Investigation of Crouzeix’s Conjecture: Numerical Results and Variational Analysis

Michael L. Overton
Courant Institute of Mathematical Sciences
New York University

Joint work with
Anne Greenbaum, University of Washington
and (the second part) with Adrian Lewis, Cornell

AIM Workshop on Crouzeix’s Conjecture
August 1, 2017
Computational Setup

Computing the Field of Values
Chebfun
Computed Extreme Points of $W(\text{diag}(J, B, D))$
The Crouzeix Ratio
Computing the Crouzeix Ratio
Nonsmooth Optimization of the Crouzeix Ratio
Nonsmooth Analysis of the Crouzeix Ratio

Concluding Remarks
Based on Kippenhahn (1951), Johnson (1978) observed that the extreme points of $W(A)$, the field of values of $A$, can be characterized as

$$\text{ext } W(A) = \{ z_\theta = v_\theta^* A v_\theta : \theta \in [0, 2\pi) \}$$

where $v_\theta$ is a normalized eigenvector corresponding to the largest eigenvalue of the Hermitian matrix

$$H_\theta = \frac{1}{2} \left( e^{i\theta} A + e^{-i\theta} A^* \right).$$
Based on Kippenhahn (1951), Johnson (1978) observed that the extreme points of $W(A)$, the field of values of $A$, can be characterized as

$$
\text{ext } W(A) = \{ z_\theta = v_\theta^* Av_\theta : \theta \in [0, 2\pi) \}
$$

where $v_\theta$ is a normalized eigenvector corresponding to the largest eigenvalue of the Hermitian matrix

$$
H_\theta = \frac{1}{2} \left( e^{i\theta} A + e^{-i\theta} A^* \right).
$$

The proof uses a supporting hyperplane argument.
Based on Kippenhahn (1951), Johnson (1978) observed that the extreme points of $W(A)$, the field of values of $A$, can be characterized as

$$\text{ext } W(A) = \{ z_\theta = v_\theta^* A v_\theta : \theta \in [0, 2\pi) \}$$

where $v_\theta$ is a normalized eigenvector corresponding to the largest eigenvalue of the Hermitian matrix

$$H_\theta = \frac{1}{2} \left( e^{i\theta} A + e^{-i\theta} A^* \right).$$

The proof uses a supporting hyperplane argument. Thus, we can compute as many extreme points as we like.
Based on Kippenhahn (1951), Johnson (1978) observed that the extreme points of $W(A)$, the field of values of $A$, can be characterized as

$$\text{ext } W(A) = \{ z_\theta = v_\theta^* A v_\theta : \theta \in [0, 2\pi) \}$$

where $v_\theta$ is a normalized eigenvector corresponding to the largest eigenvalue of the Hermitian matrix

$$H_\theta = \frac{1}{2} \left( e^{i\theta} A + e^{-i\theta} A^* \right).$$

The proof uses a supporting hyperplane argument. Thus, we can compute as many extreme points as we like. But how can we do this accurately, automatically and efficiently?
Chebfun (Trefethen et al, 2004–present) represents real- or complex-valued functions on real intervals to machine precision accuracy using Chebyshev interpolation.
Chebfun (Trefethen et al, 2004–present) represents real- or complex-valued functions on real intervals to machine precision accuracy using Chebyshev interpolation.

The necessary degree of the polynomial is determined automatically. For example, representing $\sin(\pi x)$ on $[-1, 1]$ to machine precision requires degree 19.
Chebfun

Chebfun (Trefethen et al, 2004–present) represents real- or complex-valued functions on real intervals to machine precision accuracy using Chebyshev interpolation.

The necessary degree of the polynomial is determined automatically. For example, representing $\sin(\pi x)$ on $[-1, 1]$ to machine precision requires degree 19.

Most MATLAB functions are overloaded to work with chebfun’s.
Chebfun (Trefethen et al, 2004–present) represents real- or complex-valued functions on real intervals to machine precision accuracy using Chebyshev interpolation.

The necessary degree of the polynomial is determined automatically. For example, representing $\sin(\pi x)$ on $[-1, 1]$ to machine precision requires degree 19.

Most MATLAB functions are overloaded to work with chebfun’s.

Let’s apply Chebfun’s `fov` to compute the extreme points of the field of values of the block diagonal matrix with three $2 \times 2$ blocks $J, B, D$, with $J$ a Jordan block and $D = \text{diag}(5 \pm i)$ . . .
Computed Extreme Points of $W(\text{diag}(J, B, D))$
Computed Extreme Points of $W(\text{diag}(J, B, D))$

Chebfun identifies 5 smooth “pieces” in the set of extreme points
Computed Extreme Points of $W(\text{diag}(J, B, D))$

Chebfun identifies 5 smooth “pieces” in the set of extreme points. The small circles are the interpolation points generated by Chebfun.
The Crouzeix Ratio

Let $p$ be a polynomial and define the Crouzeix ratio

$$f(p, A) = \frac{\|p\|_{W(A)}}{\|p(A)\|_2}.$$
The Crouzeix Ratio

Let $p$ be a polynomial and define the Crouzeix ratio

$$f(p, A) = \frac{\|p\|_{W(A)}}{\|p(A)\|_2}.$$

The conjecture states that $f(p, A)$ is bounded below by 0.5 independently of the polynomial degree $m$ and the matrix order $n$. 
Let \( p \) be a polynomial and define the Crouzeix ratio

\[
f(p, A) = \frac{\|p\|_{W(A)}}{\|p(A)\|_2}.
\]

The conjecture states that \( f(p, A) \) is bounded below by 0.5 independently of the polynomial degree \( m \) and the matrix order \( n \). The Crouzeix ratio \( f \) is

- A mapping from \( \mathbb{C}^{m+1} \times \mathbb{C}^{n \times n} \) to \( \mathbb{R} \) (associating polynomials \( p \in P^m \) with their vectors of coefficients \( c \in \mathbb{C}^{m+1} \) using the monomial basis)
The Crouzeix Ratio

Let $p$ be a polynomial and define the Crouzeix ratio

$$f(p, A) = \frac{\|p\|_{W(A)}}{\|p(A)\|_2}.$$ 

The conjecture states that $f(p, A)$ is bounded below by 0.5 independently of the polynomial degree $m$ and the matrix order $n$. The Crouzeix ratio $f$ is

- A mapping from $\mathbb{C}^{m+1} \times \mathbb{C}^{n \times n}$ to $\mathbb{R}$ (associating polynomials $p \in P^m$ with their vectors of coefficients $c \in \mathbb{C}^{m+1}$ using the monomial basis)
- Not convex
The Crouzeix Ratio

Let $p$ be a polynomial and define the Crouzeix ratio

$$f(p, A) = \frac{\|p\|_{W(A)}}{\|p(A)\|_2}.$$

The conjecture states that $f(p, A)$ is bounded below by $0.5$ independently of the polynomial degree $m$ and the matrix order $n$. The Crouzeix ratio $f$ is

- A mapping from $\mathbb{C}^{m+1} \times \mathbb{C}^{n \times n}$ to $\mathbb{R}$ (associating polynomials $p \in P^m$ with their vectors of coefficients $c \in \mathbb{C}^{m+1}$ using the monomial basis)
- Not convex
- Not defined if $p(A) = 0$
Let $p$ be a polynomial and define the Crouzeix ratio

$$f(p, A) = \frac{\|p\|_{W(A)}}{\|p(A)\|_2}.$$ 

The conjecture states that $f(p, A)$ is bounded below by 0.5 independently of the polynomial degree $m$ and the matrix order $n$. The Crouzeix ratio $f$ is

- A mapping from $\mathbb{C}^{m+1} \times \mathbb{C}^{n \times n}$ to $\mathbb{R}$ (associating polynomials $p \in P^m$ with their vectors of coefficients $c \in \mathbb{C}^{m+1}$ using the monomial basis)
- Not convex
- Not defined if $p(A) = 0$
- Lipschitz continuous at all other points, but not necessarily differentiable
The Crouzeix Ratio

Let \( p \) be a polynomial and define the Crouzeix ratio

\[
f(p, A) = \frac{\|p\|_{W(A)}}{\|p(A)\|_2}.
\]

The conjecture states that \( f(p, A) \) is bounded below by 0.5 independently of the polynomial degree \( m \) and the matrix order \( n \). The Crouzeix ratio \( f \) is

- A mapping from \( \mathbb{C}^{m+1} \times \mathbb{C}^{n \times n} \) to \( \mathbb{R} \) (associating polynomials \( p \in P^m \) with their vectors of coefficients \( c \in \mathbb{C}^{m+1} \) using the monomial basis)
- Not convex
- Not defined if \( p(A) = 0 \)
- Lipschitz continuous at all other points, but not necessarily differentiable
- Semialgebraic (its graph is a finite union of sets, each of which is defined by a finite system of polynomial inequalities)
Computing the Crouzeix Ratio

Numerator: use Chebfun’s `fov` (modified to return any line segments in the boundary) combined with its overloaded `polyval` and `norm(\cdot,\infty)`. 
Computing the Crouzeix Ratio

Numerator: use Chebfun’s \texttt{fov} (modified to return any line segments in the boundary) combined with its overloaded \texttt{polyval} and \texttt{norm(·,\text{inf})}.

