Scaled Relative Graphs

Ernest K. Ryu

Mathematical and Numerical Optimization
Fall 2021

Scaled relative graphs (SRG)

We present the SRG, which provides a correspondence between algebraic operations on nonlinear operators and geometric operations on subsets of the 2D plane. We can think of the SRG as a signature of an operator analogous to how eigenvalues are a signature of a matrix.

Using the SRG and Euclidean geometry, we establish averagedness and contractiveness of FPIs. (Geometric arguments are rigorous proofs, not mere illustrations.)

Outline

Basic definitions

Scaled relative graphs

Operator and SRG transformations

Averagedness coefficients

Proofs of SRG theorems

Operator classes

$\mathcal{A}$ is a class of operators if $\mathcal{A}$ is a set of operators on $\mathbb{R}^n$ for $n \in \mathbb{N}$. (Technical detail: $\mathbb{A}_1, \mathbb{A}_2 \in \mathcal{A}$, $\mathbb{A}_1 \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ $\mathbb{A}_2 \colon \mathbb{R}^m \rightrightarrows \mathbb{R}^m$, and $n \neq m$ is possible.)

Given classes of operators $\mathcal{A}$ and $\mathcal{B}$ and $\alpha > 0$, write

$$
\begin{aligned}
\mathcal{A} + \mathcal{B} &= \{\mathbb{A} + \mathbb{B} \,|\, \mathbb{A} \in \mathcal{A},\, \mathbb{B} \in \mathcal{B},\, \mathbb{A} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n,\, \mathbb{B} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n\} \\
\mathcal{A}\mathcal{B} &= \{\mathbb{A}\mathbb{B} \,|\, \mathbb{A} \in \mathcal{A},\, \mathbb{B} \in \mathcal{B},\, \mathbb{A} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n,\, \mathbb{B} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n\} \\
\mathbb{J}_{\alpha\mathcal{A}} &= \{\mathbb{J}_{\alpha\mathbb{A}} \,|\, \mathbb{A} \in \mathcal{A},\, \mathbb{A} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n\} \\
\mathbb{R}_{\alpha\mathcal{A}} &= 2\mathbb{J}_{\alpha\mathcal{A}} - \mathbb{I} = \{2\mathbb{J} - \mathbb{I} \,|\, \mathbb{J} \in \mathbb{J}_{\alpha\mathcal{A}},\, \mathbb{J} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n,\, \mathbb{I} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n\} \\
\mathcal{A}^{-1} &= \{\mathbb{A}^{-1} \,|\, \mathbb{A} \in \mathcal{A}\} \\
\alpha\mathcal{A} &= \{\alpha\mathbb{A} \,|\, \mathbb{A} \in \mathcal{A}\}.
\end{aligned}
$$

Operator classes

Class of L-Lipschitz operators:

$$\mathcal{L}_L = \left\{\mathbb{A}\colon \operatorname{dom}\mathbb{A} \to \mathbb{R}^n \mid \|\mathbb{A}x - \mathbb{A}y\|^2 \leq L^2\|x-y\|^2,\ \forall\, x, y \in \operatorname{dom}\mathbb{A} \subseteq \mathbb{R}^n,\ n \in \mathbb{N}\right\}.$$

Class of β-cocoercive operators:

$$\mathcal{C}_\beta = \left\{\mathbb{A}\colon \operatorname{dom}\mathbb{A} \to \mathbb{R}^n \mid \langle \mathbb{A}x - \mathbb{A}y, x-y\rangle \geq \beta\|\mathbb{A}x - \mathbb{A}y\|^2,\ \forall\, x, y \in \operatorname{dom}\mathbb{A} \subseteq \mathbb{R}^n,\ n \in \mathbb{N}\right\}.$$

Class of monotone operators:

$$\mathcal{M} = \left\{\mathbb{A}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n \mid \langle \mathbb{A}x - \mathbb{A}y, x-y\rangle \geq 0,\ \forall\, x, y \in \operatorname{dom}\mathbb{A},\ n \in \mathbb{N}\right\}.$$

Class of μ-strongly monotone operators:

$$\mathcal{M}_\mu = \left\{\mathbb{A}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n \mid \langle \mathbb{A}x - \mathbb{A}y, x-y\rangle \geq \mu\|x-y\|^2,\ \forall\, x, y \in \operatorname{dom}\mathbb{A},\ n \in \mathbb{N}\right\}.$$

Class of θ-averaged operators:

$$\mathcal{N}_\theta = (1-\theta)\mathbb{I} + \theta\mathcal{L}_1.$$

(No requirements on domain or maximality.)

Subdifferential operator classes

Class of μ-strongly convex and L-smooth CCP functions on $\mathbb{R}^n$:

$$\mathcal{F}_{\mu,L}$$

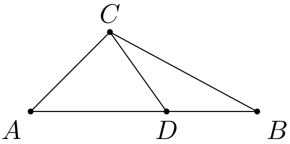
($\mu = 0$ convex but not strongly convex; $L = \infty$ not-smooth.)

Class of subdifferential operators:

$$\partial\mathcal{F}_{\mu,L} = \{\partial f \mid f \in \mathcal{F}_{\mu,L}\}$$

Basic geometry

Stewart's theorem:



lengths satisfy

$$\overline{AD} \cdot \overline{CB}^2 + \overline{DB} \cdot \overline{AC}^2 = \overline{AB} \cdot \overline{CD}^2 + \overline{AD} \cdot \overline{DB}^2 + \overline{AD}^2 \cdot \overline{DB}$$

For any $a, b \in \mathbb{R}^n$, angle function:

$$\angle(a, b) = \begin{cases} \arccos\left(\frac{\langle a,b \rangle}{\|a\|\|b\|}\right) & \text{if } a \neq 0,\ b \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Extended complex plane

Extended complex plane $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ represents the 2D plane and the point at infinity.

We avoid $\infty + \infty$, $0/0$, ∞/∞, and $0 \cdot \infty$.
Define $z + \infty = \infty$, $z/\infty = 0$, and $[z/0 = \infty$ and $z \cdot \infty = \infty$ for $z \neq 0]$.

Inversion map: $z \mapsto \bar{z}^{-1}$, a one-to-one map from $\overline{\mathbb{C}}$ to $\overline{\mathbb{C}}$.
In polar form: $re^{i\varphi} \mapsto (1/r)e^{i\varphi}$.

(Inversion preserves angle and inverts magnitude. $\bar{z}$ is complex conjugate of z.)

Outline

Basic definitions

Scaled relative graphs

Operator and SRG transformations

Averagedness coefficients

Proofs of SRG theorems

SRG of operators

Let $\mathbb{A}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, $u \in \mathbb{A}x$, and $v \in \mathbb{A}y$. Goal: understand change in ouputs $u - v$ relative to change in inputs $x - y$.

