Variational Analysis of Convexly Generated Spectral Max Functions

James V Burke
Mathematics, University of Washington

Joint work with
Julie Eation (UW), Adrian Lewis (Cornell),
Michael Overton (NYU)

Universität Bonn
Haursdorff Center for Mathematics
May 22, 2017

Convexly Generated Spectral Max Functions

$\mathfrak{f} : \mathbb{C}^{n\times n} \to \mathbb{R} \cup \{+\infty\} =: \overline{\mathbb{R}}$

$$\mathfrak{f}(X) := \max\{f(\lambda) \mid \lambda \in \mathbb{C} \text{ and } \det(\lambda I - X) = 0\},$$

where $f : \mathbb{C} \to \overline{\mathbb{R}}$ is closed, proper, convex.

Convexly Generated Spectral Max Functions

$\mathfrak{f} : \mathbb{C}^{n \times n} \to \mathbb{R} \cup \{+\infty\} =: \overline{\mathbb{R}}$

$$\mathfrak{f}(X) := \max\{f(\lambda) \,|\, \lambda \in \mathbb{C} \text{ and } \det(\lambda I - X) = 0\},$$

where $f : \mathbb{C} \to \overline{\mathbb{R}}$ is closed, proper, convex.

Spectral abscissa: $f(\cdot) = \mathrm{Re}(\cdot)$ and we write $\mathfrak{f} = \alpha$.

Convexly Generated Spectral Max Functions

$\mathfrak{f} : \mathbb{C}^{n \times n} \to \mathbb{R} \cup \{+\infty\} =: \overline{\mathbb{R}}$

$$\mathfrak{f}(X) := \max\{f(\lambda) \mid \lambda \in \mathbb{C} \text{ and } \det(\lambda I - X) = 0\},$$

where $f : \mathbb{C} \to \overline{\mathbb{R}}$ is closed, proper, convex.

Spectral abscissa: $f(\cdot) = \mathrm{Re}(\cdot)$ and we write $\mathfrak{f} = \alpha$.

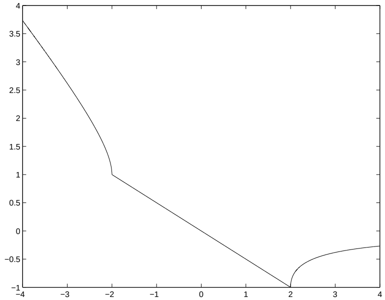
Spectral radius: $f(\cdot) = |\cdot|$ and we write $\mathfrak{f} = \rho$.

Example: Damped Oscillator: $w'' + \mu w' + w = 0$

$$u' = \begin{bmatrix} 0 & 1 \\ -1 & -\mu \end{bmatrix} u, \qquad u = \begin{pmatrix} w \\ w' \end{pmatrix}$$

$$A(\mu) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \mu \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\Lambda(A(\mu)) = \left\{ \frac{-\mu \pm \sqrt{\mu^2 - 4}}{2} \right. , \quad \alpha(A(\mu)) = \begin{cases} \frac{-\mu + \sqrt{\mu^2 - 4}}{2} & , \ |\mu| > 2 \, , \\ -\mu/2 & , \ |\mu| \leq 2 \, . \end{cases}$$



Fundamentally Non-Lipschitzian

$$N(\varepsilon) = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & & \\ & & & & & & \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \\ \epsilon & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix}_{n\times n}$$

$$\det[\lambda I - N(\varepsilon)] \ = \ \lambda^n - \varepsilon \quad \Longrightarrow \quad \lambda_k := (\varepsilon)^{1/n} \mathsf{e}^{2\pi k \mathsf{i}/n} \ k = 0, \ldots, n-1.$$

“Active eigenvalue” hypothesis

($\mathcal{H}$) For all active eigenvalues λ one of the following holds:

(A) *f is quadratic, or f is $\mathcal{C}^2$ and positive definite*
(B) $\operatorname{rspan}(\partial f(\lambda)) = \mathbb{C}$

“Active eigenvalue” hypothesis

($\mathcal{H}$) For all active eigenvalues λ one of the following holds:

(A) *f is quadratic, or f is $\mathcal{C}^2$ and positive definite*
(B) $\operatorname{rspan}(\partial f(\lambda)) = \mathbb{C}$

For $\phi : \mathbb{C} \to \overline{\mathbb{R}}$, define $\tilde{\phi} : \mathbb{R}^2 \to \overline{\mathbb{R}}$ by

$$\tilde{\phi} = \phi \circ \Theta,$$

where $\Theta : \mathbb{R}^2 \to \mathbb{C}$ is the $\mathbb{R}$-linear transformation

$$\Theta(x) = x_1 + \mathsf{i}\, x_2.$$

Note $\quad \Theta^{-1}\mu = \Theta^*\mu = \begin{bmatrix} \operatorname{Re}\mu \\ \operatorname{Im}\mu \end{bmatrix}.$

ϕ is $\mathbb{R}$-*differentiable* if $\tilde{\phi}$ is differentiable, and by the chain rule

$$\nabla\phi(\zeta) = \Theta\nabla\tilde{\phi}(\Theta^*\zeta).$$

Similarly,

$$\nabla^2\phi(\zeta) = \Theta\nabla^2\tilde{\phi}(\Theta^*\zeta)\Theta^*.$$

Key Result: B-Eaton (2016)

Suppose that $\widetilde{X} \in \mathbb{C}^{n\times n}$ *is such that hypothesis* $(\mathcal{H})$ *holds at all active eigenvalues* λ *of* $\widetilde{X}$ *with* $\partial f(\lambda) \neq \{0\}$.