Denominator: use \texttt{MATLAB}'s standard \texttt{polyvalm} and \texttt{norm(·,2)}. 
Numerator: use Chebfun’s `fov` (modified to return any line segments in the boundary) combined with its overloaded `polyval` and `norm(·,inf)`.

Denominator: use MATLAB’s standard `polyvalm` and `norm(·,2)`.

The main cost is the construction of the chebfun defining the field of values.
Nonsmooth Optimization of the Crouzeix Ratio

Computational Setup

- Nonsmooth Optimization of the Crouzeix Ratio
- Nonsmoothness of the Crouzeix Ratio
- BFGS Experiments
- Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)
- Final Fields of Values for Lowest Computed $f$
- Optimizing over both $p$ and $A$: Final $f(p, A)$
- Is the Ratio 0.5 Attained?
- Final Fields of Values for $f$ Closest to 1
- Why is the Crouzeix Ratio One?
- Results for Larger Dimension $n$ and Degree $n - 1$
- Optimizing the Radius Ratio Instead
  - Optimizing the Radius Ratio, $n = 2$
  - Optimizing the
There are three possible sources of nonsmoothness in the Crouzeix ratio $f$.
There are three possible sources of nonsmoothness in the Crouzeix ratio $f$

- When the max value of $|p(\zeta)|$ on $\text{bd } W(A)$ is attained at more than one point $\zeta$ (the most important, as this frequently occurs at apparent minimizers)
There are three possible sources of nonsmoothness in the Crouzeix ratio $f$

- When the max value of $|p(\zeta)|$ on $bd W(A)$ is attained at more than one point $\zeta$ (the most important, as this frequently occurs at apparent minimizers)

- Even if such $\zeta$ is unique, when the normalized vector $v$ for which $v^*Av = \zeta$ is not unique up to a scalar, implying that the maximum eigenvalue of the corresponding $H_\theta$ matrix has multiplicity two or more (does not seem to occur at minimizers)
Nonsmoothness of the Crouzeix Ratio

There are three possible sources of nonsmoothness in the Crouzeix ratio $f$

- When the max value of $|p(\zeta)|$ on $\text{bd } W(A)$ is attained at more than one point $\zeta$ (the most important, as this frequently occurs at apparent minimizers)
- Even if such $\zeta$ is unique, when the normalized vector $v$ for which $v^* A v = \zeta$ is not unique up to a scalar, implying that the maximum eigenvalue of the corresponding $H_\theta$ matrix has multiplicity two or more (does not seem to occur at minimizers)
- When the maximum singular value of $p(A)$ has multiplicity two or more (does not seem to occur at minimizers)
There are three possible sources of nonsmoothness in the Crouzeix ratio $f$

- When the max value of $|p(\zeta)|$ on $\text{bd } W(A)$ is attained at more than one point $\zeta$ (the most important, as this frequently occurs at apparent minimizers)

- Even if such $\zeta$ is unique, when the normalized vector $v$ for which $v^*Av = \zeta$ is not unique up to a scalar, implying that the maximum eigenvalue of the corresponding $H_\theta$ matrix has multiplicity two or more (does not seem to occur at minimizers)

- When the maximum singular value of $p(A)$ has multiplicity two or more (does not seem to occur at minimizers)

In all of these cases the gradient of $f$ is not defined.

But in practice, none of these cases ever occur, except the first one in the limit.
BFGS (Broyden, Fletcher, Goldfarb and Shanno, all independently in 1970), is the standard quasi-Newton algorithm for minimizing smooth (continuously differentiable) functions.
BFGS (Broyden, Fletcher, Goldfarb and Shanno, all independently in 1970), is the standard quasi-Newton algorithm for minimizing smooth (continuously differentiable) functions. It works by building an approximation to the Hessian of the function using gradient differences, and has a well known superlinear convergence property under a regularity condition.
BFGS (Broyden, Fletcher, Goldfarb and Shanno, all independently in 1970), is the standard quasi-Newton algorithm for minimizing smooth (continuously differentiable) functions. It works by building an approximation to the Hessian of the function using gradient differences, and has a well known superlinear convergence property under a regularity condition. Although its global convergence theory is limited to the convex case (Powell, 1976), it generally finds local minimizers efficiently in the nonconvex case too, although there are pathological counterexamples.
BFGS (Broyden, Fletcher, Goldfarb and Shanno, all independently in 1970), is the standard quasi-Newton algorithm for minimizing smooth (continuously differentiable) functions. It works by building an approximation to the Hessian of the function using gradient differences, and has a well known superlinear convergence property under a regularity condition. Although its global convergence theory is limited to the convex case (Powell, 1976), it generally finds local minimizers efficiently in the nonconvex case too, although there are pathological counterexamples.

Remarkably, this property seems to extend to nonsmooth functions too, with a linear (geometric) convergence rate, although the convergence theory is extremely limited (Lewis and Overton, 2013). It builds a very ill conditioned “Hessian” approximation, with “infinitely large” curvature in some directions and finite curvature in other directions.
We have run many experiments searching for local minimizers of the Crouzeix ratio using BFGS.
Experiments

We have run many experiments searching for local minimizers of the Crouzeix ratio using BFGS.

For fixed $n$, optimize over $A$ with order $n$ and $p$ of $\deg \leq n - 1$, running BFGS for a maximum of 1000 iterations from each of 100 randomly generated starting points.
Experiments

We have run many experiments searching for local minimizers of the Crouzeix ratio using BFGS.

For fixed $n$, optimize over $A$ with order $n$ and $p$ of deg $\leq n - 1$, running BFGS for a maximum of 1000 iterations from each of 100 randomly generated starting points.

We restrict $p$ to have real coefficients and $A$ to be real, in Hessenberg form (all but one superdiagonal is zero).
We have run many experiments searching for local minimizers of the Crouzeix ratio using BFGS.

For fixed $n$, optimize over $A$ with order $n$ and $p$ of deg $\leq n - 1$, running BFGS for a maximum of 1000 iterations from each of 100 randomly generated starting points.

We restrict $p$ to have real coefficients and $A$ to be real, in Hessenberg form (all but one superdiagonal is zero).

We have obtained similar results for $p$ with complex coefficients and complex $A$ (then can take $A$ to be triangular).
Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Computational Setup
Nonsmooth Optimization of the Crouzeix Ratio
Nonsmoothness of the Crouzeix Ratio
BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$
Optimizing over both $p$ and $A$: Final $f(p, A)$
Is the Ratio 0.5 Attained?
Final Fields of Values for $f$ Closest to 1
Why is the Crouzeix Ratio One?
Results for Larger Dimension $n$ and Degree $n - 1$
Optimizing the Radius Ratio Instead
Optimizing the Radius Ratio, $n = 2$

Sorted final values of the Crouzeix ratio $f$
found starting from 100 randomly generated initial points.
Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension $n$ and Degree $n - 1$

Optimizing the Radius Ratio Instead

Optimizing the Radius Ratio, $n = 2$

Sorted final values of the Crouzeix ratio $f$ found starting from 100 randomly generated initial points.

Are 0.5 and 1 the only locally optimal values of $f$?
Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension $n$ and Degree $n - 1$

Optimizing the Radius Ratio Instead

Optimizing the Radius Ratio, $n = 2$

Optimizing the

Solid blue curve is boundary of field of values of final computed $A$

Blue asterisks are eigenvalues of final computed $A$

Small red circles are roots of final computed $p$
Final Fields of Values for Lowest Computed $f$

Computational Setup
Nonsmooth Optimization of the Crouzeix Ratio
BFGS Experiments
Optimizing over $A$ (order $n$) and $p$ (deg $\leq n-1$)

Final Fields of Values for Lowest Computed $f$
Optimizing over both $p$ and $A$: Final $f(p,A)$
Is the Ratio 0.5 Attained?
Final Fields of Values for $f$ Closest to 1
Why is the Crouzeix Ratio One?
Results for Larger Dimension $n$ and Degree $n - 1$
Optimizing the Radius Ratio Instead
Optimizing the Radius Ratio,
$n = 2$

Solid blue curve is boundary of field of values of final computed $A$
Blue asterisks are eigenvalues of final computed $A$
Small red circles are roots of final computed $p$

$n = 3, 4, 5$: two eigenvalues of $A$ and one root of $p$ nearly coincident
Optimizing over both $p$ and $A$: Final $f(p, A)$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.500000000000000000</td>
</tr>
<tr>
<td>4</td>
<td>0.500000000000000000</td>
</tr>
<tr>
<td>5</td>
<td>0.500000000000000014</td>
</tr>
<tr>
<td>6</td>
<td>0.5000000017156953</td>
</tr>
<tr>
<td>7</td>
<td>0.500000746246673</td>
</tr>
<tr>
<td>8</td>
<td>0.500000206563813</td>
</tr>
</tbody>
</table>

$f$ is the lowest value $f(p, A)$ found over 100 runs.
Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension $n$ and Degree $n - 1$

Optimizing the Radius Ratio Instead

Optimizing the Radius Ratio, $n = 2$

Optimizing the
Is the Ratio 0.5 Attained?