For $x \neq y$, consider complex conjugate pair

$$z = \frac{\|u - v\|}{\|x - y\|} \exp\left[\pm i \angle(u - v, x - y)\right].$$

Magnitude $|z| = \frac{\|u-v\|}{\|x-y\|}$ represents size of $u - v$ relative to size of $x - y$. Angle $\angle(u - v, x - y)$ represents how much $u - v$ is aligned with $x - y$.

Equivalently, $\operatorname{Re} z$ and $\operatorname{Im} z$ respectively represent the components of $u - v$ aligned with and perpendicular to $x - y$:

$$\operatorname{Re} z = \operatorname{sgn}(\langle u - v, x - y\rangle) \frac{\|\Pi_{\operatorname{span}\{x-y\}}(u - v)\|}{\|x - y\|} = \frac{\langle u - v, x - y\rangle}{\|x - y\|^2}$$

$$\operatorname{Im} z = \pm \frac{\|\Pi_{\{x-y\}^\perp}(u - v)\|}{\|x - y\|}$$

SRG of operators

SRG of an operator $\mathbb{A}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$:

$$\mathcal{G}(\mathbb{A}) = \left\{ \frac{\|u-v\|}{\|x-y\|} \exp\left[\pm i\angle(u-v, x-y)\right] \,\middle|\, u \in \mathbb{A}x,\ v \in \mathbb{A}y,\ x \neq y \right\}$$

$$\Big(\cup \{\infty\} \text{ if } \mathbb{A} \text{ is multi-valued} \Big)$$

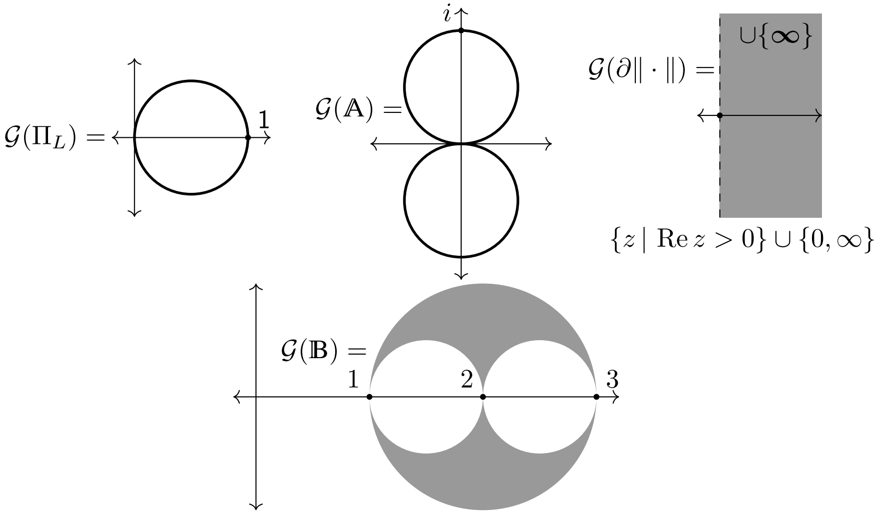
Clarification:

(i) $\mathcal{G}(\mathbb{A}) \subseteq \overline{\mathbb{C}}$

(ii) $[\infty \in \mathcal{G}(\mathbb{A})] \Leftrightarrow [\mathbb{A}x$ is multi-valued for some x.]
(Since $(x, u), (y, v) \in \mathbb{A}$ with $x = y$ and $u \neq v$,
$|z| = \frac{\|u-v\|}{0} = \infty$ and $u - v$ is infinitely larger than $x - y = 0$.)

(iii) $\pm$ makes $\mathcal{G}(\mathbb{A})$ symmetric about the real axis. (Include $\pm$ because $\angle(u - v, x - y)$ always returns a nonnegative angle.)

Examples: SRG of operators

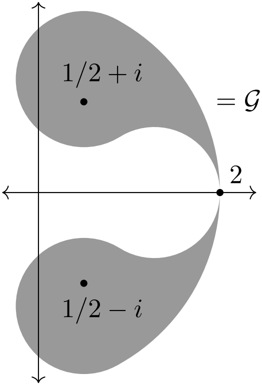


$\Pi_L \colon \mathbb{R}^2 \to \mathbb{R}^2$ the projection onto a line L; $\mathbb{A} \colon \mathbb{R}^2 \to \mathbb{R}^2$ is $\mathbb{A}(u, v) = (0, u)$; $\partial\|\cdot\|$ for $n \geq 2$; and $\mathbb{B} \colon \mathbb{R}^3 \to \mathbb{R}^3$ is $\mathbb{B}(u, v, w) = (u, 2v, 3w)$; Shapes obtained through direct calculations.

SRG and eigenvalues

For linear operators, SRG generalizes eigenvalues:
$\Lambda(\mathbb{A}) \subseteq \mathcal{G}(\mathbb{A})$, if $\mathbb{A} \in \mathbb{R}^{n\times n}$ and $n = 1$ or $n \geq 3$.

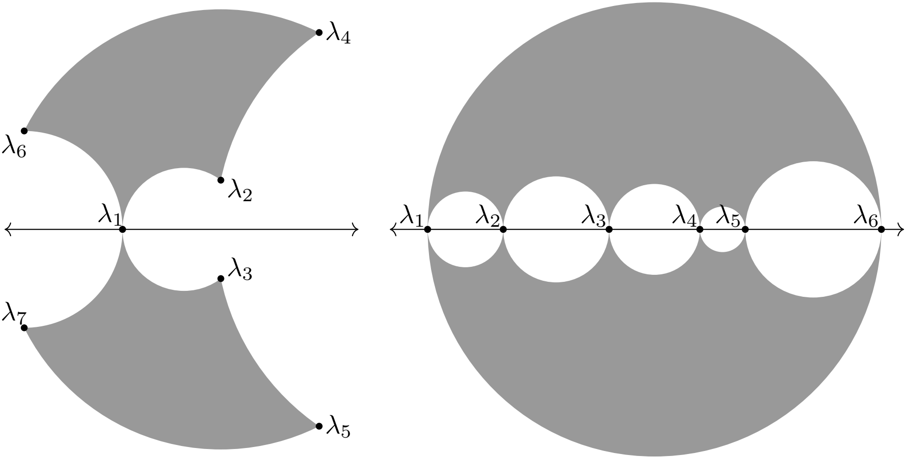
SRG of a 3×3 matrix. The three points denote the eigenvalues.



$$= \mathcal{G}\left(\begin{bmatrix} 1/2 & 2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right)$$

Example: SRG of normal matrices

For normal matrices, multiplicity of eigenvalues do not affect the SRG.