Then $\mathfrak{f}$ *is subdifferentially regular at* $\widetilde{X}$ *if and only if the active eigenvalues of* $\widetilde{X}$ *are nonderogatory.*

Key Result: B-Eaton (2016)

Suppose that $\widetilde{X} \in \mathbb{C}^{n \times n}$ *is such that hypothesis* $(\mathcal{H})$ *holds at all active eigenvalues* λ *of* $\widetilde{X}$ *with* $\partial f(\lambda) \neq \{0\}$*.*

Then $\mathfrak{f}$ *is subdifferentially regular at* $\widetilde{X}$ *if and only if the active eigenvalues of* $\widetilde{X}$ *are nonderogatory.*

Nonderogatory Matrices:

An eigenvalue if *nonderogatory* if it has only one eigenvector.

Key Result: B-Eaton (2016)

Suppose that $\widetilde{X} \in \mathbb{C}^{n\times n}$ is such that hypothesis $(\mathcal{H})$ holds at all active eigenvalues λ of $\widetilde{X}$ with $\partial f(\lambda) \neq \{0\}$.

Then $\mathfrak{f}$ is subdifferentially regular at $\widetilde{X}$ if and only if the active eigenvalues of $\widetilde{X}$ are nonderogatory.

Nonderogatory Matrices:

An eigenvalue if *nonderogatory* if it has only one eigenvector.

The set of nonderogatory matrices is an open dense set in $\mathbb{C}^{n\times n}$.

Key Result: B-Eaton (2016)

Suppose that $\widetilde{X} \in \mathbb{C}^{n\times n}$ is such that hypothesis $(\mathcal{H})$ holds at all active eigenvalues λ of $\widetilde{X}$ with $\partial f(\lambda) \neq \{0\}$.

Then $\mathfrak{f}$ is subdifferentially regular at $\widetilde{X}$ if and only if the active eigenvalues of $\widetilde{X}$ are nonderogatory.

Nonderogatory Matrices:

An eigenvalue if *nonderogatory* if it has only one eigenvector.

The set of nonderogatory matrices is an open dense set in $\mathbb{C}^{n\times n}$.

If matrices are stratified by Jordan structure, then within every strata the submanifold of nonderogatory matrices has the same co-dimension as the strata.

Key Result: B-Eaton (2016)

Suppose that $\widetilde{X} \in \mathbb{C}^{n\times n}$ is such that hypothesis $(\mathcal{H})$ holds at all active eigenvalues λ of $\widetilde{X}$ with $\partial f(\lambda) \neq \{0\}$.

Then $\mathfrak{f}$ is subdifferentially regular at $\widetilde{X}$ if and only if the active eigenvalues of $\widetilde{X}$ are nonderogatory.

Nonderogatory Matrices:

An eigenvalue if *nonderogatory* if it has only one eigenvector.

The set of nonderogatory matrices is an open dense set in $\mathbb{C}^{n\times n}$.

If matrices are stratified by Jordan structure, then within every strata the submanifold of nonderogatory matrices has the same co-dimension as the strata.

Spectral abscissa case, B-Overton (2001).

Convexly Generated Polynomial Root Max Functions

The characteristic polynomial mapping $\Phi_n : \mathbb{C}^{n\times n} \to \mathbb{P}^n$:

$$\Phi_n(X)(\lambda) := \det(\lambda I - X).$$

Polynomial root max function generated by f is the mapping $\mathbf{f} : \mathbb{P}^n \to \overline{\mathbb{R}}$ defined by

$$\mathbf{f}(p) := \max\{f(\lambda) \,|\, \lambda \in \mathbb{C} \text{ and } p(\lambda) = 0\}.$$

Key Result:
Suppose that $p \in \mathbb{P}^n$ is such that ($\mathcal{H}$) holds at all active roots λ of p with $\partial f(\lambda) \neq \{0\}$.

Then $\mathbf{f}$ is subdifferentially regular at p.

Convexly Generated Polynomial Root Max Functions

The characteristic polynomial mapping $\Phi_n : \mathbb{C}^{n\times n} \to \mathbb{P}^n$:

$$\Phi_n(X)(\lambda) := \det(\lambda I - X).$$

Polynomial root max function generated by f is the mapping $\mathbf{f} : \mathbb{P}^n \to \overline{\mathbb{R}}$ defined by

$$\mathbf{f}(p) := \max\{f(\lambda) \,|\, \lambda \in \mathbb{C} \text{ and } p(\lambda) = 0\}.$$

Key Result:

Suppose that $p \in \mathbb{P}^n$ is such that $(\mathcal{H})$ holds at all active roots λ of p with $\partial f(\lambda) \neq \{0\}$.

Then $\mathbf{f}$ is subdifferentially regular at p.

The abscissa case, B-Overton (2001).

