J_1) \subset \{\text{Re}(z) < 0\}$, $\Lambda(J_2) \subset \{\text{Re}(z) > 0\}$. Then 

$$\text{sign}(A) := Z \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} Z^{-1} := S$$

Can show $A = SN$ where $S = A(A^2)^{-1/2}$, $N = (A^2)^{1/2}$. 

NM; Householder, 27May05 – p. 11/41
Structured Matrix Sign Decomposition

**Defn:** Suppose $A = Z \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} Z^{-1}$ has no pure imaginary eigenvalues, and $\Lambda(J_1) \subset \{\text{Re}(z) < 0\}$, $\Lambda(J_2) \subset \{\text{Re}(z) > 0\}$. Then $\text{sign}(A) := Z \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} Z^{-1} := S$

Can show $A = SN$ where $S = A(A^2)^{-1/2}$, $N = (A^2)^{1/2}$.

**Theorem 2.** Suppose the sign decomposition $A = SN$ exists.

(a) $A \in G \implies S \in G$ and $N \in G$.
(b) $A \in L \implies S \in L$ and $N \in J$.
(c) $A \in J \implies S \in J$ and $N \in J$.

Holds in any scalar product space.
Structured Matrix Sign Decomposition

**Defn:** Suppose $A = Z \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} Z^{-1}$ has no pure imaginary eigenvalues, and $\Lambda(J_1) \subset \{\text{Re}(z) < 0\}$, $\Lambda(J_2) \subset \{\text{Re}(z) > 0\}$. Then

$$\text{sign}(A) := Z \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} Z^{-1} := S$$

Can show $A = SN$ where $S = A(A^2)^{-1/2}$, $N = (A^2)^{1/2}$.

**Theorem 2.** Suppose the sign decomposition $A = SN$ exists.

(a) $A \in G \implies S \in G$ and $N \in G$.
(b) $A \in L \implies S \in L$ and $N \in J$.
(c) $A \in J \implies S \in J$ and $N \in J$.

Holds in any scalar product space.

**Proof.** Use $N = (A^2)^{1/2}$, and structured square root result.
Two Important Classes of Scalar Products

In the literature you find a variety of assumptions imposed on the scalar product.

It is essential to understand whether these assumptions are related.
Orthosymmetric Scalar Products

Theorem 3. Let $\langle \cdot, \cdot \rangle_M$ be a scalar product on $\mathbb{K}^n$. The following are equivalent:

- Adjoint is involutory, that is, $(A^*)^* = A$ for all $A \in \mathbb{K}^{n \times n}$.
- Vector orthogonality is a symmetric relation, that is,
  $$\langle x, y \rangle_M = 0 \iff \langle y, x \rangle_M = 0,$$
  for all $x, y \in \mathbb{K}^n$.
- $\mathbb{K}^{n \times n} = \mathbb{L} \oplus \mathbb{J}$.
- For bilinear forms, $M^T = \pm M$. For sesquilinear forms, $M^* = \alpha M$ with $\alpha \in \mathbb{C}, |\alpha| = 1$;
Orthosymmetric Scalar Products

Theorem 3. Let $\langle \cdot, \cdot \rangle_M$ be a scalar product on $\mathbb{K}^n$. The following are equivalent:

- Adjoint is involutory, that is, $(A^*)^* = A$ for all $A \in \mathbb{K}^{n \times n}$.
- Vector orthogonality is a symmetric relation, that is,
  $$\langle x, y \rangle_M = 0 \iff \langle y, x \rangle_M = 0,$$
  for all $x, y \in \mathbb{K}^n$.
- $\mathbb{K}^{n \times n} = L \oplus J$.
- For bilinear forms, $M^T = \pm M$. For sesquilinear forms, $M^* = \alpha M$ with $\alpha \in \mathbb{C}, |\alpha| = 1$;

Defn: $\langle \cdot, \cdot \rangle_M$ is orthosymmetric if any one (and hence all) of the above properties hold.
Unitary Scalar Products