(Left) SRG of an $n \times n$ normal matrix with one distinct real eigenvalue and three distinct complex conjugate eigenvalue pairs. (Right) SRG of an $n \times n$ symmetric matrix with distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_6$.

SRG of operator classes

SRG of a collection of operators:

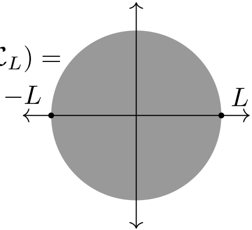
$$\mathcal{G}(\mathcal{A}) = \bigcup_{\mathbb{A} \in \mathcal{A}} \mathcal{G}(\mathbb{A}).$$

(Theorems are usually stated with operator classes. For example, “If $\mathbb{A}$ is $\beta/2$-cocoercive, i.e., if $\mathbb{A} \in \mathcal{C}_{\beta/2}$, then $\mathbb{I} - \mathbb{A}$ is nonexpansive.”)

Theorem 19.

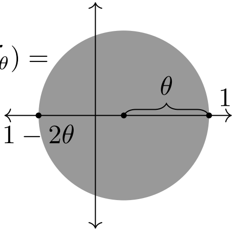
Let $\mu, \beta, L \in (0, \infty)$ and $\theta \in (0, 1)$. Then

$\mathcal{G}(\mathcal{L}_L) =$



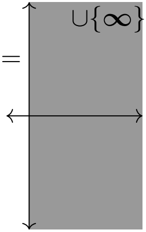
$\{z \in \mathbb{C} \mid |z|^2 \leq L^2\}$

$\mathcal{G}(\mathcal{N}_\theta) =$



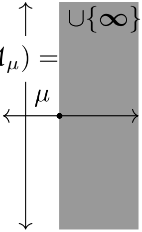
$\{z \in \mathbb{C} \mid |z|^2 + (1 - 2\theta) \leq 2(1 - \theta)\,\mathrm{Re}\, z\}$

$\mathcal{G}(\mathcal{M}) =$



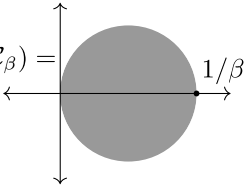
$\{z \in \mathbb{C} \mid \mathrm{Re}\, z \geq 0\} \cup \{\infty\}$

$\mathcal{G}(\mathcal{M}_\mu) =$



$\{z \in \mathbb{C} \mid \mathrm{Re}\, z \geq \mu\} \cup \{\infty\}$

$\mathcal{G}(\mathcal{C}_\beta) =$

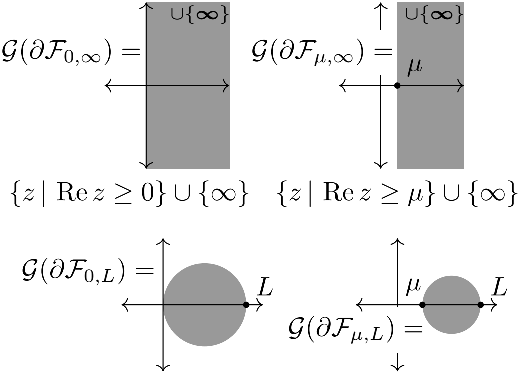


$\{z \in \mathbb{C} \mid \mathrm{Re}\, z \geq \beta |z|^2\}$

SRG of subdifferential operators

Theorem 20.

Let $0 < \mu < L < \infty$. *Then*



SRG-full classes

An operator defines its SRG. Conversely, can we examine the SRG and conclude something about the operator?

A class of operators $\mathcal{A}$ is *SRG-full* if

$$\mathbb{A} \in \mathcal{A} \quad \Leftrightarrow \quad \mathcal{G}(\mathbb{A}) \subseteq \mathcal{G}(\mathcal{A}).$$

Since $\mathbb{A} \in \mathcal{A} \Rightarrow \mathcal{G}(\mathbb{A}) \subseteq \mathcal{G}(\mathcal{A})$ holds by definition,
$\mathcal{G}(\mathbb{A}) \subseteq \mathcal{G}(\mathcal{A}) \Rightarrow \mathbb{A} \in \mathcal{A}$ is the substance of SRG-fullness.

Essentially, a class is SRG-full if it can be fully characterized by its SRG; we can check membership $\mathbb{A} \in \mathcal{A}$ by verifying (through geometric arguments) containment $\mathcal{G}(\mathbb{A}) \subseteq \mathcal{G}(\mathcal{A})$ in the 2D plane.

SRG-full classes

SRG-fullness assumes the desirable property $\mathcal{G}(\mathbb{A}) \subseteq \mathcal{G}(\mathcal{A}) \Rightarrow \mathbb{A} \in \mathcal{A}$.
The following characterizes classes with this property.

Theorem 21.

An operator class $\mathcal{A}$ is SRG-full if it is defined by

$$\mathbb{A} \in \mathcal{A} \quad \Leftrightarrow \quad h\left(\|u-v\|^2, \|x-y\|^2, \langle u-v, x-y\rangle\right) \leq 0, \quad \forall u \in \mathbb{A}x,\, v \in \mathbb{A}y$$

for some nonnegative homogeneous function $h\colon \mathbb{R}^3 \to \mathbb{R}$.

h is nonnegative homogeneous if $\theta h(a,b,c) = h(\theta a, \theta b, \theta c)$ for all $\theta \geq 0$.
(We do not assume h is smooth.)

Example: SRG-full classes

When a class $\mathcal{A}$ is defined by h as in Theorem 21, we say h *represents* $\mathcal{A}$.

μ-strongly monotone class $\mathcal{M}_\mu$ represented by $h(a,b,c) = \mu b - c$:

$$\mathbb{A} \in \mathcal{M}_\mu \quad \Leftrightarrow \quad \mu\|x-y\|^2 \leq \langle u-v, x-y\rangle, \quad \forall u \in \mathbb{A}x,\, v \in \mathbb{A}y.$$

Firmly nonexpansive class $\mathcal{N}_{1/2}$ represented by $h(a,b,c) = a - c$:

$$\mathbb{A} \in \mathcal{N}_{1/2} \quad \Leftrightarrow \quad \|u-v\|^2 \leq \langle u-v, x-y\rangle, \quad \forall u \in \mathbb{A}x,\, v \in \mathbb{A}y.$$

By Theorem 21, $\mathcal{M}$, $\mathcal{M}_\mu$, $\mathcal{C}_\beta$, $\mathcal{L}_L$, and $\mathcal{N}_\theta$ are all SRG-full.