 Independently, Crabb (1970! thanks Abbas!), Crouzeix and Choi showed that the ratio 0.5 is attained if $p(\zeta) = \zeta^{n-1}$ and $A$ is the $n$ by $n$ “CCC” matrix

$$\Xi_n = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \text{ if } n = 2, \text{ or } \begin{bmatrix} 0 & \sqrt{2} \\ \cdot & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & 1 \\ \cdot & \cdot \\ \sqrt{2} & 0 \end{bmatrix} \text{ if } n > 2$$

for which $W(A)$ is the closed unit disk $\overline{D}$. 
Is the Ratio 0.5 Attained?

Independently, Crabb (1970! thanks Abbas!), Crouzeix and Choi showed that the ratio 0.5 is attained if $p(\zeta) = \zeta^{n-1}$ and $A$ is the $n$ by $n$ “CCC” matrix

$$\Xi_n = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \text{ if } n = 2, \text{ or } \begin{bmatrix} 0 & \sqrt{2} \\ \vdots & 1 \\ \vdots & \vdots \\ \sqrt{2} & 0 \end{bmatrix} \text{ if } n > 2$$

for which $W(A)$ is the closed unit disk $\overline{D}$.

We find that computed minimizers have the form

$$A = \lambda I + \alpha U \text{diag}(\Xi, B)U^T + E,$$

where $k \geq 2$ (usually $k = 2$), $\alpha \neq 0$, $U$ is orthogonal, $W(B) \subset \overline{D}$, $\|E\|$ is small and $|c_j|$ is small for $j \geq k$. 
Is the Ratio 0.5 Attained?

Independently, Crabb (1970! thanks Abbas!), Crouzeix and Choi showed that the ratio 0.5 is attained if \( p(\zeta) = \zeta^{n-1} \) and \( A \) is the \( n \) by \( n \) “CCC” matrix

\[
\Xi_n = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{if } n = 2, \text{ or } \quad \begin{bmatrix} 0 & \sqrt{2} \\ \cdot & 1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & 1 \\ \cdot & \sqrt{2} \\ 0 \end{bmatrix} \quad \text{if } n > 2
\]

for which \( W(A) \) is the closed unit disk \( \overline{D} \).

We find that computed minimizers have the form \(^1\)

\[
A = \lambda I + \alpha U \text{diag}(\Xi_k, B) U^T + E, \quad p(\zeta) = c_{n-1} \zeta^{n-1} + \ldots + c_1 \zeta + c_0
\]

where \( k \geq 2 \) (usually \( k = 2 \)), \( \alpha \neq 0 \), \( U \) is orthogonal, \( W(B) \subset \overline{D} \), \( \|E\| \) is small and \( |c_j| \) is small for \( j \geq k \).

A. Salemi pointed out that our conjecture that these (with \( E = 0 \), \( c_j = 0 \)) are the only cases where \( f(p, A) = 0.5 \) with \( p \) equal to a monomial \( \zeta^m \) follows from Crabb.
Is the Ratio 0.5 Attained?

Independently, Crabb (1970! thanks Abbas!), Crouzeix and Choi showed that the ratio 0.5 is attained if \( p(\zeta) = \zeta^{n-1} \) and \( A \) is the \( n \) by \( n \) “CCC” matrix

\[
\Xi_n = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{if } n = 2, \quad \text{or} \quad \begin{bmatrix} 0 & \sqrt{2} \\ \vdots & 1 \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & 1 \\ \sqrt{2} & 0 \end{bmatrix} \quad \text{if } n > 2
\]

for which \( W(A) \) is the closed unit disk \( \overline{D} \).

We find that computed minimizers have the form \(^1\)

\[
A = \lambda I + \alpha U \text{diag}(\Xi_k, B) U^T + E, \quad p(\zeta) = c_{n-1} \zeta^{n-1} + \ldots + c_1 \zeta + c_0
\]

where \( k \geq 2 \) (usually \( k = 2 \)), \( \alpha \neq 0 \), \( U \) is orthogonal, \( W(B) \subset \overline{D} \), \( \|E\| \) is small and \( |c_j| \) is small for \( j \geq k \).

A. Salemi pointed out that our conjecture that these (with \( E = 0, c_j = 0 \)) are the only cases where \( f(p, A) = 0.5 \) with \( p \) equal to a monomial \( \zeta^m \) follows from Crabb. Conjecture: these (with \( E = 0, c_j = 0 \)) are the only cases where \( f(p, A) = 0.5 \), allowing any polynomial \( p \).
Is the Ratio 0.5 Attained?

Independently, Crabb (1970! thanks Abbas!), Crouzeix and Choi showed that the ratio 0.5 is attained if $p(\zeta) = \zeta^{n-1}$ and $A$ is the $n$ by $n$ “CCC” matrix

$$
\Xi_n = \begin{bmatrix}
0 & 2 \\
0 & 0
\end{bmatrix} \text{ if } n = 2, \quad \begin{bmatrix}
0 & \sqrt{2} \\
0 & 0 \\
0 & 0 \\
\ddots & \ddots & \ddots & \ddots \\
0 & 0 & \ddots & \ddots & \ddots \\
0 & 0 & \cdots & \ddots & \ddots & \ddots \\
0 & 0 & \cdots & \cdots & \ddots & \ddots \\
0 & 0 & \cdots & \cdots & \cdots & \ddots \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 \\
0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 \\
\end{bmatrix} \text{ if } n > 2
$$

for which $W(A)$ is the closed unit disk $\overline{D}$.

We find that computed minimizers have the form

$$
A = \lambda I + \alpha U \text{diag}(\Xi_k, B) U^T + E, \quad p(\zeta) = c_{n-1}\zeta^{n-1} + \ldots + c_1\zeta + c_0
$$

where $k \geq 2$ (usually $k = 2$), $\alpha \neq 0$, $U$ is orthogonal, $W(B) \subset \overline{D}$, $\|E\|$ is small and $|c_j|$ is small for $j \geq k$.

A. Salemi pointed out that our conjecture that these (with $E = 0$, $c_j = 0$) are the only cases where $f(p, A) = 0.5$ with $p$ equal to a monomial $\zeta^m$ follows from Crabb.

Conjecture: these (with $E = 0$, $c_j = 0$) are the only cases where $f(p, A) = 0.5$, allowing any polynomial $p$.

However, this is not true if we allow $p$ to be any analytic function. (Crouzeix has a complete analysis for $n = 3$.)
Is the Ratio 0.5 Attained?

Independently, Crabb (1970! thanks Abbas!), Crouzeix and Choi showed that the ratio 0.5 is attained if \( p(\zeta) = \zeta^{n-1} \) and \( A \) is the \( n \) by \( n \) “CCC” matrix

\[
\Xi_n = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \text{ if } n = 2, \text{ or } \begin{bmatrix} 0 & \sqrt{2} & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \sqrt{2} \\ 0 \end{bmatrix} \text{ if } n > 2
\]

for which \( W(A) \) is the closed unit disk \( \overline{D} \).

We find that computed minimizers have the form

\[
A = \lambda I + \alpha U \text{diag}(\Xi_k, B)U^T + E, \quad p(\zeta) = c_{n-1}\zeta^{n-1} + \ldots + c_1\zeta + c_0
\]

where \( k \geq 2 \) (usually \( k = 2 \)), \( \alpha \neq 0 \), \( U \) is orthogonal, \( W(B) \subset \overline{D} \), \( \|E\| \) is small and \( |c_j| \) is small for \( j \geq k \).

A. Salemi pointed out that our conjecture that these (with \( E = 0 \), \( c_j = 0 \)) are the only cases where \( f(p, A) = 0.5 \) with \( p \) equal to a monomial \( \zeta^m \) follows from Crabb. Conjecture: these (with \( E = 0 \), \( c_j = 0 \)) are the only cases where \( f(p, A) = 0.5 \), allowing any polynomial \( p \).

However, this is not true if we allow \( p \) to be any analytic function. (Crouzeix has a complete analysis for \( n = 3 \).)

Note: \( f \) is nonsmooth at these pairs \((p, A)\) because \(|p|\) is constant on the boundary of \( W(A) \).
Is the Ratio 0.5 Attained?

Independently, Crabb (1970! thanks Abbas!), Crouzeix and Choi showed that the ratio 0.5 is attained if \( p(\zeta) = \zeta^{n-1} \) and \( A \) is the \( n \) by \( n \) "CCC" matrix

\[
\Xi_n = \begin{bmatrix}
0 & 2 \\
0 & 0
\end{bmatrix}
\text{ if } n = 2, \text{ or }
\begin{bmatrix}
0 & \sqrt{2} & 1 \\
& & \\
& & \\
& & \sqrt{2}
\end{bmatrix}
\text{ if } n > 2
\]

for which \( W(A) \) is the closed unit disk \( \overline{D} \).