**Theorem 4.** Let $\langle \cdot, \cdot \rangle_M$ be a scalar product on $\mathbb{K}^n$. The following are equivalent:

- $(A^*)^* = (A^*)^*$ for all $A \in \mathbb{K}^{n \times n}$.
- $U$ unitary $\Rightarrow$ $U^*$ is unitary.
- $H$ hpd $\Rightarrow$ $H^*$ is hpd.
- $M = \alpha V$ for some unitary $V$ and $\alpha > 0$. 
Unitary Scalar Products

Theorem 4. Let $\langle \cdot, \cdot \rangle_M$ be a scalar product on $\mathbb{K}^n$. The following are equivalent:

- $(A^*)^* = (A^*)^*$ for all $A \in \mathbb{K}^{n\times n}$.
- $U$ unitary $\Rightarrow U^*$ is unitary.
- $H$ hpd $\Rightarrow H^*$ is hpd.
- $M = \alpha V$ for some unitary $V$ and $\alpha > 0$.

Defn: $\langle \cdot, \cdot \rangle_M$ is unitary if any one (and hence all) of the above properties hold.
Structured Polar Decomposition

Theorem 5. In a *unitary* scalar product space, $A \in \mathbb{G} \Rightarrow U, \ H \in \mathbb{G}$.
Structured Polar Decomposition

**Theorem 6.** In a *unitary* scalar product space, \( A \in \mathbb{G} \Rightarrow U, H \in \mathbb{G} \).

**Proof.** \( U \) unitary \( \Rightarrow U^* \) unitary, \( H \) hpd \( \Rightarrow H^* \) hpd

\[
A \in \mathbb{G} \Rightarrow A = A^{-*} \\
\Rightarrow A = UH = (UH)^{-*} = U^{-*}H^{-*}
\]

gives us two polar decompositions of \( A \). Uniqueness of the polar factors implies \( U = U^{-*}, H = H^{-*} \). So \( U, H \in \mathbb{G} \). \( \square \)
Theorem 7. Suppose $A$ is nonsingular with polar decomposition $A = UH$. Let $S$ denote either the Lie algebra or the Jordan algebra associated with a unitary scalar product space. Then $A \in S \implies U \in S$.

In general, can’t say much about $H$ other than hpsd. But when $A^*A = AA^*$, more can be said.
Structured Polar Decomposition, ctd

Theorem 7. Suppose $A$ is nonsingular with polar decomposition $A = U H$. Let $S$ denote either the Lie algebra or the Jordan algebra associated with a unitary scalar product space. Then $A \in S \Rightarrow U \in S$.

In general, can’t say much about $H$ other than hpsd. But when $A^* A = AA^*$, more can be said.

The singular case and more is addressed in our paper *Structured Factorizations in Scalar Product Spaces*. 
A Different Kind of Decomposition

Analogue of the polar factorization.

\[
A \in K^{n \times n} \text{ is a } \textit{general} \text{ matrix.} \\
A = WS, \text{ where } W \in G, S \in J.
\]

Analogue of the positive semi-definiteness condition on \( S \)??
A Different Kind of Decomposition

Analogue of the polar factorization.

\( A \in \mathbb{K}^{n \times n} \) is a \emph{general} matrix.

\[ A = W S, \quad \text{where} \quad W \in \mathbb{G}, \ S \in \mathbb{J}. \]

Analogue of the positive semi-definiteness condition on \( S \).

Various approaches, various conditions

- Kaplansky (’90)
- Horn, Merino (’95)
- Bolshakov, vanderMee, Ran, Reichstein, Rodman (’97, etc)
- Ikramov (’01)
- Iserles, Zanna (’02)
- Zanna, Munthe-Kaas (’02)
- Mehl, Ran, Rodman (’05)
**Computable Generalized Polar Decomposition**

**Defn:** Let $\mathbb{K}^n$ be a scalar product space. For $A \in \mathbb{K}^{n \times n}$, a generalized polar decomposition (GPD) is a factorization

$$A = WS,$$

where $W \in \mathbb{G}$, $S \in \mathbb{J}$, and $\text{sign}(S) = I$. 