Example: Class of subdifferentials are not SRG-full

Classes $\partial\mathcal{F}_{0,\infty}$, $\partial\mathcal{F}_{\mu,\infty}$, $\partial\mathcal{F}_{0,L}$, and $\partial\mathcal{F}_{\mu,L}$ are *not* SRG-full.

For example,

$$\mathbb{A}(z_1, z_2) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_2 \\ z_2 \end{bmatrix}$$

satisfies $\mathcal{G}(\mathbb{A}) = \{-i, i\} \subseteq \mathcal{G}(\partial\mathcal{F}_{0,\infty})$, but $\mathbb{A} \notin \partial\mathcal{F}_{0,\infty}$.

(If $\nabla f = \mathbb{A}$ for a function f, then $D\mathbb{A} = \nabla^2 f$ must be symmetric.)

Role of maximality

Maximality is mostly orthogonal to the notion of the SRG: non-maximal operators have a well-defined SRGs, and SRG-full classes contain non-maximal operators.

This separation allows the geometric analyses via SRGs being entangled with the subtleties of maximality.

Outline

Basic definitions

Scaled relative graphs

Operator and SRG transformations

Averagedness coefficients

Proofs of SRG theorems

Intersection

Theorem 22.
If $\mathcal{A}$ and $\mathcal{B}$ are SRG-full classes, then $\mathcal{A} \cap \mathcal{B}$ is SRG-full, and

$$\mathcal{G}(\mathcal{A} \cap \mathcal{B}) = \mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B}).$$

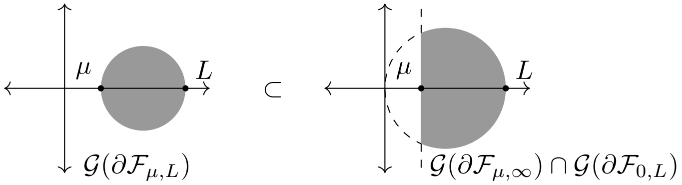
The substance of Theorem 22 is $\mathcal{G}(\mathcal{A} \cap \mathcal{B}) \supseteq \mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B})$ since $\mathcal{G}(\mathcal{A} \cap \mathcal{B}) \subseteq \mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B})$ holds by definition, regardless of SRG-fullness.

Example: Non-SRG-full example

Theorem 22 does not apply when the operator classes are not SRG-full. For example, although

$$\partial\mathcal{F}_{\mu,L} = \partial\mathcal{F}_{\mu,\infty} \cap \partial\mathcal{F}_{0,L}$$

we have the strict containment



Scaling and translation

Theorem 23.

Let $\alpha \in \mathbb{R}$ and $\alpha \neq 0$. If $\mathcal{A}$ is a class of operators, then

$$\mathcal{G}(\alpha\mathcal{A}) = \alpha\mathcal{G}(\mathcal{A}), \qquad \mathcal{G}(\mathbb{I} + \mathcal{A}) = 1 + \mathcal{G}(\mathcal{A}).$$

If $\mathcal{A}$ is furthermore SRG-full, then $\alpha\mathcal{A}$ and $\mathbb{I} + \mathcal{A}$ are SRG-full.

Note a class of operators can consist of a single operator. So

$$\mathcal{G}(\alpha\mathbb{A}) = \alpha\mathcal{G}(\mathbb{A}), \qquad \mathcal{G}(\mathbb{I} + \mathbb{A}) = 1 + \mathcal{G}(\mathbb{A})$$

Convergence analysis: gradient descent

Consider

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x),$$

where f is μ-strongly convex and L-smooth with $0 < \mu < L < \infty$.

Gradient descent

$$x^{k+1} = x^k - \alpha \nabla f(x^k)$$

converges with rate

$$\|x^k - x^\star\| \leq \left(\max\{|1 - \alpha\mu|, |1 - \alpha L|\}\right)^k \|x^0 - x^\star\|$$

for $\alpha \in (0, 2/L)$ by the following Proposition 2.

Convergence analysis: gradient descent

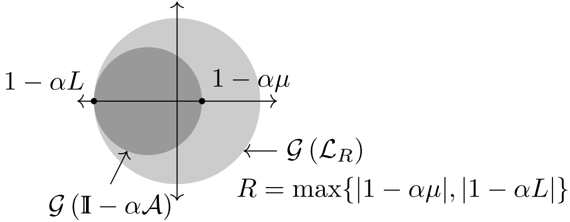
Proposition 2.

Let $0 < \mu < L < \infty$ *and* $\alpha \in (0, \infty)$. *If* $\mathcal{A} = \partial \mathcal{F}_{\mu,L}$, *then* $\mathbb{I} - \alpha\mathcal{A} \subseteq \mathcal{L}_R$ *for*

$$R = \max\{|1-\alpha\mu|, |1-\alpha L|\}.$$

Result is tight in the sense that $\mathbb{I} - \alpha\mathcal{A} \nsubseteq \mathcal{L}_R$ *for any smaller value of* R.

Proof. By Theorems 20 and 23, we have the geometry



The containment of $\mathcal{G}(\mathbb{I} - \alpha\mathcal{A})$ holds for R and fails for smaller R. Since $\mathcal{L}_R$ is SRG-full by Theorem 21, the containment of the SRG in $\overline{\mathbb{C}}$ equivalent to the containment of the class. $\square$

Convergence analysis: Forward step method

Consider

$$\underset{x \in \mathbb{R}^n}{\text{find}} \quad 0 \in \mathbb{A}x,$$

where $\mathbb{A}\colon \mathbb{R}^n \to \mathbb{R}^n$. Consider the forward step method

$$x^{k+1} = x^k - \alpha \mathbb{A} x^k$$

under the following two setups.

If $\mathbb{A}$ is μ-strongly monotone and L-Lipschitz with $0 < \mu < L < \infty$,

$$\|x^k - x^\star\| \le \left(1 - 2\alpha\mu + \alpha^2 L^2\right)^{k/2} \|x^0 - x^\star\|$$

for $\alpha \in (0, 2\mu/L^2)$ by Proposition 3.