We find that computed minimizers have the form \(^1\)

\[
A = \lambda I + \alpha U \text{diag}(\Xi_k, B)U^T + E, \quad p(\zeta) = c_{n-1}\zeta^{n-1} + \ldots + c_1\zeta + c_0
\]

where \( k \geq 2 \) (usually \( k = 2 \)), \( \alpha \neq 0 \), \( U \) is orthogonal, \( W(B) \subset \overline{D} \), \( \|E\| \) is small and \( |c_j| \) is small for \( j \geq k \).

A. Salemi pointed out that our conjecture that these (with \( E = 0, c_j = 0 \)) are the only cases where \( f(p, A) = 0.5 \) with \( p \) equal to a monomial \( \zeta^m \) follows from Crabb. Conjecture: these (with \( E = 0, c_j = 0 \)) are the only cases where \( f(p, A) = 0.5 \), allowing any polynomial \( p \).

However, this is not true if we allow \( p \) to be any analytic function. (Crouzeix has a complete analysis for \( n = 3 \).)

Note: \( f \) is nonsmooth at these pairs \((p, A)\) because \(|p| \) is constant on the boundary of \( W(A) \).

Final Fields of Values for $f$ Closest to 1

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS

Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension $n$ and Degree $n - 1$

Optimizing the Radius Ratio Instead

Optimizing the Radius Ratio, $n = 2$

Optimizing the
Final Fields of Values for $f$ Closest to 1

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension $n$ and Degree $n - 1$

Optimizing the Radius Ratio Instead

Optimizing the Radius Ratio, $n = 2$

Optimizing the

Ice cream cone shape:

exactly one eigenvalue at a vertex of the field of values
Why is the Crouzeix Ratio One?

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension $n$ and Degree $n - 1$

Optimizing the Radius Ratio Instead

Optimizing the Radius Ratio, $n = 2$
Why is the Crouzeix Ratio One?

Because for this computed local minimizer, \( A \) is nearly unitarily similar to a block diagonal matrix

\[
\text{diag}(\lambda, B), \quad \lambda \in \mathbb{R}
\]

so

\[
W(A) \approx \text{conv}(\lambda, W(B))
\]

with \( \lambda \) \textit{active} and the block \( B \) \textit{inactive}, that is:

- \( \|p\|_{W(A)} \) is attained only at \( \lambda \)
- \( |p(\lambda)| > \|p(B)\|_2 \)
Why is the Crouzeix Ratio One?

Because for this computed local minimizer, $A$ is nearly unitarily similar to a block diagonal matrix

$$\text{diag}(\lambda, B), \quad \lambda \in \mathbb{R}$$

so

$$W(A) \approx \text{conv}(\lambda, W(B))$$

with $\lambda$ active and the block $B$ inactive, that is:

- $\|p\|_{W(A)}$ is attained only at $\lambda$
- $|p(\lambda)| > \|p(B)\|_2$

So, $\|p\|_{W(A)} = |p(\lambda)| = \|p(A)\|_2$ and hence $f(p, A) = 1$. 
Why is the Crouzeix Ratio One?

Because for this computed local minimizer, $A$ is nearly unitarily similar to a block diagonal matrix

$$\text{diag}(\lambda, B), \quad \lambda \in \mathbb{R}$$

so

$$W(A) \approx \text{conv}(\lambda, W(B))$$

with $\lambda$ active and the block $B$ inactive, that is:

- $\|p\|_{W(A)}$ is attained only at $\lambda$
- $|p(\lambda)| > \|p(B)\|_2$

So, $\|p\|_{W(A)} = |p(\lambda)| = \|p(A)\|_2$ and hence $f(p, A) = 1$.

Furthermore, $f$ is differentiable at this pair $(p, A)$, with zero gradient. Thus, such $(p, A)$ is a smooth stationary point of $f$. 
Why is the Crouzeix Ratio One?

Because for this computed local minimizer, $A$ is nearly unitarily similar to a block diagonal matrix

$$\text{diag}(\lambda, B), \quad \lambda \in \mathbb{R}$$

so

$$W(A) \approx \text{conv}(\lambda, W(B))$$

with $\lambda$ active and the block $B$ inactive, that is:

- $\|p\|_{W(A)}$ is attained only at $\lambda$
- $|p(\lambda)| > \|p(B)\|_2$

So, $\|p\|_{W(A)} = |p(\lambda)| = \|p(A)\|_2$ and hence $f(p, A) = 1$.

Furthermore, $f$ is differentiable at this pair $(p, A)$, with zero gradient. Thus, such $(p, A)$ is a smooth stationary point of $f$.

This doesn’t imply that it is a local minimizer, but the numerical results make this evident.
Why is the Crouzeix Ratio One?

Because for this computed local minimizer, $A$ is nearly unitarily similar to a block diagonal matrix

$$\text{diag} (\lambda, B), \quad \lambda \in \mathbb{R}$$

so

$$W(A) \approx \text{conv}(\lambda, W(B))$$

with $\lambda$ active and the block $B$ inactive, that is:

- $\|p\|_{W(A)}$ is attained only at $\lambda$
- $|p(\lambda)| > \|p(B)\|_2$

So, $\|p\|_{W(A)} = |p(\lambda)| = \|p(A)\|_2$ and hence $f(p, A) = 1$.

Furthermore, $f$ is differentiable at this pair $(p, A)$, with zero gradient. Thus, such $(p, A)$ is a smooth stationary point of $f$.

This doesn’t imply that it is a local minimizer, but the numerical results make this evident.

As $n$ increases, ice cream cone stationary points become increasingly common and it becomes very difficult to reduce $f$ below 1.
Results for Larger Dimension $n$ and Degree $n - 1$

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Sorted final values of the Crouzeix ratio $f$

found starting from many randomly generated initial points.
Results for Larger Dimension $n$ and Degree $n - 1$

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Sorted final values of the Crouzeix ratio $f$

found starting from many randomly generated initial points.

There are other locally optimal values of $f$ between 0.5 and 1!
We learned yesterday from Michel and César that Crouzeix’s conjecture is true if and only if the numerical radius of $p(A)$ satisfies

$$r(p(A)) \leq \frac{5}{4} \|p\|_{W(A)}$$

for all matrices $A$ and polynomials $p$. So, we can instead minimize the “radius ratio”

$$\frac{\|p\|_{W(A)}}{r(p(A))}$$

over $A$ with order $n$ and $p$ with degree $\leq m$.

The results on the next slide were obtained late last night and early this morning.
Optimizing the Radius Ratio, $n = 2$

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio
Nonsmoothness of the Crouzeix Ratio
BFGS
Experiments
Optimizing over $A$ (order $n$) and $p$ ($\text{deg} \leq n - 1$)
Final Fields of Values for Lowest Computed $f$
Optimizing over both $p$ and $A$: Final $f(p, A)$
Is the Ratio 0.5 Attained?
Final Fields of Values for $f$ Closest to 1
Why is the Crouzeix Ratio One?
Results for Larger Dimension $n$ and Degree $n - 1$
Optimizing the Radius Ratio Instead

Sorted final values of the radius ratio $f$, $n = 2$

<table>
<thead>
<tr>
<th>$m$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>0.815</td>
<td>0.808</td>
<td>0.805</td>
<td>0.803</td>
<td>0.802</td>
<td>0.801</td>
</tr>
</tbody>
</table>
Optimizing the Radius Ratio, $n = 2$

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS Experiments

Optimizing over $A$ (order $n$) and $p$ (deg $\leq n - 1$)

Final Fields of Values for Lowest Computed $f$

Optimizing over both $p$ and $A$: Final $f(p, A)$

Is the Ratio 0.5 Attained?

Final Fields of Values for $f$ Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension $n$ and Degree $n - 1$

Optimizing the Radius Ratio Instead

---

*Sorted* final values of the radius ratio $f$, $n = 2$

Lowest values of radius ratio, $n = 2$

<table>
<thead>
<tr>
<th>$m$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>0.815</td>
<td>0.808</td>
<td>0.805</td>
<td>0.803</td>
<td>0.802</td>
<td>0.801</td>
</tr>
</tbody>
</table>

To approximate $\frac{4}{5}$ well we apparently need to allow unbounded degree, in sharp contrast to minimizing the Crouzeix ratio, where degree $m = n - 1$ is enough to attain 0.5.
Optimizing the Radius Ratio, \( n > 2 \)

### Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

#### Nonsmoothness of the Crouzeix Ratio

BFGS Experiments

Optimizing over \( A \) (order \( n \)) and \( p \) (deg \( \leq n - 1 \))

Final Fields of Values for Lowest Computed \( f \)

Optimizing over both \( p \) and \( A \): Final \( f(p, A) \)

Is the Ratio 0.5 Attained?