The Plan

$$\begin{array}{ccc}
\mathbb{C}^{n\times n} & \xrightarrow{\ \mathfrak{f}\ } & \overline{\mathbb{R}} \\
{\scriptstyle G(0,\cdot)}\downarrow & & \uparrow{\scriptstyle \mathbf{f}} \\
\mathbb{S}_{\tilde{p}} & \xrightarrow[F_{\tilde{p}}]{} & \mathbb{P}^{\tilde{n}}
\end{array}$$

$G(0, \widetilde{X}) := (0, e_{(n_1, \tilde{\lambda}_1)}, \ldots, e_{(n_m, \tilde{\lambda}_m)})$ are the "active factors" of the characteristic polynomial $\Phi_n(\widetilde{X})$.

The mapping $G : \mathbb{C} \times \mathbb{C}^{n\times n} \to \mathbb{S}_{\tilde{p}}$ takes a matrix $\widetilde{X} \in \mathbb{C}^{n\times n}$ to the "active factors" (of degree $\tilde{n} \leq n$) associated with its characteristic polynomial $\Phi_n(\widetilde{X})$.

$\tilde{p}$ has a local factorization based at its roots giving rise to a *factorization space* $\mathbb{S}_{\tilde{p}}$ and an associated diffeomorphism $F_{\tilde{p}} : \mathbb{S}_{\tilde{p}} \to \mathbb{P}^{\tilde{n}}$.

Apply a nonsmooth chain rule.

The Plan

$$
\begin{array}{ccc}
\mathbb{C}^{n\times n} & \xrightarrow{\mathfrak{f}} & \overline{\mathbb{R}} \\
{\scriptstyle G(0,\cdot)}\downarrow & & \uparrow{\scriptstyle \mathbf{f}} \\
\mathbb{S}_{\tilde{p}} & \xrightarrow[F_{\tilde{p}}]{} & \mathbb{P}^{\tilde{n}}
\end{array}
$$

$$\tilde{p}(\lambda) := \prod_{j=1}^{m} (\lambda - \tilde{\lambda}_j)^{n_j}$$

$$e_{(n_j,\tilde{\lambda}_j)}(\lambda) := (\lambda - \tilde{\lambda}_j)^{n_j}$$

$G(0,\widetilde{X}) := (0, e_{(n_1,\tilde{\lambda}_1)}, \ldots, e_{(n_m,\tilde{\lambda}_m)})$ are the "active factors" of the characteristic polynomial $\Phi_n(\widetilde{X})$.

The mapping $G : \mathbb{C} \times \mathbb{C}^{n\times n} \to \mathbb{S}_{\tilde{p}}$ takes a matrix $\widetilde{X} \in \mathbb{C}^{n\times n}$ to the "active factors" (of degree $\tilde{n} \leq n$) associated with its characteristic polynomial $\Phi_n(\widetilde{X})$.

$\tilde{p}$ has a local factorization based at its roots giving rise to a *factorization space* $\mathbb{S}_{\tilde{p}}$ and an associated diffeomorphism $F_{\tilde{p}} : \mathbb{S}_{\tilde{p}} \to \mathbb{P}^{\tilde{n}}$.

Apply a nonsmooth chain rule.

Inner Products

Lemma (Inner product construction)

Let $\mathbb{L}_1$ *and* $\mathbb{L}_2$ *be finite dimensional vector spaces over* $\mathbb{F} = \mathbb{C}$ *or* $\mathbb{R}$, *and suppose*

$$L : \mathbb{L}_1 \to \mathbb{L}_2 \quad \textit{is an } \mathbb{F}\textit{-linear isomorphism.}$$

If $\mathbb{L}_2$ *has inner product* $\langle \cdot, \cdot \rangle_2$, *then the bilinear form* $\mathcal{B} : \mathbb{L}_1 \times \mathbb{L}_1 \to \mathbb{F}$ *given by*

$$\mathcal{B}(x, y) := \langle Lx, Ly \rangle_2 \quad \forall\ x, y \in \mathbb{L}_1$$

is an inner product on $\mathbb{L}_1$.
The adjoint mappings $L^\star : \mathbb{L}_2 \to \mathbb{L}_1$ *and* $(L^{-1})^\star : \mathbb{L}_1 \to \mathbb{L}_2$ *with respect to the inner products* $\langle \cdot, \cdot \rangle_1 := \mathcal{B}(x, y)$ *and* $\langle \cdot, \cdot \rangle_2$ *satisfy*

$$L^\star = L^{-1} \quad \textit{and} \quad (L^{-1})^\star = L.$$

Inner Products

Lemma (Inner product construction)

Let $\mathbb{L}_1$ *and* $\mathbb{L}_2$ *be finite dimensional vector spaces over* $\mathbb{F} = \mathbb{C}$ *or* $\mathbb{R}$*, and suppose*

$$L : \mathbb{L}_1 \to \mathbb{L}_2 \qquad \textit{is an } \mathbb{F}\textit{-linear isomorphism.}$$

If $\mathbb{L}_2$ *has inner product* $\langle \cdot, \cdot \rangle_2$*, then the bilinear form* $\mathcal{B} : \mathbb{L}_1 \times \mathbb{L}_1 \to \mathbb{F}$ *given by*

$$\mathcal{B}(x, y) := \langle Lx, Ly \rangle_2 \quad \forall\, x, y \in \mathbb{L}_1$$

is an inner product on $\mathbb{L}_1$*.*
The adjoint mappings $L^\star : \mathbb{L}_2 \to \mathbb{L}_1$ *and* $(L^{-1})^\star : \mathbb{L}_1 \to \mathbb{L}_2$ *with respect to the inner products* $\langle \cdot, \cdot \rangle_1 := \mathcal{B}(x, y)$ *and* $\langle \cdot, \cdot \rangle_2$ *satisfy*

$$L^\star = L^{-1} \qquad \textit{and} \qquad (L^{-1})^\star = L.$$

Real and complex inner products, $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle^{\mathfrak{c}}$, resp.ly.