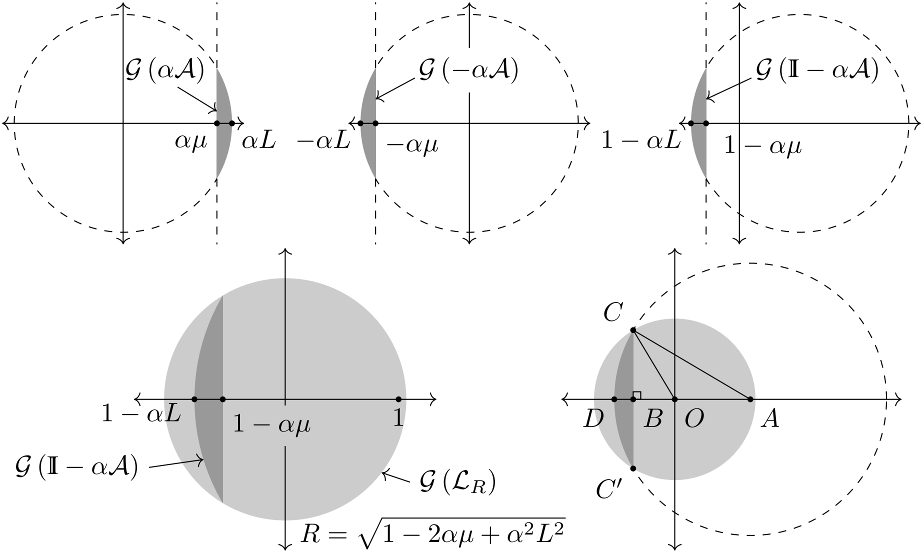
Proposition 3.

Let $0 < \mu < L < \infty$ *and* $\alpha \in (0, \infty)$. *If* $\mathcal{A} = \mathcal{M}_\mu \cap \mathcal{L}_L$, *then* $\mathbb{I} - \alpha\mathcal{A} \subseteq \mathcal{L}_R$ *for*

$$R = \sqrt{1 - 2\alpha\mu + \alpha^2 L^2}.$$

Result is tight in the sense that $\mathbb{I} - \alpha\mathcal{A} \nsubseteq \mathcal{L}_R$ *for any smaller value of* R.

Proof. First consider the case $\alpha\mu > 1$. By Theorems 19 and 23, we have



To clarify, O is the center of the circle with radius $\overline{OC}$ (lighter shade) and A is the center of the circle with radius $\overline{AC} = \overline{AD}$ defining the inner region (darker shade).

Proof of Proposition 3

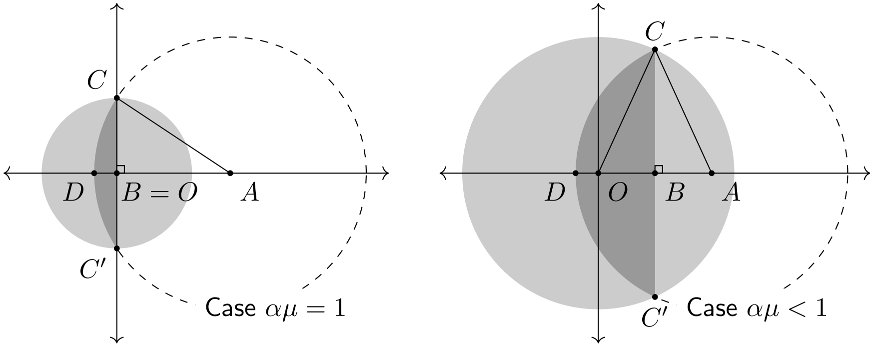
With 2 applications of the Pythagorean theorem, we get

$$\begin{aligned}\overline{OC}^2 &= \overline{CB}^2 + \overline{BO}^2 = \overline{AC}^2 - \overline{BA}^2 + \overline{BO}^2 \\ &= (\alpha L)^2 - (\alpha\mu)^2 + (1-\alpha\mu)^2 = 1 - 2\alpha\mu + \alpha^2 L^2.\end{aligned}$$

Since $\overline{C'C}$ is a chord of circle O, it is within the circle. Since 2 non-identical circles intersect at no more than 2 points, and since D is within circle O, arc $\widehat{CDC'}$ is within circle O. Finally, the region bounded by $\overline{C'C} \cup \widehat{CDC'}$ (darker shade) is within circle O (lighter shade).

Proof of Proposition 3

The previous diagram illustrates the case $\alpha\mu > 1$. When $\alpha\mu = 1$ and $\alpha\mu < 1$, the geometries are slightly different but same arguments hold:



The containment holds for R and fails for smaller R. Since $\mathcal{L}_R$ is SRG-full by Theorem 21, the containment of the SRG in $\overline{\mathbb{C}}$ equivalent to the containment of the class. □

Convergence analysis: Forward step method

If $\mathbb{A}$ is μ-strongly monotone and β-cocoercive with $0 < \mu < 1/\beta < \infty$,

$$\|x^k - x^\star\| \leq \left(1 - 2\alpha\mu + \alpha^2\mu/\beta\right)^{k/2} \|x^0 - x^\star\|$$

for $\alpha \in (0, 2\beta)$ by Proposition 4.

Proposition 4.

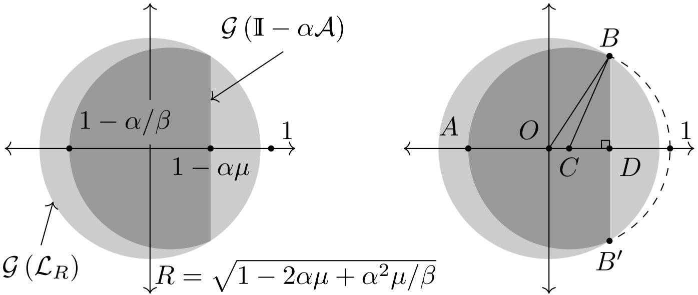
Let $0 < \mu < 1/\beta < \infty$ *and* $\alpha \in (0, 2\beta)$. *If* $\mathcal{A} = \mathcal{M}_\mu \cap \mathcal{C}_\beta$, *then*
$\mathbb{I} - \alpha\mathcal{A} \subseteq \mathcal{L}_R$ *for*

$$R = \sqrt{1 - 2\alpha\mu + \alpha^2\mu/\beta}.$$

Result is tight in the sense that $\mathbb{I} - \alpha\mathcal{A} \nsubseteq \mathcal{L}_R$ *for any smaller value of* R.

Proof of Proposition 4

Proof. First consider the case $\mu < 1/(2\beta)$. By Theorems 19 and 23,



To clarify, O is the center of the circle with radius $\overline{OB}$ (lighter shade) and C is the center of the circle with radius $\overline{AC} = \overline{CB}$ defining the inner region (darker shade).