Final Fields of Values for \( f \) Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension \( n \) and Degree \( n - 1 \)

Optimizing the Radius Ratio Instead

Optimizing the Radius Ratio, \( n = 2 \)

---

Sorted final values of the radius ratio \( f \)
Optimizing the Radius Ratio, \( n > 2 \)

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmoothness of the Crouzeix Ratio

BFGS

Experiments

Optimizing over \( A \) (order \( n \)) and \( p \) (deg \( \leq n - 1 \))

Final Fields of Values for Lowest Computed \( f \)

Optimizing over both \( p \) and \( A \): Final \( f(p, A) \)

Is the Ratio 0.5 Attained?

Final Fields of Values for \( f \) Closest to 1

Why is the Crouzeix Ratio One?

Results for Larger Dimension \( n \) and Degree \( n - 1 \)

Optimizing the Radius Ratio Instead

Optimizing the Radius Ratio, \( n = 2 \)

Sorted final values of the radius ratio \( f \)

The space is even more littered with ice-cream cone local minimizers than earlier!
Nonsmooth Analysis of the Crouzeix Ratio

The Clarke Subdifferential
The Gradient or Subgradients of the Crouzeix Ratio
Regularity
Simplest Case where Crouzeix Ratio is Nonsmooth
(\( \hat{c}, \hat{A} \)) is a Nonsmooth Stationary Point of \( f(\cdot, \cdot) \)
The General Case
(\( \hat{c}, \hat{A} \)) is a Nonsmooth Stationary Point of \( f(\cdot, \cdot) \)
Is the Crouzeix Ratio Globally Clarke Regular?
Assume $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is locally Lipschitz, and let $D = \{ x \in \mathbb{R}^n : h \text{ is differentiable at } x \}$. 

The Clarke Subdifferential

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmooth Analysis of the Crouzeix Ratio

The Clarke Subdifferential

The Gradient or Subgradients of the Crouzeix Ratio

Regularity

Simplest Case where Crouzeix Ratio is Nonsmooth

$(\hat{c}, \hat{A})$ is a Nonsmooth Stationary Point of $f(\cdot, \cdot)$

The General Case

$(\hat{c}, \hat{A})$ is a Nonsmooth Stationary Point of $f(\cdot, \cdot)$

Is the Crouzeix Ratio Globally Clarke Regular?

Concluding Remarks
The Clarke Subdifferential

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmooth Analysis of the Crouzeix Ratio

The Clarke Subdifferential

The Gradient or Subgradients of the Crouzeix Ratio

Regularity

Simplest Case where Crouzeix Ratio is Nonsmooth

(\hat{c}, \hat{A}) is a Nonsmooth Stationary Point of $f(\cdot, \cdot)$

The General Case

(\hat{c}, \hat{A}) is a Nonsmooth Stationary Point of $f(\cdot, \cdot)$

Is the Crouzeix Ratio Globally Clarke Regular?

Assume $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is locally Lipschitz, and let $D = \{x \in \mathbb{R}^n : h \text{ is differentiable at } x\}$.

Rademacher’s Theorem: $\mathbb{R}^n \setminus D$ has measure zero.
The Clarke Subdifferential

Assume $h : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz, and let $D = \{ x \in \mathbb{R}^n : h \text{ is differentiable at } x \}$. Rademacher’s Theorem: $\mathbb{R}^n \setminus D$ has measure zero.

The Clarke subdifferential, or set of subgradients, of $h$ at $\bar{x}$ is

$$\partial h(\bar{x}) = \text{conv} \left\{ \lim_{x \to \bar{x}, x \in D} \nabla h(x) \right\}. $$
Assume $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is locally Lipschitz, and let $D = \{ x \in \mathbb{R}^n : h \text{ is differentiable at } x \}$. Rademacher’s Theorem: $\mathbb{R}^n \setminus D$ has measure zero.

The Clarke subdifferential, or set of subgradients, of $h$ at $\bar{x}$ is

$$\partial h(\bar{x}) = \text{conv} \left\{ \lim_{x \rightarrow \bar{x}, x \in D} \nabla h(x) \right\}.$$ 

F.H. Clarke, 1973 (he used the name “generalized gradient”).
The Clarke Subdifferential

Assume \( h : \mathbb{R}^n \to \mathbb{R} \) is locally Lipschitz, and let \( D = \{ x \in \mathbb{R}^n : h \text{ is differentiable at } x \} \).

Rademacher’s Theorem: \( \mathbb{R}^n \setminus D \) has measure zero.

The Clarke subdifferential, or set of subgradients, of \( h \) at \( \bar{x} \) is

\[
\partial h(\bar{x}) = \text{conv} \left\{ \lim_{x \to \bar{x}, x \in D} \nabla h(x) \right\}.
\]

F.H. Clarke, 1973 (he used the name “generalized gradient”).

If \( h \) is continuously differentiable at \( \bar{x} \), then \( \partial h(\bar{x}) = \{ \nabla h(\bar{x}) \} \).
The Clarke Subdifferential

Assume \( h : \mathbb{R}^n \to \mathbb{R} \) is locally Lipschitz, and let \( D = \{ x \in \mathbb{R}^n : h \text{ is differentiable at } x \} \).

Rademacher’s Theorem: \( \mathbb{R}^n \setminus D \) has measure zero.

The Clarke subdifferential, or set of subgradients, of \( h \) at \( \bar{x} \) is

\[
\partial h(\bar{x}) = \text{conv} \left\{ \lim_{x \to \bar{x}, x \in D} \nabla h(x) \right\}.
\]

F.H. Clarke, 1973 (he used the name “generalized gradient”).

If \( h \) is continuously differentiable at \( \bar{x} \), then \( \partial h(\bar{x}) = \{ \nabla h(\bar{x}) \} \).

If \( h \) is convex, \( \partial h \) is the subdifferential of convex analysis.
The Clarke Subdifferential

Assume $h : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz, and let $D = \{ x \in \mathbb{R}^n : h \text{ is differentiable at } x \}$.

Rademacher’s Theorem: $\mathbb{R}^n \setminus D$ has measure zero.

The Clarke subdifferential, or set of subgradients, of $h$ at $\bar{x}$ is

$$\partial h(\bar{x}) = \text{conv} \left\{ \lim_{x \to \bar{x}, x \in D} \nabla h(x) \right\}.$$

F.H. Clarke, 1973 (he used the name “generalized gradient”).

If $h$ is continuously differentiable at $\bar{x}$, then $\partial h(\bar{x}) = \{ \nabla h(\bar{x}) \}$.

If $h$ is convex, $\partial h$ is the subdifferential of convex analysis.

We say $\bar{x}$ is Clarke stationary for $h$ if $0 \in \partial h(\bar{x})$ (a nonsmooth stationary point if $\in \partial h(\bar{x})$ contains more than one vector).
The Clarke Subdifferential

Assume $h : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz, and let $D = \{ x \in \mathbb{R}^n : h \text{ is differentiable at } x \}$.

Rademacher’s Theorem: $\mathbb{R}^n \setminus D$ has measure zero.

The Clarke subdifferential, or set of subgradients, of $h$ at $\bar{x}$ is

$$\partial h(\bar{x}) = \text{conv} \left\{ \lim_{x \to \bar{x}, x \in D} \nabla h(x) \right\}.$$

F.H. Clarke, 1973 (he used the name “generalized gradient”).

If $h$ is continuously differentiable at $\bar{x}$, then $\partial h(\bar{x}) = \{ \nabla h(\bar{x}) \}$.

If $h$ is convex, $\partial h$ is the subdifferential of convex analysis.

We say $\bar{x}$ is Clarke stationary for $h$ if $0 \in \partial h(\bar{x})$ (a nonsmooth stationary point if $\in \partial h(\bar{x})$ contains more than one vector).

Clarke stationarity is a necessary condition for local or global optimality.
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of
\[
\max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta.
\]
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of
\[
\max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta.
\]

- the gradient of \( p(z_\theta) \) w.r.t. the coefficients of \( p \)
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of
\[
\max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta.
\]

- the gradient of \( p(z_\theta) \) w.r.t. the coefficients of \( p \)
- the gradient of \( p(z_\theta) \) w.r.t. \( z_\theta \)
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of
\[ \max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta. \]

- the gradient of \( p(z_\theta) \) w.r.t. the coefficients of \( p \)
- the gradient of \( p(z_\theta) \) w.r.t. \( z_\theta \)
- the gradient of \( z_\theta(A) = v_\theta^* A v_\theta \) w.r.t. \( A \)
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of

\[ \max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta. \]

- the gradient of \( p(z_\theta) \) w.r.t. the coefficients of \( p \)
- the gradient of \( p(z_\theta) \) w.r.t. \( z_\theta \)
- the gradient of \( z_\theta(A) = v_\theta^* A v_\theta \) w.r.t. \( A \)

If the max of \( |p(z_\theta)| \) is attained by a unique point \( \hat{\theta} \), then all these are evaluated at \( \hat{\theta} \) and combined with the gradient of \( |\cdot| \) to obtain the gradient of the numerator.
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of
\[
\max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta.
\]

- the gradient of \( p(z_\theta) \) w.r.t. the coefficients of \( p \)
- the gradient of \( p(z_\theta) \) w.r.t. \( z_\theta \)
- the gradient of \( z_\theta(A) = v_\theta^* A v_\theta \) w.r.t. \( A \)

If the max of \( |p(z_\theta)| \) is attained by a unique point \( \hat{\theta} \), then all these are evaluated at \( \hat{\theta} \) and combined with the gradient of \(| \cdot |\) to obtain the gradient of the numerator.