Factorization Spaces

Given $\tilde{p} \in \mathbb{M}^{\tilde{n}}$ (monic polynomials of degree at most n), write

$$\tilde{p} := \prod_{j=1}^{m} e_{(n_j, \tilde{\lambda}_j)},$$

where $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m$ are the distinct roots of $\tilde{p}$, ordered lexicographically with multiplicities $n_1, \ldots, n_m$, and the monomials $e_{(\ell, \tilde{\lambda}_j)}$ are defined by

$$e_{(\ell, \lambda_0)}(\lambda) := (\lambda - \lambda_0)^{\ell}.$$

The factorization space $\mathbb{S}_{\tilde{p}}$ for $\tilde{p}$ is given by

$$\mathbb{S}_{\tilde{p}} := \mathbb{C} \times \mathbb{P}^{n_1 - 1} \times \mathbb{P}^{n_2 - 1} \times \cdots \times \mathbb{P}^{n_m - 1},$$

where the component indexing for elements of $\mathbb{S}_{\tilde{p}}$ starts at zero so that the jth component is an element of $\mathbb{P}^{n_j - 1}$.

Taylor Mappings $\mathcal{T}_{\tilde{p}} : \mathbb{S}_{\tilde{p}} \to \mathbb{C}^{\tilde{n}+1}$

For each $\lambda_0 \in \mathbb{C}$, the scalar *Taylor maps* $\tau_{(k,\lambda_0)} : \mathbb{P}^{\tilde{n}} \to \mathbb{C}$ are

$$\tau_{(k,\lambda_0)}(q) := q^{(k)}(\lambda_0)/k!\ ,\ \text{for } k = 0, 1, 2, \ldots, \tilde{n},$$

where $q^{(\ell)}$ is the ℓth derivative of q.

Define the $\mathbb{C}$-linear isomorphism $\mathcal{T}_{\tilde{p}} : \mathbb{S}_{\tilde{p}} \to \mathbb{C}^{\tilde{n}+1}$ by

$$\mathcal{T}_{\tilde{p}}(u) := [\mu_0, (\tau_{(n_1-1,\tilde{\lambda}_1)}(u_1), \ldots, \tau_{(0,\tilde{\lambda}_1)}(u_1)), \ldots, (\tau_{(n_m-1,\tilde{\lambda}_m)}(u_m), \ldots, \tau_{(0,\tilde{\lambda}_m)}(u_m))]^T.$$

$\mathcal{T}_{\tilde{p}}$ induces and inner product on $\mathbb{S}_{\tilde{p}}$

By the inner product construction Lemma we have

$$\langle u, w \rangle^{\mathfrak{c}}_{\mathbb{S}_{\tilde{p}}} := \langle \mathcal{T}_{\tilde{p}}(u), \mathcal{T}_{\tilde{p}}(w) \rangle^{\mathfrak{c}}_{\mathbb{C}^{\tilde{n}+1}}, \text{ for all } u, w \in \mathbb{S}_{\tilde{p}}$$

is an inner product on $\mathbb{S}_{\tilde{p}}$ with

$$\mathcal{T}^{\star}_{\tilde{p}} = \mathcal{T}^{-1}_{\tilde{p}}$$

with respect to the inner products $\langle \cdot, \cdot \rangle^{\mathfrak{c}}_{\mathbb{S}_{\tilde{p}}}$ and $\langle \cdot, \cdot \rangle^{\mathfrak{c}}_{\mathbb{C}^{\tilde{n}+1}}$.

Here $\langle z, w \rangle_{\mathbb{C}^{\tilde{n}+1}} := \operatorname{Re} \langle z, w \rangle^{\mathfrak{c}}_{\mathbb{C}^{\tilde{n}+1}}$ with

$$\langle z, w \rangle^{\mathfrak{c}}_{\mathbb{C}^{\tilde{n}+1}} = \left\langle (z_0, \ldots, z_{\tilde{n}})^T, (w_0, \ldots, w_{\tilde{n}})^T \right\rangle^{\mathfrak{c}}_{\mathbb{C}^{\tilde{n}+1}} := \sum_{j=0}^{\tilde{n}} \bar{z}_j w_j .$$

The Diffeomorphism $F_{\tilde{p}} : \mathbb{S}_{\tilde{p}} \to \mathbb{P}^{\tilde{n}}$

Recall

$$\tilde{p} := \prod_{j=1}^{m} e_{(n_j,\tilde{\lambda}_j)}$$

and

$$\mathbb{S}_{\tilde{p}} := \mathbb{C} \times \mathbb{P}^{n_1-1} \times \mathbb{P}^{n_2-1} \times \cdots \times \mathbb{P}^{n_m-1}.$$

Define

$$F_{\tilde{p}}(q_0, q_1, q_2, \ldots, q_m) := (1+q_0)\prod_{j=1}^{m}(e_{(n_j,\tilde{\lambda}_j)} + q_j).$$

Then $F_{\tilde{p}}$ is a local diffeomorphism over $\mathbb{C}$ with $F_{\tilde{p}}(0) = \tilde{p}$ and

$$F'_{\tilde{p}}(0)(\omega_0, w_1, w_2, \ldots, w_m) = \omega_0\tilde{p} + \sum_{j=1}^{m} r_j w_j,$$

where $r_j := \tilde{p}/e_{(n_j,\tilde{\lambda}_j)}$.