Proof of Proposition 4

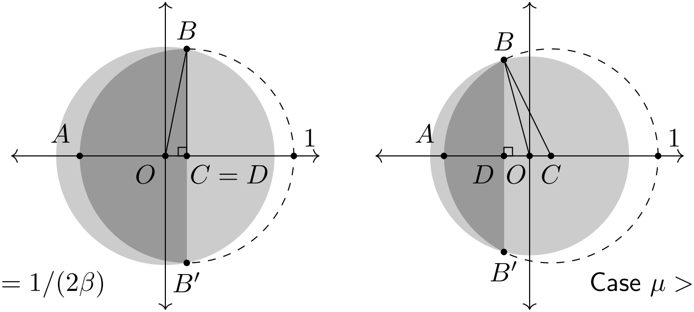
With two applications of the Pythagorean theorem, we get

$$\begin{aligned}\overline{OB}^2 &= \overline{OD}^2 + \overline{DB}^2 = \overline{OD}^2 + \overline{BC}^2 - \overline{CD}^2 \\ &= (1-\alpha\mu)^2 + (\alpha/(2\beta))^2 - (\alpha/(2\beta) - \alpha\mu)^2 = 1 - 2\alpha\mu + \alpha^2\mu/\beta.\end{aligned}$$

Since $\overline{B'B}$ is a chord of circle O, it is within the circle. Since 2 non-identical circles intersect at at most 2 points, and since A is within circle O, arc $\widehat{BAB'}$ is within circle O. Finally, the region bounded by $\overline{B'B} \cup \widehat{BAB'}$ (darker shade) is within circle O (lighter shade).

Proof of Proposition 4

When $\mu = 1/(2\beta)$ and $\mu > 1/(2\beta)$, geometries are slightly different but same arguments hold:



Case $\mu = 1/(2\beta)$

Case $\mu > 1/(2\beta)$

The containment holds for R and fails for smaller R. Since $\mathcal{L}_R$ is SRG-full by Theorem 21, the containment of the SRG in $\overline{\mathbb{C}}$ equivalent to the containment of the class. □

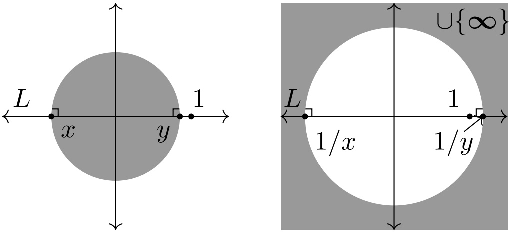
Inversive geometry

Generalized circles consist of (finite) circles and lines with $\{\infty\}$.
(A line is like a circle with infinite radius.)

Inversion maps generalized circles to generalized circles.

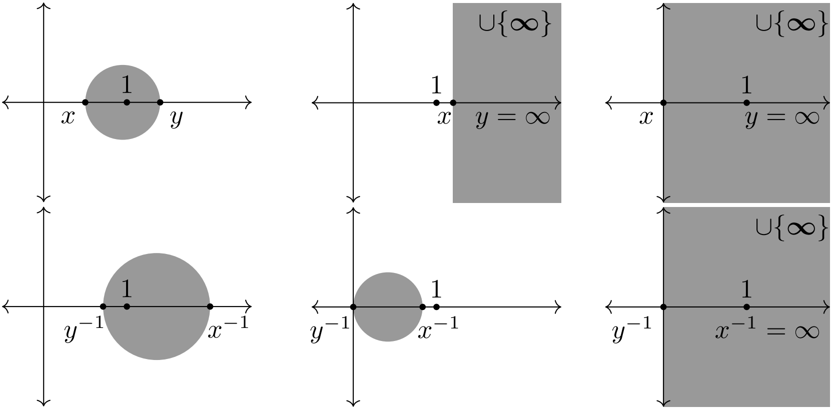
In complex analysis, the inversion map is known as the Möbius transformation. In classical Euclidean geometry, inversive geometry considers generally the inversion of the 2D plane about any circle but our inversion map $z \mapsto \bar{z}^{-1}$ is the inversion about the unit circle.

Inverting generalized circles



1. Draw a line L through the origin orthogonally intersecting the generalized circle. This means L intersects the boundary perpendicularly, which implies L goes through the circle's center when the generalized circle is finite.
2. Let $-\infty < x < y \leq \infty$ represent the signed distance of the intersecting points from the origin along this line. If the generalized circle is a line, then $y = \infty$.
3. Draw a generalized circle orthogonally intersecting L at $(1/x)$ and $(1/y)$.
4. When inverting a region with a generalized circle as the boundary, pick a point on L within the interior of the region to determine on which side of the boundary the inverted interior lies.

Example: Inverting generalized circles



Inversion

We relate inversion of operators with inversion (reciprocal) of complex numbers and utilize inversive geometry.

Theorem 24.

If $\mathcal{A}$ is a class of operators, then

$$\mathcal{G}(\mathcal{A}^{-1}) = (\mathcal{G}(\mathcal{A}))^{-1}.$$

If $\mathcal{A}$ is furthermore SRG-full, then $\mathcal{A}^{-1}$ is SRG-full.

To clarify, $(\mathcal{G}(\mathcal{A}))^{-1} = \{z^{-1} \mid z \in \mathcal{G}(\mathcal{A})\} \subseteq \overline{\mathbb{C}}$.

Note $(\mathcal{G}(\mathcal{A}))^{-1} = (\overline{\mathcal{G}(\mathcal{A})})^{-1}$, since $\mathcal{G}(\mathcal{A})$ is symmetric about real axis, so we write the simpler $(\mathcal{G}(\mathcal{A}))^{-1}$ even though inversion map is $z \mapsto \bar{z}^{-1}$.

Convergence analysis: proximal point

Consider

$$\underset{x\in\mathbb{R}^n}{\text{find}} \quad 0 \in \mathbb{A}x,$$

where $\mathbb{A}$ is maximal μ-strongly monotone. PPM

$$x^{k+1} = \mathbb{J}_{\alpha\mathbb{A}} x^k$$

converges exponentially with rate

$$\|x^k - x^\star\| \leq \left(\frac{1}{1+\alpha\mu}\right)^k \|x^0 - x^\star\|$$

for $\alpha > 0$ by the following Proposition 5.

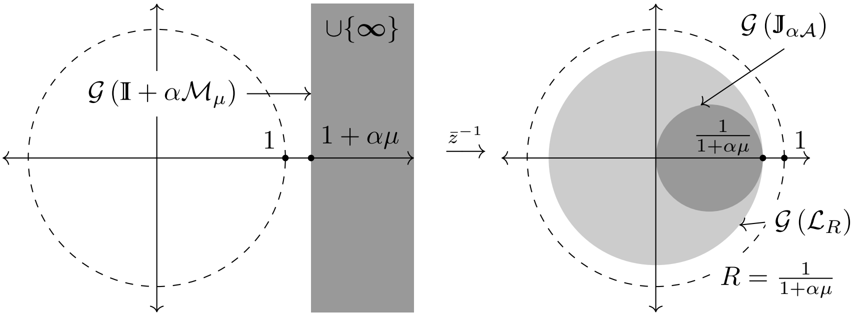
Proposition 5.