Computable Generalized Polar Decomposition

Defn: Let $\mathbb{K}^n$ be a scalar product space. For $A \in \mathbb{K}^{n \times n}$, a generalized polar decomposition (GPD) is a factorization

$$A = WS,$$

where $W \in \mathbb{G}$, $S \in \mathbb{J}$, and $\text{sign}(S) = I$.

Theorem 8. For $\mathbb{K}^n$ equipped with an orthosymmetric scalar product, a matrix $A \in \mathbb{K}^{n \times n}$ has a GPD iff $A^*A$ has no eigenvalues on $\mathbb{R}^-$. Whenever this GPD exists it is unique.
Holds in any scalar product space

**Theorem 9.** Let $A \in \mathbb{G}$. The eigenvalues of $A$ come in pairs $\lambda$ and $1/\lambda$ for bilinear forms, and in pairs $\lambda$ and $1/\bar{\lambda}$ for sesquilinear forms. In both cases these pairs have the same Jordan structure, and hence the same algebraic and geometric multiplicities.
Holds in any scalar product space

**Theorem 9.** Let $A \in G$. The eigenvalues of $A$ come in pairs $\lambda$ and $1/\lambda$ for bilinear forms, and in pairs $\lambda$ and $1/\bar{\lambda}$ for sesquilinear forms. In both cases these pairs have the same Jordan structure, and hence the same algebraic and geometric multiplicities.

Well known for real symplectics, but holds more generally.

**Corollary 2.** Let $A \in G$, where $G$ is the automorphism group of a real bilinear form. Then the eigenvalues of $A$ come in quartets $\lambda, 1/\lambda, \bar{\lambda}, 1/\bar{\lambda}$. 
Theorem 10. Let $A \in \mathbb{L}$ or $A \in \mathbb{J}$. Then the eigenvalues of $A$ occur in pairs as shown below, with the same Jordan structure for each eigenvalue in a pair.

<table>
<thead>
<tr>
<th></th>
<th>Bilinear</th>
<th>Sesquilinear</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A \in \mathbb{L}$</td>
<td>$\lambda, -\lambda$</td>
<td>$\lambda, -\bar{\lambda}$</td>
</tr>
<tr>
<td>$A \in \mathbb{J}$</td>
<td>no structure</td>
<td>$\lambda, \bar{\lambda}$</td>
</tr>
</tbody>
</table>
Theorem 10. Let $A \in \mathbb{L}$ or $A \in \mathbb{J}$. Then the eigenvalues of $A$ occur in pairs as shown below, with the same Jordan structure for each eigenvalue in a pair.

<table>
<thead>
<tr>
<th></th>
<th>Bilinear</th>
<th>Sesquilinear</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A \in \mathbb{L}$</td>
<td>$\lambda, -\lambda$</td>
<td>$\lambda, -\bar{\lambda}$</td>
</tr>
<tr>
<td>$A \in \mathbb{J}$</td>
<td>no structure</td>
<td>$\lambda, \bar{\lambda}$</td>
</tr>
</tbody>
</table>

Theorem 11. Let $\mathbb{J}$ be the Jordan algebra of any skew-symmetric bilinear form on $\mathbb{R}^n$ or $\mathbb{C}^n$. Then for any $A \in \mathbb{J}$, the eigenvalues of $A$ all have even multiplicity. Moreover, for any $m > 0$ and eigenvalue $\lambda$, the number of $m \times m$ Jordan blocks corresponding to $\lambda$ in the Jordan form for $A$ is even.
Theorem 12. Let $G$ be the automorphism group of a unitary scalar product, or of a scalar product $\langle \cdot, \cdot \rangle_M$ with $M^2 = \alpha I$ such that $M$ is real when $\langle \cdot, \cdot \rangle_M$ is complex bilinear. Then every $A \in G$ has reciprocally paired singular values.