$F'_{\tilde{p}}(0)^{-1}$ induces and inner product on $\mathbb{P}^{\tilde{n}}$

Since $F'_{\tilde{p}}(0) : \mathbb{S}_{\tilde{p}} \to \mathbb{P}^{\tilde{n}}$ is a $\mathbb{C}$-linear isomorphism, the inner product construction Lemma tells us that $F'_{\tilde{p}}(0)^{-1}$ induces an inner product on $\mathbb{P}^{\tilde{n}}$ through the inner product $\langle \cdot, \cdot \rangle^{\mathfrak{c}}_{\mathbb{S}_{\tilde{p}}}$ by setting

$$\begin{aligned}\langle z, v \rangle^{\mathfrak{c}}_{(\mathbb{P}^{\tilde{n}}, \tilde{p})} &:= \left\langle F'_{\tilde{p}}(0)^{-1} z, F'_{\tilde{p}}(0)^{-1} v \right\rangle^{\mathfrak{c}}_{\mathbb{S}_{\tilde{p}}} \\ &= \left\langle \mathcal{T}_{\tilde{p}}(F'_{\tilde{p}}(0)^{-1} z), \mathcal{T}_{\tilde{p}}(F'_{\tilde{p}}(0)^{-1} v) \right\rangle^{\mathfrak{c}}_{\mathbb{C}^{\tilde{n}+1}} .\end{aligned}$$

With respect to these inner products, we have

$$(F'_{\tilde{p}}(0)^{-1})^{\star} = F'_{\tilde{p}}(0), \quad \text{and} \quad (\mathcal{T}_{\tilde{p}} \circ F'_{\tilde{p}}(0)^{-1})^{\star} = F'_{\tilde{p}}(0) \circ \mathcal{T}_{\tilde{p}}^{-1}.$$

Every $\tilde{p} \in \mathbb{M}^{\tilde{n}}$ induces and inner product on $\mathbb{P}^{\tilde{n}}$ in this way.

The subdifferential of $\mathbf{f}$ (B-Eaton (2012))

Recall that $\mathbf{f} : \mathbb{P}^n \to \overline{\mathbb{R}}$ is given by

$$\mathbf{f}(p) := \max\{f(\lambda) \,|\, \lambda \in \mathbb{C} \text{ and } p(\lambda) = 0\}.$$

Let $\tilde{p} \in \mathbb{P}^{\tilde{n}} \cap \operatorname{dom}(\mathbf{f})$ have degree $\tilde{n}$ with decomposition

$$\tilde{p} := \prod_{j=1}^{m} e_{(n_j, \tilde{\lambda}_j)},$$

and set $\Xi_{\tilde{p}} := \{\lambda_1, \ldots, \lambda_m\}$ and

$$\mathcal{A}_{\mathbf{f}}(\tilde{p}) := \{\lambda_j \in \Xi_p \,|\, f(\lambda_j) = \mathbf{f}(p)\}\,.$$

Assume that every active root $\lambda_j \in \mathcal{A}_{\mathbf{f}}(\tilde{p})$ satisfies the *active root hypotheses* with $\partial f(\lambda) \neq \{0\}$. Then, with respect to $\langle \cdot, \cdot \rangle^{\mathfrak{c}}_{\mathbb{S}_{\tilde{p}}}$,

$$\partial \mathbf{f}(\tilde{p}) = F'_{\tilde{p}}(0) \circ \mathcal{T}_{\tilde{p}}^{-1}(D_{\tilde{p}}) \subset \mathbb{P}^{\tilde{n}},$$

where

$$D_{\tilde{p}} := \operatorname{conv}\left(\{0\} \times \bigtimes_{j=1}^{m} \Gamma(n_j, \tilde{\lambda}_j)\right) \ \subset \ \mathbb{C}^{\tilde{n}+1}.$$

The sets $\Gamma(n_j, \tilde{\lambda}_j)$ (B-Lewis-Overton (2005))

$$\Gamma(n_j, \tilde{\lambda}_j) := \begin{cases} (-\nabla f(\tilde{\lambda}_j)/n_j) \times \mathcal{D}(n_j, \tilde{\lambda}_j) \times \mathbb{C}^{n_j-2} & \text{if } f \ \mathcal{C}^2 \text{ at } \tilde{\lambda}_j, \\ (-\partial f(\tilde{\lambda}_j)/n_j) \times \mathcal{Q}(\tilde{\lambda}_j) \times \mathbb{C}^{n_j-2} & \text{if } f \text{ nonsmooth at } \tilde{\lambda}_j \end{cases}$$

$$\mathcal{D}(n_j, \tilde{\lambda}_j) := \{\theta \mid \langle \theta, (\nabla f(\tilde{\lambda}_j))^2 \rangle_{\mathbb{C}} \le \langle \mathsf{i}\nabla f(\tilde{\lambda}_j), \nabla^2 f(\tilde{\lambda}_j)(\mathsf{i}\nabla f(\tilde{\lambda}_j)) \rangle_{\mathbb{C}} / n_j\}$$

$$\mathcal{Q}(\tilde{\lambda}_j) := -\mathrm{cone}\,(\partial f(\tilde{\lambda}_j)^2) + \mathsf{i}(\mathrm{rspan}\,(\partial f(\tilde{\lambda}_j)^2)),$$

where $\partial f(\tilde{\lambda}_j)^2 := \{g^2 | g \in \partial f(\tilde{\lambda}_j)\}$.