Let $\mu \in (0,\infty)$ and $\alpha \in (0,\infty)$. If $\mathcal{A} = \mathcal{M}_\mu$, then $\mathbb{J}_{\alpha\mathcal{A}} \subseteq \mathcal{L}_R$ for

$$R = \frac{1}{1+\alpha\mu}.$$

Result is tight in the sense that $\mathbb{J}_{\alpha\mathcal{A}} \nsubseteq \mathcal{L}_R$ for any smaller value of R.

Proof. By Theorems 19, 23, and 24, we have



The containment holds for R and fails for smaller R. Since $\mathcal{L}_R$ is SRG-full by Theorem 21, the containment of the SRG in $\overline{\mathbb{C}}$ equivalent to the containment of the class. □

Convergence analysis: DRS

Consider

$$\underset{x \in \mathbb{R}^n}{\text{find}} \quad 0 \in (\mathbb{A} + \mathbb{B})x,$$

where $\mathbb{A} \in \mathcal{M}_\mu \cap \mathcal{C}_\beta$ and $\mathbb{B} \in \mathcal{M}$ are maximal monotone. DRS

$$z^{k+1} = \left(\tfrac{1}{2}\mathbb{I} + \tfrac{1}{2}\mathbb{R}_{\alpha\mathbb{A}}\mathbb{R}_{\alpha\mathbb{B}}\right) z^k$$

converges exponentially with rate

$$\|z^k - z^\star\| \leq \left(\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{4\alpha\mu}{1 + 2\alpha\mu + \alpha^2\mu/\beta}}\right)^k \|z^0 - z^\star\|$$

for $\alpha > 0$ by the following Proposition 6 and Exercise 13.9.

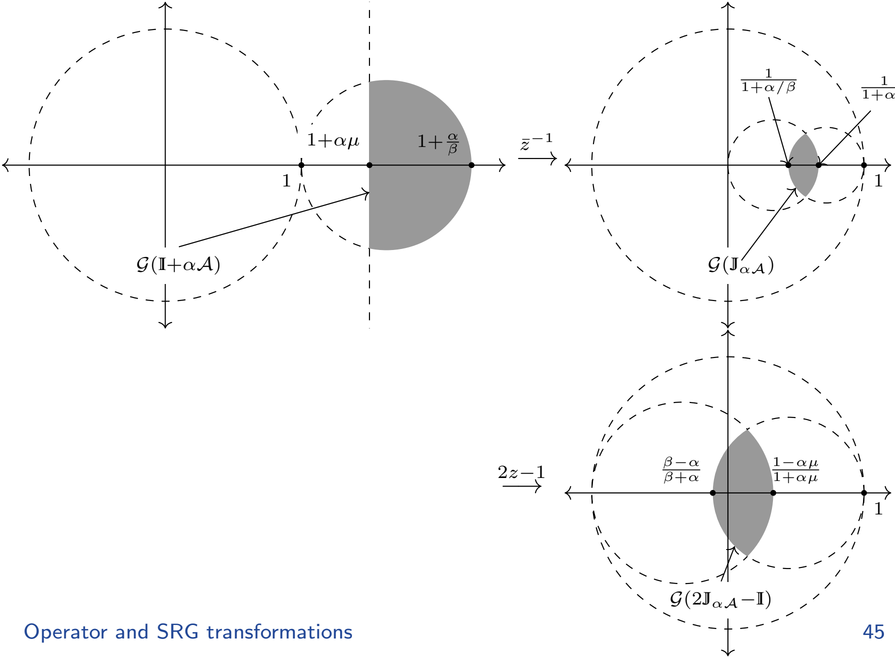
Proposition 6.

Let $0 < \mu < 1/\beta < \infty$ *and* $\alpha \in (0, \infty)$*. If* $\mathcal{A} = \mathcal{M}_\mu \cap \mathcal{C}_\beta$*, then* $\mathbb{R}_{\alpha\mathcal{A}} \subseteq \mathcal{L}_R$ *for*

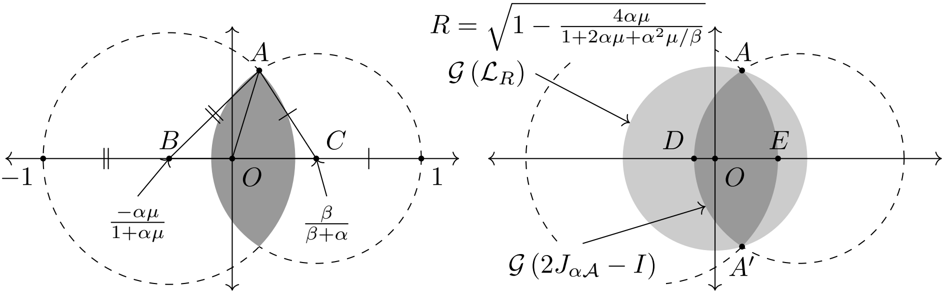
$$R = \sqrt{1 - \frac{4\alpha\mu}{1 + 2\alpha\mu + \alpha^2\mu/\beta}}.$$

Result is tight in the sense that $\mathbb{R}_{\alpha\mathcal{A}} \not\subseteq \mathcal{L}_R$ *for any smaller value of* R*.*

Proof. By Theorems 19, 23, and 24,



A closer look gives us



To clarify, B is the center of the circle with radius $\overline{BA}$ and C is the center of the circle with radius $\overline{CA}$.

By Stewart's theorem, we have

$$\overline{OA}^2 = \frac{\overline{OC}\cdot\overline{AB}^2 + \overline{BO}\cdot\overline{CA}^2 - \overline{BO}\cdot\overline{OC}\cdot\overline{BC}}{\overline{BC}}$$
$$= \frac{\frac{\beta}{\alpha+\beta}\left(1-\frac{\alpha\mu}{1+\alpha\mu}\right)^2 + \frac{\alpha\mu}{1+\alpha\mu}\left(1-\frac{\beta}{\alpha+\beta}\right)^2 - \frac{\beta}{\alpha+\beta}\frac{\alpha\mu}{1+\alpha\mu}\left(\frac{\beta}{\alpha+\beta}+\frac{\alpha\mu}{1+\alpha\mu}\right)}{\frac{\beta}{\alpha+\beta}+\frac{\alpha\mu}{1+\alpha\mu}}$$
$$= 1 - \frac{4\alpha\mu}{1+2\alpha\mu+\alpha^2\mu/\beta}.$$

Since 2 non-identitcal circles intersect at at most 2 points, and since D is within circle B, arc $\widehat{ADA'}$ is within circle O. By the same reasoning, arc $\widehat{A'EA}$ is within circle O. Finally, the region bounded by $\widehat{ADA'} \cup \widehat{A'EA}$ (darker shade) is within circle O (lighter shade).