Otherwise, need to take the convex hull of these gradients over all maximizing \( \theta \) to get the subgradients of the numerator.
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of
\[
\max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta.
\]

- the gradient of \( p(z_\theta) \) w.r.t. the coefficients of \( p \)
- the gradient of \( p(z_\theta) \) w.r.t. \( z_\theta \)
- the gradient of \( z_\theta(A) = v_\theta^* A v_\theta \) w.r.t. \( A \)

If the max of \( |p(z_\theta)| \) is attained by a unique point \( \hat{\theta} \), then all these are evaluated at \( \hat{\theta} \) and combined with the gradient of \( | \cdot | \) to obtain the gradient of the numerator.

Otherwise, need to take the convex hull of these gradients over all maximizing \( \theta \) to get the subgradients of the numerator.

For the denominator, combine:
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of

\[
\max_{\theta \in [0,2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta.
\]

- the gradient of \( p(z_\theta) \) w.r.t. the coefficients of \( p \)
- the gradient of \( p(z_\theta) \) w.r.t. \( z_\theta \)
- the gradient of \( z_\theta(A) = v_\theta^* A v_\theta \) w.r.t. \( A \)

If the \( \max \) of \(|p(z_\theta)|\) is attained by a unique point \( \hat{\theta} \), then all these are evaluated at \( \hat{\theta} \) and combined with the gradient of \(|\cdot|\) to obtain the gradient of the numerator.

Otherwise, need to take the \textit{convex hull} of these gradients over all maximizing \( \theta \) to get the subgradients of the numerator.

For the denominator, combine:

- the gradient or subgradients of the 2-norm (maximum singular value) of a matrix (involves the singular vectors)
For the numerator, we need the variational properties of
\[
\max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta.
\]
- the gradient of \(p(z_\theta)\) w.r.t. the coefficients of \(p\)
- the gradient of \(p(z_\theta)\) w.r.t. \(z_\theta\)
- the gradient of \(z_\theta(A) = v_\theta^* A v_\theta\) w.r.t. \(A\)

If the max of \(|p(z_\theta)|\) is attained by a unique point \(\hat{\theta}\), then all
these are evaluated at \(\hat{\theta}\) and combined with the gradient of \(|\cdot|\)
to obtain the gradient of the numerator.

Otherwise, need to take the \textit{convex hull} of these gradients over
all maximizing \(\theta\) to get the subgradients of the numerator.

For the denominator, combine:
- the gradient or subgradients of the 2-norm (maximum singular value) of a matrix (involves the singular vectors)
- the gradient of the matrix polynomial \(p(A)\) w.r.t. \(A\) (involves differentiating \(A^k\) w.r.t. \(A\), resulting in Kronecker products).
The Gradient or Subgradients of the Crouzeix Ratio

For the numerator, we need the variational properties of

$$\max_{\theta \in [0, 2\pi]} |p(z_\theta)| \quad \text{where} \quad z_\theta = v_\theta^* A v_\theta.$$ 

- the gradient of $p(z_\theta)$ w.r.t. the coefficients of $p$
- the gradient of $p(z_\theta)$ w.r.t. $z_\theta$
- the gradient of $z_\theta(A) = v_\theta^* A v_\theta$ w.r.t. $A$

If the max of $|p(z_\theta)|$ is attained by a unique point $\hat{\theta}$, then all these are evaluated at $\hat{\theta}$ and combined with the gradient of $| \cdot |$ to obtain the gradient of the numerator.

Otherwise, need to take the convex hull of these gradients over all maximizing $\theta$ to get the subgradients of the numerator.

For the denominator, combine:

- the gradient or subgradients of the 2-norm (maximum singular value) of a matrix (involves the singular vectors)
- the gradient of the matrix polynomial $p(A)$ w.r.t. $A$ (involves differentiating $A^k$ w.r.t. $A$, resulting in Kronecker products).

Finally, use the quotient rule.
A directionally differentiable, locally Lipschitz function $h$ is *regular* (in the sense of Clarke, 1975) near a point $x$ when its directional derivative $x \mapsto h'(x; d)$ is upper semicontinuous there for every fixed direction $d$. 
A directionally differentiable, locally Lipschitz function $h$ is *regular* (in the sense of Clarke, 1975) near a point $x$ when its directional derivative $x \mapsto h'(x; d)$ is upper semicontinuous there for every fixed direction $d$.

In this case $0 \in \partial h(x)$ is equivalent to the first-order optimality condition $h'(x, d) \geq 0$ for all directions $d$. 

---

**Concluding Remarks**
A directionally differentiable, locally Lipschitz function $h$ is \textit{regular} (in the sense of Clarke, 1975) near a point $x$ when its directional derivative $x \mapsto h'(x; d)$ is upper semicontinuous there for every fixed direction $d$.

In this case $0 \in \partial h(x)$ is equivalent to the first-order optimality condition $h'(x, d) \geq 0$ for all directions $d$.

- All convex functions are regular
Regularity

A directionally differentiable, locally Lipschitz function $h$ is **regular** (in the sense of Clarke, 1975) near a point $x$ when its directional derivative $x \mapsto h'(x; d)$ is upper semicontinuous there for every fixed direction $d$.

In this case $0 \in \partial h(x)$ is equivalent to the first-order optimality condition $h'(x, d) \geq 0$ for all directions $d$.

- All convex functions are regular
- All continuously differentiable functions are regular

Concluding Remarks
A directionally differentiable, locally Lipschitz function $h$ is \textit{regular} (in the sense of Clarke, 1975) near a point $x$ when its directional derivative $x \mapsto h'(x; d)$ is upper semicontinuous there for every fixed direction $d$.

In this case $0 \in \partial h(x)$ is equivalent to the first-order optimality condition $h'(x, d) \geq 0$ for all directions $d$.

- All convex functions are regular
- All continuously differentiable functions are regular
- Nonsmooth concave functions, e.g. $h(x) = -|x|$, are not regular.
Optimize over complex monic linear polynomials \( p(\zeta) \equiv c + \zeta \) and complex matrices with order \( n = 2 \). Let \( f(p, A) \equiv f(c, A) \), where now \( f : \mathbb{C} \times \mathbb{C}^{2 \times 2} \rightarrow \mathbb{R} \).
Simplest Case where Crouzeix Ratio is Nonsmooth

Optimize over complex monic linear polynomials \( p(\zeta) \equiv c + \zeta \) and complex matrices with order \( n = 2 \). Let \( f(p, A) \equiv f(c, A) \), where now \( f : \mathbb{C} \times \mathbb{C}^{2 \times 2} \rightarrow \mathbb{R} \).

Let \( \hat{c} = 0 \) (\( \hat{p}(\zeta) = \zeta \)) and \( \hat{A} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \), so \( W(\hat{A}) = D \), the unit disk, and hence \( |p(\zeta)| \) is maximized everywhere on the unit circle, with \( f(\zeta) \) nonsmooth at \( (\hat{c}, \hat{A}) \) and \( f(\hat{c}, \hat{A}) = 1/2 \).
Optimize over complex monic linear polynomials $p(\zeta) \equiv c + \zeta$ and complex matrices with order $n = 2$. Let $f(p, A) \equiv f(c, A)$, where now $f : \mathbb{C} \times \mathbb{C}^{2 \times 2} \rightarrow \mathbb{R}$.

Let $\hat{c} = 0$ ($\hat{p}(\zeta) = \zeta$) and $\hat{A} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$, so $W(\hat{A}) = \mathcal{D}$, the unit disk, and hence $|p(\zeta)|$ is maximized everywhere on the unit circle, with $f$ nonsmooth at $(\hat{c}, \hat{A})$ and $f(\hat{c}, \hat{A}) = 1/2$.

**Theorem.** The Crouzeix ratio $f$ is regular at $(\hat{c}, \hat{A})$, with

$$\partial f(\hat{c}, \hat{A}) = \text{conv}_{\theta \in [0, 2\pi)} \left\{ \left( \frac{1}{2} e^{-i\theta}, \frac{1}{4} \begin{bmatrix} e^{-i\theta} & 0 \\ e^{-2i\theta} & e^{-i\theta} \end{bmatrix} \right) \right\}$$
Corollary.

\[ 0 \in \partial f(\hat{c}, \hat{A}) \]
Corollary.

$$0 \in \partial f(\hat{c}, \hat{A})$$

Proof: the vectors inside the convex hull defined by $\theta = 0$, $2\pi/3$ and $4\pi/3$ sum to zero.
\((\hat{c}, \hat{A})\) is a Nonsmooth Stationary Point of \(f(\cdot, \cdot)\)

**Corollary.**

\[0 \in \partial f(\hat{c}, \hat{A})\]

Proof: the vectors inside the convex hull defined by \(\theta = 0, 2\pi/3\) and \(4\pi/3\) sum to zero.