Jordan Decomposition

Let

$$\widetilde{\Xi} := \{\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m\}$$

be a subset of the distinct eigenvalues of $\widetilde{X} \in \mathbb{C}^{n \times n}$. The Jordan structure of $\widetilde{X}$ relative to these eigenvalues is given by

$$J := \widetilde{P}\widetilde{X}\widetilde{P}^{-1} = \mathrm{Diag}\,(\widetilde{B}, J_1, \ldots, J_m),$$

where

$$J_j := \mathrm{Diag}\,(J_j^{(1)}, \ldots, J_j^{(q_j)})$$

and $J_j^{(k)}$ is an $m_{jk} \times m_{jk}$ Jordan block

$$J_j^{(k)} := \tilde{\lambda}_j I_{m_{jk}} + N_{jk}, \quad k = 1, \ldots, q_j, \; j = 1, \ldots, m,$$

where $N_{jk} \in \mathbb{C}^{m_{jk} \times m_{jk}}$ is the nilpotent matrix given by ones on the superdiagonal and zeros elsewhere, and $I_{m_{jk}} \in \mathbb{C}^{m_{jk} \times m_{jk}}$ is the identity matrix. With this notation, q_j is the geometric multiplicity of the eigenvalue $\tilde{\lambda}_j$.

Arnold Form: Nonderogatory Case

There exists a neighborhood Ω of $\widetilde{X} \in \mathbb{C}^{n\times n}$ and smooth maps $P : \Omega \to \mathbb{C}^{n\times n}$, $B : \Omega \to \mathbb{C}^{n_0\times n_0}$ and, for $j \in \{1, \ldots, m\}$ and $s \in \{0, 1, \ldots, n_j - 1\}$, $\lambda_{js} : \Omega \to \mathbb{C}$ such that

$$
\begin{aligned}
P(X)XP(X)^{-1} &= \mathrm{Diag}\,(B(X), 0, \ldots, 0) + \textstyle\sum_{j=1}^{m} \check{J}_j(X) \in \mathbb{C}^{n\times n},\\
\lambda_{js}(\widetilde{X}) &= 0, \quad s = 0, 1, \ldots, n_j - 1,\\
P(\widetilde{X}) &= \widetilde{P}, \quad B(\widetilde{X}) = \widetilde{B}, \quad \text{and}\\
P(\widetilde{X})\widetilde{X}P(\widetilde{X})^{-1} &= \mathrm{Diag}\,(B(\widetilde{X}), J_1, \ldots, J_m),
\end{aligned}
$$

where

$$
\begin{aligned}
\check{J}_j(X) &:= \tilde{\lambda}_j J_{j0} + J_{j1} + \textstyle\sum_{s=0}^{n_j-1} \lambda_{js}(X) J_{js}^*,\\
J_{js} &:= \mathrm{Diag}\,(0, \ldots, 0, N_j^s, 0, \ldots, 0), \quad \text{and}\\
J_{j0} &:= \mathrm{Diag}\,(0, \ldots, 0, I_{n_j}, 0, \ldots, 0),
\end{aligned}
$$

with N_j^s and I_{n_j} in the $\tilde{\lambda}_j$ diagonal block. Finally, the functions λ_{js} are uniquely defined on Ω, though the maps P and B are not unique.

Observation on $\lambda_{js}(\widetilde{X})$

Arnold form illustrates a fundamental difference between the symmetric and nonsymmetric cases.

In the symmetric case, the matrices are unitarily diagonalizable so there are no nilpotent matrices N_j and the mappings λ_{js} reduce to the eigenvalue mapping λ_j.

In this case, a seminal result due to Adrian Lewis shows that the variational properties depend only on the eigenvalues (up to the orbit).

On the other hand, in the nonsymmetric case they depend on the entire family of functions λ_{js}.