Remarks:

- includes all the classical matrix groups
- but not a characterization
Theorem 12. Let $G$ be the automorphism group of a unitary scalar product, or of a scalar product $\langle \cdot , \cdot \rangle_M$ with $M^2 = \alpha I$ such that $M$ is real when $\langle \cdot , \cdot \rangle_M$ is complex bilinear. Then every $A \in G$ has reciprocally paired singular values.

Remarks:

- includes all the classical matrix groups
- but not a characterization

Have some positive results for $A \in L, J$, when $\langle \cdot , \cdot \rangle_M$ both unitary and orthosymmetric.
Structured SVD

If $A \in \mathbf{G}$, does $A$ have some SVD for which all three factors are also in $\mathbf{G}$?
Structured SVD

If $A \in \mathbb{G}$, does $A$ have some SVD for which all three factors are also in $\mathbb{G}$?

Yes, for some specific $\mathbb{G}$, but not in general.

- real symplectic, conjugate symplectic (Xu 03)
- complex symplectic (DSM, NM, Mehrmann)
- analytic symplectic SVD (DSM, NM, Mehrmann)
Structured SVD

- If $A \in G$, does $A$ have some SVD for which all three factors are also in $G$?

- Yes, for some specific $G$, but not in general.
  - real symplectic, conjugate symplectic (Xu 03)
  - complex symplectic (DSM, NM, Mehrmann)
  - analytic symplectic SVD (DSM, NM, Mehrmann)

- No.
  - complex orthogonal ($A^{-1} = A^T$)
  - real and complex pseudo-orthogonal, pseudo unitary
    $$(A^{-1} = \Sigma_{p,q} A^T \Sigma_{p,q}, \ A^{-1} = \Sigma_{p,q} A^* \Sigma_{p,q})$$

- Real pseudo-orthogonals (Lorentz matrices) have a structured CS-like decomposition (Higham 03), smooth CS-like decomposition (Dieci, Elia, Lopez 04);
Part I

Characterize functions that preserve matrix groups
i.e., characterize $f$ such that $f(G) \subseteq G$
Computation on Matrix Groups

Part I
Characterize functions that preserve matrix groups
i.e., characterize $f$ such that $f(G) \subseteq G$

Part II
Design structure preserving iterations

$A, f(A), f(f(A)), \ldots \quad A \in G$

that converge to a structured factor of $A$,
e.g., $\text{sign}(A), A^{1/2}, \ldots$
Matrix Functions

- start with scalar function
- extend to a matrix function using any of the usual techniques, Jordan canonical form, Hermite interpolation, etc.
- Convention: whenever $f(A)$ appears, assume $A \in \text{domain}(f)$. 
Bilinear Case

Theorem 13.
(a) For any $f$ and $A \in \mathbb{K}^{n \times n}$, $f(A^*) = f(A)^*$. 
(b) For $A \in G$, $f(A) \in G$ \iff $f(A^{-1}) = f(A)^{-1}$.

Proof. (a) We have

$$f(A^*) = f(M^{-1}A^TM) = M^{-1}f(A^T)M = M^{-1}f(A)^TM = f(A)^*.$$ 

(b) For $A \in G$, consider

$$f(A)^* \begin{array}{c} \parallel \\ \parallel \end{array} f(A^*)$$

$$\parallel \\

f(A^{-1})$$

\end{array}$$

\end{document}
Bilinear Case

Theorem 13.
(a) For any $f$ and $A \in \mathbb{K}^{n \times n}$, $f(A^*) = f(A)^*$.  
(b) For $A \in \mathbb{G}$, $f(A) \in \mathbb{G}$ iff $f(A^{-1}) = f(A)^{-1}$.

Proof. (a) We have

$$f(A^*) = f(M^{-1} A^T M) = M^{-1} f(A^T) M = M^{-1} f(A)^T M = f(A)^*.$$ 

(b) For $A \in \mathbb{G}$, consider

$$f(A)^* = f(A^*)$$

$$|| = ||$$

$$f(A)^{-1} = f(A^{-1})$$
Sesquilinear Case

Theorem 14. $f(A^-*) = f(A)^-*$ iff $f(G) \subseteq G$. 