Actually, we knew this must be true as Crouzeix’s conjecture is known to hold for \(n = 2\), and hence \((\hat{c}, \hat{A})\) is a global minimizer of \(f(\cdot, \cdot)\), but we can extend the result to larger values of \(m, n\), for which we don’t know whether the conjecture holds.
The General Case

Optimize over complex polynomials \( p(\zeta) \equiv c_0 + \cdots + c_m \zeta^m \) and complex matrices with order \( n \). Let \( f(p, A) \equiv f(c, A) \), where \( f : \mathbb{C}^{m+1} \times \mathbb{C}^{n \times n} \rightarrow \mathbb{R} \).
The General Case

Optimize over complex polynomials $p(\zeta) \equiv c_0 + \cdots + c_m \zeta^m$ and complex matrices with order $n$. Let $f(p, A) \equiv f(c, A)$, where $f : \mathbb{C}^{m+1} \times \mathbb{C}^{n \times n} \to \mathbb{R}$.

Let $\hat{c} = [0, 0, \ldots, 1]$, corresponding to the polynomial $\zeta^{n-1}$, and $\hat{A} = \Xi_n$, the CCC matrix of order $n$ so $W(\hat{A}) = D$, the unit disk, and hence $f(\hat{c}, \hat{A}) = 1/2$. 

Concluding Remarks
The General Case

Optimize over complex polynomials \( p(\zeta) \equiv c_0 + \cdots + c_m \zeta^m \) and complex matrices with order \( n \). Let \( f(p, A) \equiv f(c, A) \), where \( f : \mathbb{C}^{m+1} \times \mathbb{C}^{n \times n} \to \mathbb{R} \).

Let \( \hat{c} = [0, 0, \ldots, 1] \), corresponding to the polynomial \( z^{n-1} \), and \( \hat{A} = \Xi_n \), the CCC matrix of order \( n \) so \( W(\hat{A}) = D \), the unit disk, and hence \( f(\hat{c}, \hat{A}) = 1/2 \).

**Theorem.** The Crouzeix ratio on \( (c, A) \in \mathbb{C}^{m+1} \times \mathbb{C}^{n \times n} \) is regular at \( (\hat{c}, \hat{A}) \) with

\[
\partial f(\hat{c}, \hat{A}) = \text{conv}_{\theta \in [0, 2\pi)} \left\{ (y_\theta, Y_\theta) \right\}
\]

where

\[
y_\theta = \frac{1}{2} \left[ z^m, z^{m-1}, \ldots, z, 0 \right]^T
\]

and \( Y_\theta \) \( n \times n \) matrix

\[
Y_\theta = \frac{1}{4} \begin{bmatrix}
z & 0 & \sqrt{2}z^{-1} & \sqrt{2}z^{-2} & \cdots & \sqrt{2}z^{3-n} & z^{2-n} \\
\sqrt{2}z^2 & 2z & 0 & 2z^{-1} & \cdots & 2z^{4-n} & \sqrt{2}z^{3-n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\sqrt{2}z^{n-2} & 2z^{n-3} & 2z^{n-4} & 2z^{n-5} & \cdots & 0 & \sqrt{2}z \\
\sqrt{2}z^{n-1} & 2z^{n-2} & 2z^{n-3} & 2z^{n-4} & \cdots & 2z & 0 \\
z^n & \sqrt{2}z^{n-1} & \sqrt{2}z^{n-2} & \sqrt{2}z^{n-3} & \cdots & \sqrt{2}z^2 & z
\end{bmatrix}
\]

with \( z = e^{-i\theta} \).
The pair $\hat{c}, \hat{A}$ is a Nonsmooth Stationary Point of $f(\cdot, \cdot)$.

**Corollary.**

$$0 \in \partial f(\hat{c}, \hat{A})$$

so, for any $n$, the pair $\hat{c}, \hat{A}$ is a nonsmooth stationary point of $f$. 

---

**Computational Setup**

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmooth Analysis of the Crouzeix Ratio

The Clarke Subdifferential

The Gradient or Subgradients of the Crouzeix Ratio

Regularity

Simplest Case where Crouzeix Ratio is Nonsmooth

$(\hat{c}, \hat{A})$ is a Nonsmooth Stationary Point of $f(\cdot, \cdot)$

The General Case

$(\hat{c}, \hat{A})$ is a Nonsmooth Stationary Point of $f(\cdot, \cdot)$

Is the Crouzeix Ratio Globally Clarke Regular?

---

**Concluding Remarks**
\((\hat{c}, \hat{A})\) is a Nonsmooth Stationary Point of \(f(\cdot, \cdot)\)

**Corollary.**

\[ 0 \in \partial f(\hat{c}, \hat{A}) \]

so, for any \(n\), the pair \((\hat{c}, \hat{A})\) is a nonsmooth stationary point of \(f\).

**Proof.** The convex combination

\[
\frac{1}{n+1} \sum_{k=0}^{n} \left( \frac{y_{2k\pi}/(n+1)}{Y_{2k\pi}/(n+1)} \right)
\]

is zero.
**Corollary.**

\[ 0 \in \partial f(\hat{c}, \hat{A}) \]

so, for any \( n \), the pair \((\hat{c}, \hat{A})\) is a nonsmooth stationary point of \( f \).

**Proof.** The convex combination

\[
\frac{1}{n + 1} \sum_{k=0}^{n} \left( y_{2k\pi/(n+1)}, Y_{2k\pi/(n+1)} \right)
\]

is zero.

This is a necessary condition for \((\hat{c}, \hat{A})\) to be a local (or global) minimizer of \( f \) on \( \mathbb{R}^{m+1} \times \mathbb{R}^{n \times n} \). This is a new result for \( n > 2 \).
\((\hat{c}, \hat{A})\) is a Nonsmooth Stationary Point of \(f(\cdot, \cdot)\)

**Corollary.**

\[0 \in \partial f(\hat{c}, \hat{A})\]

so, for any \(n\), the pair \((\hat{c}, \hat{A})\) is a nonsmooth stationary point of \(f\).

**Proof.** The convex combination

\[
\frac{1}{n + 1} \sum_{k=0}^{n} \left( y_{2k\pi/(n+1)}, Y_{2k\pi/(n+1)} \right)
\]

is zero.

This is a necessary condition for \((\hat{c}, \hat{A})\) to be a local (or global) minimizer of \(f\) on \(\mathbb{R}^{m+1} \times \mathbb{R}^{n \times n}\). This is a new result for \(n > 2\). And by regularity, it implies that the directional derivative \(f'(\cdot, d) \geq 0\) for all directions \(d\).
Is the Crouzeix Ratio Globally Clarke Regular?

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmooth Analysis of the Crouzeix Ratio

The Clarke Subdifferential

The Gradient or Subgradients of the Crouzeix Ratio

Regularity

Simplest Case where Crouzeix Ratio is Nonsmooth

\((\hat{c}, \hat{A})\) is a Nonsmooth Stationary Point of \(f(\cdot, \cdot)\)

The General Case

\((\hat{c}, \hat{A})\) is a Nonsmooth Stationary Point of \(f(\cdot, \cdot)\)

Concluding Remarks
Is the Crouzeix Ratio Globally Clarke Regular?

No. Let \( \tilde{p}(\zeta) = \zeta \) and

\[
\tilde{A} = \begin{bmatrix}
0 & \sqrt{2} & 0 \\
0 & 0 & \sqrt{2} \\
0 & 0 & 0
\end{bmatrix}
\]

for which \( W(\tilde{A}) \) is a disk and \( f(\tilde{p}, \tilde{A}) = 1/\sqrt{2} \). The Crouzeix ratio \( f \) is not regular at \( (\tilde{p}, \tilde{A}) \).
Is the Crouzeix Ratio Globally Clarke Regular?

No. Let \( \tilde{p}(\zeta) = \zeta \) and

\[
\tilde{A} = \begin{bmatrix}
0 & \sqrt{2} & 0 \\
0 & 0 & \sqrt{2} \\
0 & 0 & 0
\end{bmatrix}
\]

for which \( W(\tilde{A}) \) is a disk and \( f(\tilde{p}, \tilde{A}) = 1/\sqrt{2} \).

The Crouzeix ratio \( f \) is not regular at \( (\tilde{p}, \tilde{A}) \).

Plot of the denominator \( \beta \), the numerator \( \tau \) and the Crouzeix ratio \( f \) evaluated at \( (\tilde{p}, \tilde{A} + t\tilde{A}^2), \ t \in [-2, 2] \).
Concluding Remarks
Summary

Minimizing the Crouzeix ratio $f$ over $p$ and $A$, BFGS mostly converged either to *nonsmooth stationary* values of 0.5 associated with the CCC matrix (with field of values a disk), or *smooth stationary* values of 1 (with “ice cream cone” fields of values).
Minimizing the Crouzeix ratio $f$ over $p$ and $A$, BFGS mostly converged either to *nonsmooth stationary* values of $0.5$ associated with the CCC matrix (with field of values a disk), or *smooth stationary* values of $1$ (with “ice cream cone” fields of values).