$\nabla \lambda_{js}(\widetilde{X})$ (B-Lewis-Overton (2001))

The gradients of the functions $\lambda_{js} : \mathbb{C}^{n \times n} \to \mathbb{C}$ are given by

$$\nabla \lambda_{js}(\widetilde{X}) = (n_j - s)^{-1} \widetilde{P}^* J_{js}^* \widetilde{P}^{-*},$$

with respect to the inner product $\langle \cdot, \cdot \rangle^{\mathfrak{c}}_{\mathbb{C}^{n \times n}}$.

$$J_{js} := \mathrm{Diag}\,(0, \ldots, 0, N_j^s, 0, \ldots, 0)$$

$$N_j := \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & & \\ & & & & & & \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix}_{n_j \times n_j}$$

Derivatives of Characteristic Factors

$$\Phi_n(\widetilde{X}) = \Phi_{\tilde{n}}(\widehat{J}(\widetilde{X}))\Phi_{n_0}(B(\widetilde{X})) = \tilde{p}\,\Phi_{n_0}(B(\widetilde{X}))$$

$$\Phi_{\tilde{n}}(\widehat{J}(\widetilde{X})) = \prod_{j=1}^{m} \Phi_{n_j}(\widehat{J}_j(X))$$

$$\widehat{J}_j(X) = \widehat{J}_j(X) := \tilde{\lambda}_j I_{n_j} + J_j + \textstyle\sum_{s=0}^{n_j-1} \lambda_{js}(X)(J_j^s)^*$$

Derivatives of Characteristic Factors

$$\Phi_n(\widetilde{X}) = \Phi_{\tilde{n}}(\widehat{J}(\widetilde{X}))\Phi_{n_0}(B(\widetilde{X})) = \tilde{p}\,\Phi_{n_0}(B(\widetilde{X}))$$

$$\Phi_{\tilde{n}}(\widehat{J}(\widetilde{X})) = \prod_{j=1}^{m} \Phi_{n_j}(\widehat{J}_j(X))$$

$$\widehat{J}_j(X) = \widehat{J}_j(X) := \tilde{\lambda}_j I_{n_j} + J_j + \textstyle\sum_{s=0}^{n_j-1} \lambda_{js}(X)(J_j^s)^*$$

Theorem (B-Eaton (2016))

$$\Phi_{n_j}(\widehat{J}_j)(X) = e_{(n_j,\tilde{\lambda}_j)} - \sum_{s=0}^{n_j-1} (n_j - s)\lambda_{js}(X) e_{(n_j-s-1,\tilde{\lambda}_j)} + o(\lambda_{j0}(X), \ldots, \lambda_{j(n_j-1)}(X))$$

and so

$$\left((\Phi_{n_j}(\widehat{J}_j))'(X)\right)^{\star} = -\textstyle\sum_{s=0}^{n_j-1} \widetilde{P}^* J_{js}^* \widetilde{P}^{-*}\,\tau_{(n_j-s-1,\tilde{\lambda}_j)}.$$

The mapping $G : \mathbb{C} \times \Omega \to \mathbb{S}_{\tilde{p}}$

$$G(\zeta, X) := (\zeta, g_1(X), \ldots, g_m(X)),$$

where

$$g_j(X) := \Phi_{n_j}(\widehat{J}_j)(X) - e_{(n_j, \tilde{\lambda}_j)}.$$

Then

$$\mathfrak{f}(X) = (\mathbf{f} \circ F_{\tilde{p}} \circ G)(\zeta, X) \qquad (\forall\ \zeta \neq 0).$$

The mapping $G : \mathbb{C} \times \Omega \to \mathbb{S}_{\tilde{p}}$

$$G(\zeta, X) := (\zeta, g_1(X), \ldots, g_m(X)),$$

where

$$g_j(X) := \Phi_{n_j}(\widehat{J}_j)(X) - e_{(n_j, \tilde{\lambda}_j)}.$$

Then

$$\mathfrak{f}(X) = (\mathbf{f} \circ F_{\tilde{p}} \circ G)(\zeta, X) \qquad (\forall\ \zeta \neq 0).$$

We have

$$(G'(\zeta, \widetilde{X}))^\star = R \circ \mathcal{T}_{\tilde{p}},$$

where $R : \mathbb{C}^{\tilde{n}+1} \to \mathbb{C} \times \mathbb{C}^{n \times n}$ be the $\mathbb{C}$-linear transformation

$$R(v) := \left(v_0,\ -\sum_{j=1}^m \sum_{s=0}^{n_j-1} v_{js} \widetilde{P}^* J_{js}^* \widetilde{P}^{-*} \right),$$

for all $v := (v_0, v_{10}, \ldots, v_{1(n_1-1)}, \ldots, v_{m0}, \ldots, v_{m(n_m-1)}) \in \mathbb{C}^{\tilde{n}+1}$.

$\mathfrak{f}(X) = (\mathbf{f} \circ F_{\tilde{p}} \circ G)(\zeta, X)$

$$\begin{array}{ccc}
\mathbb{C}^{n\times n} & \xrightarrow{\ \mathfrak{f}\ } & \overline{\mathbb{R}} \\
{\scriptstyle G(0,\cdot)}\Big\downarrow & & \Big\uparrow{\scriptstyle \mathbf{f}} \\
\mathbb{S}_{\tilde{p}} & \xrightarrow[F_{\tilde{p}}]{} & \mathbb{P}^{\tilde{n}}
\end{array}$$

Define

$$\hat{\mathfrak{f}}(\zeta, X) := (\mathbf{f} \circ F_{\tilde{p}} \circ G)(\zeta, X).$$