Meromorphic Functions, Bilinear Case

Meromorphic: analytic on $\mathbb{C}$ except for isolated poles.

When is $f(A^{-1}) = f(A)^{-1}$ for all $A \in \mathbb{G}$?
Meromorphic Functions, Bilinear Case

Meromorphic: analytic on \( \mathbb{C} \) except for isolated poles.

When is \( f(A^{-1}) = f(A)^{-1} \) for all \( A \in \mathbb{G} \)?

**Defn:** If poly \( p \) has degree \( m \) then \( \text{rev} p(z) := x^m p(1/z) \).
Meromorphic Functions, Bilinear Case

Meromorphic: analytic on $\mathbb{C}$ except for isolated poles.

When is $f(A^{-1}) = f(A)^{-1}$ for all $A \in G$?

**Defn:** If poly $p$ has degree $m$ then

$$\text{rev}_p(z) := x^m p(1/z).$$

**Theorem 17.** For bilinear forms on $\mathbb{K}^n$, a meromorphic $f$ satisfies $f(G) \subseteq G$ for all $G$ iff $f$ is rational and of the form

$$f(z) = \pm z^k p(z)/\text{rev}_p(z),$$

for some $k \in \mathbb{Z}$ and some monic $p$ with $p(0) \neq 0$, where $p \in \mathbb{K}[z]$.

Starting point for proof:

A diagonal $\Rightarrow f(z)f(1/z) \equiv 1$ is necessary as a scalar function.
Meromorphic Functions, Sesquilinear Case

When is $f(A^{-*}) = f(A)^{-*}$ for all $A \in \mathbb{G}$?

If $p$ has degree $m$ then $\text{rev} p(z) := z^m p(1/z)$.

**Theorem 18.** For sesquilinear forms on $\mathbb{C}^n$, a meromorphic $f$ satisfies $f(G) \subseteq G$ for all $G$ iff $f$ is rational and of the form

$$f(z) = \alpha z^k p(z) / \text{rev} \overline{p}(z),$$

for some $\alpha \in \mathbb{C}$, $|\alpha| = 1$, $k \in \mathbb{Z}$ and some monic $p$ with $p(0) \neq 0$, where $p \in \mathbb{K}[z]$.

Starting point for proof:
A diagonal $\Rightarrow f(x)f(1/x) \equiv 1$ is necessary.
Analytic Functions

Theorem 19. Let $f$ be analytic on an open subset $\Omega \subseteq \mathbb{C}$ such that each connected component of $\Omega$ is closed under the map $z \mapsto 1/\bar{z}$. Then $f$ is structure preserving for all $G$ associated with a sesquilinear form iff $f$ is structure preserving for the unitary group $U(n)$. 
Part II

Design structure preserving iterations that converge to a structured factor of $A$
Examples of Matrix Sign Iterations

- \( f_1(x) = \frac{1}{2}(x + x^{-1}) = \frac{x(x^2 + 1)}{2x^2} \)

- \( f_2(x) = \frac{x(3 + x^2)}{1 + 3x^2} \)

- \( f_3(x) = \frac{x(5 + 10x^2 + x^4)}{1 + 10x^2 + 5x^4} \)

▶ Observe all are of the form \( Xh(X^2) \).

▶ \( f_2, f_3 \) are structure preserving and converge cubically and quintically respectively.

▶ Kenney & Laub (1991) gave an infinite family of iterations for \( \text{sign}(A) \). All of the form \( Xh(X^2) \).
Class of Square Root Iterations

Theorem 20. Suppose $X_{k+1} = X_k h(X_k^2)$ converges to $\text{sign}(X_0)$ with order $m$, whenever $\text{sign}(X_0)$ exists. Consider the coupled iteration

$$
Y_{k+1} = Y_k h(Z_k Y_k), \quad Y_0 = A,
$$
$$
Z_{k+1} = h(Z_k Y_k) Z_k, \quad Z_0 = I.
$$

If $A^{1/2}$ exists, then

$$
\begin{bmatrix}
Y_k \\
Z_k
\end{bmatrix} \to \begin{bmatrix}
A^{1/2} \\
A^{-1/2}
\end{bmatrix}
$$

as $k \to \infty$, with order $m$.