Both Chebfun and BFGS perform remarkably reliably despite nonsmoothness that can occur either in the boundary of the field of values (w.r.t. the complex plane) or in the Crouzeix ratio $f$ (w.r.t the polynomial-matrix space).
Summary

Minimizing the Crouzeix ratio $f$ over $p$ and $A$, BFGS mostly converged either to non-smooth stationary values of 0.5 associated with the CCC matrix (with field of values a disk), or smooth stationary values of 1 (with “ice cream cone” fields of values).

Both Chebfun and BFGS perform remarkably reliably despite nonsmoothness that can occur either in the boundary of the field of values (w.r.t. the complex plane) or in the Crouzeix ratio $f$ (w.r.t. the polynomial-matrix space).

Using nonsmooth variational analysis, we proved regularity and Clarke stationarity of the Crouzeix ratio, with value 0.5, at pairs $(\hat{p}, \hat{A})$, where $\hat{p}$ is the monomial $\zeta^{n-1}$ and $\hat{A}$ is a CCC matrix of order $n$, a necessary condition for local or global optimality.
Minimizing the Crouzeix ratio $f$ over $p$ and $A$, BFGS mostly converged either to non-smooth stationary values of 0.5 associated with the CCC matrix (with field of values a disk), or smooth stationary values of 1 (with “ice cream cone” fields of values). Both Chebfun and BFGS perform remarkably reliably despite nonsmoothness that can occur either in the boundary of the field of values (w.r.t. the complex plane) or in the Crouzeix ratio $f$ (w.r.t the polynomial-matrix space).

Using nonsmooth variational analysis, we proved regularity and Clarke stationarity of the Crouzeix ratio, with value 0.5, at pairs $(\hat{p}, \hat{A})$, where $\hat{p}$ is the monomial $\zeta^{n-1}$ and $\hat{A}$ is a CCC matrix of order $n$, a necessary condition for local or global optimality.

We also found $(\tilde{p}, \tilde{A})$ for which the Crouzeix ratio is not regular.
Minimizing the Crouzeix ratio $f$ over $p$ and $A$, BFGS mostly converged either to \textit{nonsmooth stationary} values of 0.5 associated with the CCC matrix (with field of values a disk), or \textit{smooth stationary} values of 1 (with “ice cream cone” fields of values).

Both Chebfun and BFGS perform remarkably reliably despite nonsmoothness that can occur either in the boundary of the field of values (w.r.t. the complex plane) or in the Crouzeix ratio $f$ (w.r.t the polynomial-matrix space).

Using nonsmooth variational analysis, we proved \textit{regularity} and \textit{Clarke stationarity} of the Crouzeix ratio, with value 0.5, at pairs $(\hat{p}, \hat{A})$, where $\hat{p}$ is the monomial $\zeta^{n-1}$ and $\hat{A}$ is a CCC matrix of order $n$, a necessary condition for local or global optimality.

We also found $(\tilde{p}, \tilde{A})$ for which the Crouzeix ratio is \textit{not regular}.

The results strongly support Crouzeix’s conjecture: the globally minimal value of the Crouzeix ratio $f(p, A)$ is 0.5.
**Computational:** extend our nonsmooth optimization of the Crouzeix ratio from polynomials to analytic functions, using the Blaschke product and conformal map techniques discussed by Anne.
Challenges

**Computational:** extend our nonsmooth optimization of the Crouzeix ratio from polynomials to analytic functions, using the Blaschke product and conformal map techniques discussed by Anne.

Doable, but complicated, especially if we want to use gradients.
Challenges

**Computational:** extend our nonsmooth optimization of the Crouzeix ratio from polynomials to analytic functions, using the Blaschke product and conformal map techniques discussed by Anne.

Doable, but complicated, especially if we want to use gradients.

**Theoretical:** could the Clarke stationarity result be extended to include analytic functions?
**Challenges**

**Computational:** extend our nonsmooth optimization of the Crouzeix ratio from polynomials to analytic functions, using the Blaschke product and conformal map techniques discussed by Anne.

Doable, but complicated, especially if we want to use gradients.

**Theoretical:** could the Clarke stationarity result be extended to include analytic functions?

Possibly, but much more complicated, especially as the optimal families seem to be hard to describe even for $n = 4$ (Crouzeix).
**Challenges**

**Computational**: extend our nonsmooth optimization of the Crouzeix ratio from polynomials to analytic functions, using the Blaschke product and conformal map techniques discussed by Anne.

Doable, but complicated, especially if we want to use gradients.

**Theoretical**: could the Clarke stationarity result be extended to include analytic functions?

Possibly, but much more complicated, especially as the optimal families seem to be hard to describe even for $n = 4$ (Crouzeix).

**Theoretical**: could the Clarke stationarity result be extended to prove local optimality of the $n \times n$ CCC matrix and associated polynomial $\zeta^{n-1}$?
Challenges

**Computational:** extend our nonsmooth optimization of the Crouzeix ratio from polynomials to analytic functions, using the Blaschke product and conformal map techniques discussed by Anne.

Doable, but complicated, especially if we want to use gradients.

**Theoretical:** could the Clarke stationarity result be extended to include analytic functions?

Possibly, but much more complicated, especially as the optimal families seem to be hard to describe even for $n = 4$ (Crouzeix).

**Theoretical:** could the Clarke stationarity result be extended to prove *local* optimality of the $n \times n$ CCC matrix and associated polynomial $\zeta^{n-1}$?

Possibly, but this will also be complicated as it involves (at least) second derivatives.
Challenges

**Computational:** extend our nonsmooth optimization of the Crouzeix ratio from polynomials to analytic functions, using the Blaschke product and conformal map techniques discussed by Anne.

Doable, but complicated, especially if we want to use gradients.

**Theoretical:** could the Clarke stationarity result be extended to include analytic functions?

Possibly, but much more complicated, especially as the optimal families seem to be hard to describe even for \( n = 4 \) (Crouzeix).

**Theoretical:** could the Clarke stationarity result be extended to prove *local* optimality of the \( n \times n \) CCC matrix and associated polynomial \( \zeta^{n-1} \)?

Possibly, but this will also be complicated as it involves (at least) second derivatives.

**Holy Grail:** could the variational analytic approach be used to prove *global* optimality?
Computational: extend our nonsmooth optimization of the Crouzeix ratio from polynomials to analytic functions, using the Blaschke product and conformal map techniques discussed by Anne.

Doable, but complicated, especially if we want to use gradients.

Theoretical: could the Clarke stationarity result be extended to include analytic functions?

Possibly, but much more complicated, especially as the optimal families seem to be hard to describe even for $n = 4$ (Crouzeix).

Theoretical: could the Clarke stationarity result be extended to prove local optimality of the $n \times n$ CCC matrix and associated polynomial $\zeta^{n-1}$?

Possibly, but this will also be complicated as it involves (at least) second derivatives.

Holy Grail: could the variational analytic approach be used to prove global optimality?

Not at all clear how, but if the conjecture is not true, there exists a matrix $A$ and analytic function $p$ for which the Crouzeix ratio is less than 0.5 (and greater than or equal to $1/(1 + \sqrt{2})$). Either there must be a pair $(p, A)$ at which the Crouzeix ratio is globally minimized, and therefore Clarke stationary, or the minimal value is attained only in the limit.
The first organizer said
The first organizer said

The second organizer said
A Probability Story

Computational Setup

Nonsmooth Optimization of the Crouzeix Ratio

Nonsmooth Analysis of the Crouzeix Ratio

Concluding Remarks

Summary

Challenges

A Probability Story

Our Papers

The first organizer said

The second organizer said

The third organizer said
A Probability Story

- The first organizer said
- The second organizer said
- The third organizer said
- The genie said
A Probability Story

- The first organizer said
- The second organizer said
- The third organizer said
- The genie said
- The organizers responded
Our Papers

A. Greenbaum and M.L. Overton
*Investigation of Crouzeix’s Conjecture via Nonsmooth Optimization*
Linear Alg. Appl., 2017

A. Greenbaum, A.S. Lewis and M.L. Overton
*Variational Analysis of the Crouzeix Ratio*

N. Guglielmi, M.L. Overton and G.W. Stewart
*An Efficient Algorithm for Computing the Generalized Null Space Decomposition*

A.S. Lewis and M.L. Overton
*Nonsmooth Optimization via Quasi-Newton Methods*
Our Papers

A. Greenbaum and M.L. Overton
Investigation of Crouzeix’s Conjecture via Nonsmooth Optimization
Linear Alg. Appl., 2017

A. Greenbaum, A.S. Lewis and M.L. Overton
Variational Analysis of the Crouzeix Ratio

N. Guglielmi, M.L. Overton and G.W. Stewart
An Efficient Algorithm for Computing the Generalized Null Space Decomposition

A.S. Lewis and M.L. Overton
Nonsmooth Optimization via Quasi-Newton Methods

Many thanks to Nick Trefethen for suggesting studying this problem using Chebfun and Anthony Austin and Nick Hale for Chebfun support.