Then, with respect to the Frobenius inner product on $\mathbb{C}^{n\times n}$,

$$\begin{aligned}
\partial\hat{\mathfrak{f}}(\zeta, \widetilde{X}) &= \partial\hat{\mathfrak{f}}(0, \widetilde{X}) \\
&= G'(0, \widetilde{X})^{\star} \circ (F'_{\tilde{p}}(0))^{\star} \circ \partial\mathbf{f}(F_{\tilde{p}}(G(0, \widetilde{X}))) \\
&= [R \circ \mathcal{T}_{\tilde{p}} \circ (F'_{\tilde{p}}(0))^{-1}] \circ F'_{\tilde{p}}(0) \circ \mathcal{T}_{\tilde{p}}^{-1}(D_{\tilde{p}}) \\
&= R(D_{\tilde{p}}).
\end{aligned}$$

Ideas behind computing $\partial \mathbf{f}$

The polynomial abscissa: $\mathsf{a}(p) := \max\{\operatorname{Re}(\lambda) \mid p(\lambda) = 0\}$.
We have $(v(\lambda), \eta) \in T_{\operatorname{epi}(\mathsf{a})}(\lambda^n, 0)$, where

$$v(\lambda) = b_0\lambda^n + b_1\lambda^{n-1} + b_2\lambda^{n-2} + \cdots + b_n,$$

if and only if

$$-\frac{\operatorname{Re} b_1}{n} \leq \eta, \tag{1}$$

$$\operatorname{Re} b_2 \geq 0, \tag{2}$$

$$\operatorname{Im} b_2 = 0, \text{ and} \tag{3}$$

$$b_k = 0, \text{ for } k = 3, \ldots, n. \tag{4}$$

The Gauss-Lukas Theorem (1830)

All critical points of a non-constant polynomial lie in the convex hull of the set of roots of the polynomial.

That is,

$$\mathcal{R}(p') \subset \mathrm{co}\,(\mathcal{R}(p))$$

where

$$\mathcal{R}(q) = \{\lambda \mid q(\lambda) = 0\}\,.$$

The Gauss-Lukas Theorem (1830)

All critical points of a non-constant polynomial lie in the convex hull of the set of roots of the polynomial.

That is,

$$\mathcal{R}(p') \subset \mathrm{co}\,(\mathcal{R}(p))$$

where

$$\mathcal{R}(q) = \{\lambda \mid q(\lambda) = 0\}\,.$$

Suppose $\deg p = n$ and $(v(\lambda), \eta) \in T_{\mathrm{epi}(\mathsf{a})}(p, \mu)$.
Then, by Gauss-Lucas,

$$\mathcal{R}(p^{(n-1)}) \subset \mathrm{conv}\,\mathcal{R}(p^{(n-2)}) \subset \ldots \mathrm{conv}\,\mathcal{R}(p) \subset \{\zeta \mid \langle 1, \zeta\rangle \leq \mu\}\,,$$

which implies

$$\mathsf{a}(p^{(n-1)}) \leq \mathsf{a}(p^{(n-2)}) \leq \cdots \leq \mathsf{a}(p) \leq \mu.$$

Consequences for $T_{\mathrm{epi}\,(\mathsf{a})}\,(\lambda^n, 0)$

Suppose $(v, \eta) \in T_{\mathrm{epi}\,(\mathsf{a})}\,(\lambda^n, 0)$, that is there exists

$$t_j \downarrow 0 \quad \text{and} \quad \{(p_j, \mu_j)\} \in \mathrm{epi}(\mathsf{a})$$

such that

$$t_j^{-1}((p_j, \mu_j) - (\lambda^n, 0)) \to (v, \eta).$$

Then there exists

$$\{(a_0^j, a_1^j, \ldots, a_n^j)\} \in \mathbb{C}^{n+1}$$

such that

$$p_j(\lambda) = \sum_{k=0}^{n} a_k^j \lambda^{n-k}$$

with

$$t_j^{-1}\mu_j \to \eta, \qquad t_j^{-1}(a_0^j - 1) \to b_0,$$

$$t_j^{-1} a_k^j \to b_k, \;\; k = 1, \ldots, n,$$

where

$$v(\lambda) = \sum_{k=0}^{n} b_k \lambda^{n-k}.$$

Apply Gauss-Lukas Theorem

$$\mathcal{R}(p_j^{(n-1)}) \subset \operatorname{conv} \mathcal{R}(p_j^{(n-2)}) \subset \ldots \operatorname{conv} \mathcal{R}(p_j) \subset \{\zeta \mid \langle 1, \zeta \rangle \leq \mu\},$$

for each $j = 1, 2, 3, \ldots$.
Thus, for $j = 1, 2, 3, \ldots$ and $\ell = 1, 2, \ldots, n-1$

$$\mu_j \geq \max \left\{ \operatorname{Re} \zeta \;\middle|\; p_j^{(\ell)}(\zeta) = 0 \right\}.$$

where $p_j(\lambda) = a_0^j \lambda^n + a_1^j \lambda^{n-1} + a_2^j \lambda^{n-2} + \ldots + a_n^j$.
For $\ell = n-1$, this yields

$$\mu_j \geq \max \left\{ \operatorname{Re} \zeta \;\middle|\; n! a_0^k \lambda + (n-1)! a_1^j = 0 \right\} = -\frac{1}{n} \operatorname{Re} \frac{a_1^j}{a_0^k}.$$

Hence

$$\frac{\mu_j}{t_j} \geq -\frac{1}{n} \operatorname{Re} \frac{a_1^j}{t_j a_0^k}.$$

Taking the limit in j yields

$$\eta \geq -\frac{\operatorname{Re} b_1}{n}.$$