Moreover, if $X \in \mathbb{G}$ implies $X h(X^2) \in \mathbb{G}$, then $A \in \mathbb{G}$ implies $Y_k \in \mathbb{G}$ and $Z_k \in \mathbb{G}$ for all $k$.

A Fréchet derivative based analysis shows these iterations are numerically stable.
Theorem 21. Suppose the iteration $X_{k+1} = X_k h(X_k^2)$, converges to \( \text{sign}(X_0) \) with order \( m \) whenever \( \text{sign}(X_0) \) exists. Let \( A \) be non-singular with polar decomposition \( A = U H \). Then the iteration

$$Y_{k+1} = Y_k h(Y_k^*Y_k), \quad Y_0 = A$$

converges to \( U \) with order of convergence \( m \). Furthermore, if \( A \in \mathbb{G} \) has a \textit{structured} polar decomposition, and if \( X \in \mathbb{G} \) implies \( X h(X^2) \in \mathbb{G} \), then \( Y_k \in \mathbb{G} \) for all \( k \).
Class of Generalized Polar Factor Iterations

Theorem 22. Suppose a GPD $A = WS$ exists for $A$, with respect to a given scalar product. Suppose the iteration $X_{k+1} = X_k h(X_k^2)$ converges to $\text{sign}(X_0)$ with order $m$ whenever $\text{sign}(X_0)$ exists. For sesquilinear forms assume $h(X^*) = h(X)^*$. Then the iteration

$$Y_{k+1} = Y_k h(Y_k^*Y_k), \quad Y_0 = A$$

converges to $W$ with order of convergence $m$.

A Fréchet derivative based analysis shows these iterations are numerically stable in orthosymmetric scalar products.
Lemma 1. Let \( A, B \in \mathbb{C}^{n \times n} \) and suppose that \( AB \) (and hence also \( BA \)) has no eigenvalues on \( \mathbb{R}^- \). Then

\[
\text{sign} \left( \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & C' \\ C^{-1} & 0 \end{bmatrix}, \quad \text{where} \quad C = A(BA)^{-1/2}
\]

Special cases:

- for \( A \in \mathbb{C}^{n \times n} \) with no eigenvalues on \( \mathbb{R}^- \),

\[
\text{sign} \left( \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & A^{1/2} \\ A^{-1/2} & 0 \end{bmatrix}
\]

- for nonsingular \( A \in \mathbb{C}^{n \times n} \) with unitary polar factor \( U \),

\[
\text{sign} \left( \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & U \\ U^* & 0 \end{bmatrix}
\]
Structure-preserving Iterations for Square Root

Recall: \( f(X) = X^h(X^2) \) converging to \( \text{sign}(X_0) \) yields

\[
\begin{bmatrix}
Y_{k+1} \\
Z_{k+1}
\end{bmatrix} = \begin{bmatrix}
Y_k h(Z_k Y_k) \\
h(Z_k Y_k) Z_k
\end{bmatrix}
\]

for square root.

Cubically convergent: \( f(x) = \frac{x(3 + x^2)}{3x^2 + 1} \)

\[
Y_{k+1} = Y_k (3I + Z_k Y_k)(I + 3Z_k Y_k)^{-1}, \quad Y_0 = A,
\]

\[
Z_{k+1} = (3I + Z_k Y_k)(I + 3Z_k Y_k)^{-1} Z_k, \quad Z_0 = I.
\]
Experimental Results: square root iteration

Random pseudo-orthogonal $A \in \mathbb{R}^{10 \times 10}$, $M = \text{diag}(I_6, -I_4)$, $\kappa_2(A) = 10^{10}$. 
$A$ generated using algorithm of Higham (2003) and chosen to be symmetric positive definite.

$$\text{err}(X) = \frac{\|X - A^{1/2}\|_2}{\|A^{1/2}\|}, \quad \mu_G(X) = \frac{\|X^* X - I\|_2}{\|X\|_2^2}$$

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\text{err}(Y_k)$</th>
<th>$\mu_G(Y_k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.2e+2</td>
<td>1.4e-15</td>
</tr>
<tr>
<td>1</td>
<td>1.0e+2</td>
<td>7.2e-15</td>
</tr>
<tr>
<td>2</td>
<td>3.4e+1</td>
<td>6.1e-14</td>
</tr>
<tr>
<td>3</td>
<td>1.1e+1</td>
<td>5.1e-13</td>
</tr>
<tr>
<td>4</td>
<td>3.0e+0</td>
<td>2.9e-12</td>
</tr>
<tr>
<td>5</td>
<td>5.5e-1</td>
<td>4.4e-12</td>
</tr>
<tr>
<td>6</td>
<td>2.0e-2</td>
<td>4.3e-12</td>
</tr>
<tr>
<td>7</td>
<td>2.0e-6</td>
<td>4.5e-12</td>
</tr>
<tr>
<td>8</td>
<td>2.1e-11</td>
<td>4.8e-12</td>
</tr>
</tbody>
</table>
Recall: \( f(X) = X h(X^2) \), converging to \( \text{sign}(X_0) \) yields

\[ Y_{k+1} = Y_k h(Y_k^*Y_k) \]

converging to unitary polar factor.

**Cubically convergent:**

\[
X_{k+1} = \frac{1}{3} X_k \left[ I + 8 \left( I + 3X_k^*X_k \right)^{-1} \right].
\]

**Quintically convergent:**

\[
x_{k+1} = x_k \left[ \frac{1}{5} + \frac{8}{5x_k^2 + 7 - \frac{16}{5x_k^2 + 3}} \right].
\]
Experimental Results: unitary polar factor iterations

Random symplectic $A \in \mathbb{R}^{12 \times 12}$, $\kappa_2(A) = 9.6 \times 10^4$.

$$\mu_\varnothing(A) = \frac{\|A^*A - I\|_2}{\|A\|_2^2}, \quad \mu_G(A) = \frac{\|A^*A - I\|_2}{\|A\|_2^2}.$$  

<table>
<thead>
<tr>
<th>$k$</th>
<th>Cubic</th>
<th>Quintic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu_\varnothing(X_k)$</td>
<td>$\mu_G(X_k)$</td>
</tr>
<tr>
<td>0</td>
<td>1.0e+0</td>
<td>7.0e-18</td>
</tr>
<tr>
<td>1</td>
<td>1.0e+0</td>
<td>8.9e-17</td>
</tr>
<tr>
<td>2</td>
<td>1.0e+0</td>
<td>8.1e-16</td>
</tr>
<tr>
<td>3</td>
<td>9.9e-01</td>
<td>6.3e-15</td>
</tr>
<tr>
<td>4</td>
<td>9.4e-01</td>
<td>5.0e-14</td>
</tr>
<tr>
<td>5</td>
<td>5.7e-01</td>
<td>2.8e-13</td>
</tr>
<tr>
<td>6</td>
<td>3.6e-02</td>
<td>5.2e-13</td>
</tr>
<tr>
<td>7</td>
<td>3.2e-06</td>
<td>5.3e-13</td>
</tr>
<tr>
<td>8</td>
<td>3.8e-16</td>
<td>5.3e-13</td>
</tr>
</tbody>
</table>
Summary

★ Delineated two important classes of scalar products
★ Established structured factorizations for matrices in \( G, \ L, \ J \)
★ Computable generalized polar decomposition
★ \( f \) preserves group structure if \( f(A^{-1}) = f(A)^{-1} \) (bilinear) or if \( f(A^{-*}) = f(A)^{-*} \) (sesquilinear).
★ Meromorphic functions on \( \mathbb{C} \) mapping \( G \) into itself characterized.
★ Derived new families of coupled iterations for \( A^{1/2} \), and for the unitary polar factor that are structure preserving for all matrix groups.


http://www.homepages.wmich.edu/~mackey