The containment holds for R and fails for smaller R. Since $\mathcal{L}_R$ is SRG-full by Theorem 21, the containment of the SRG in $\overline{\mathbb{C}}$ equivalent to the containment of the class. □

Convergence analysis: DRS on optimization

Consider

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) + g(x),$$

where f and g are CCP. Assume $f \in \mathcal{F}_{\mu,L}$ or $g \in \mathcal{F}_{0,\infty}$. DRS

$$\begin{aligned} x^{k+1/2} &= \mathrm{Prox}_{\alpha g}(z^k) \\ x^{k+1} &= \mathrm{Prox}_{\alpha f}(2x^{k+1/2} - z^k) \\ z^{k+1} &= z^k + x^{k+1} - x^{k+1/2} \end{aligned}$$

converges exponentially with rate

$$\|z^k - z^\star\| \le \left(\frac{1}{2} + \frac{1}{2} \max \left\{ \left| \frac{1-\alpha\mu}{1+\alpha\mu} \right|, \left| \frac{1-\alpha L}{1+\alpha L} \right| \right\} \right)^k \|z^0 - z^\star\|$$

by the following Proposition 6 and Exercise 13.9.

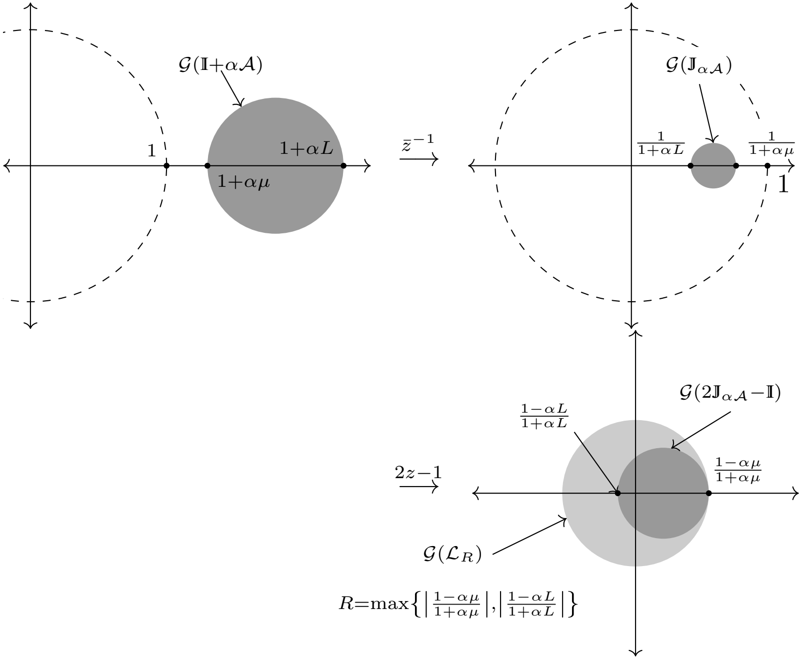
Proposition 7.

Let $0 < \mu < L < \infty$ and $\alpha \in (0, \infty)$. If $\mathcal{A} = \partial \mathcal{F}_{\mu,L}$, then $\mathbb{R}_{\alpha\mathcal{A}} \subseteq \mathcal{L}_R$ for

$$R = \max \left\{ \left| \frac{1-\alpha\mu}{1+\alpha\mu} \right|, \left| \frac{1-\alpha L}{1+\alpha L} \right| \right\}.$$

Result is tight in the sense that $\mathbb{R}_{\alpha\mathcal{A}} \not\subseteq \mathcal{L}_R$ for any smaller value of R.

Proof. By Theorems 20, 23, and 24, we have



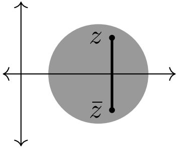
The containment holds for R and fails for smaller R. Since $\mathcal{L}_R$ is SRG-full by Theorem 21, the containment of the SRG in $\overline{\mathbb{C}}$ equivalent to the Since $\mathcal{L}_R$ is SRG-full by Theorem 21, the containment of the SRG in $\overline{\mathbb{C}}$ equivalent to the containment of the class. $\square$

Sum of operators

Lline segment between $z, w \in \mathbb{C}$:

$$[z, w] = \{\theta z + (1-\theta)w \,|\, \theta \in [0,1]\}.$$

SRG-full class $\mathcal{A}$ satisfies the chord property if $z \in \mathcal{G}(\mathcal{A}) \setminus \{\infty\}$ implies $[z, \bar{z}] \subseteq \mathcal{G}(\mathcal{A})$.



Theorem 25.

Let $\mathcal{A}, \mathcal{B}$ be SRG-full classes such that $\infty \notin \mathcal{G}(A), \infty \notin \mathcal{G}(B)$. Then

$$\mathcal{G}(\mathcal{A} + \mathcal{B}) \supseteq \mathcal{G}(\mathcal{A}) + \mathcal{G}(\mathcal{B}).$$

If $\mathcal{A}$ or $\mathcal{B}$ furthermore satisfies the chord property, then

$$\mathcal{G}(\mathcal{A} + \mathcal{B}) = \mathcal{G}(\mathcal{A}) + \mathcal{G}(\mathcal{B}).$$

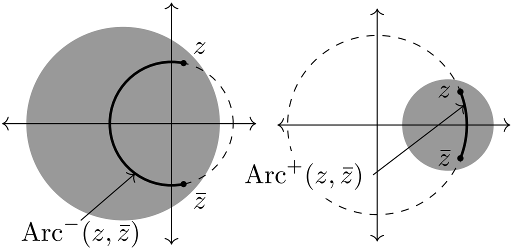
Composition of operators

Right-hand arc between $z \in \mathbb{C}$ and $\bar{z}$:

$$\mathrm{Arc}^+(z,\bar{z}) = \left\{ re^{i(1-2\theta)\varphi} \,\middle|\, z = re^{i\varphi},\ \varphi \in (-\pi,\pi],\ \theta \in [0,1],\ r \geq 0 \right\}$$

Left-hand arc between $z \in \mathbb{C}$ and $\bar{z}$:

$$\mathrm{Arc}^-(z,\bar{z}) = -\mathrm{Arc}^+(-z,-\bar{z}).$$



An SRG-full class $\mathcal{A}$ respectively satisfies the left-arc property and right-arc property if $z \in \mathcal{G}(\mathcal{A}) \setminus \{\infty\}$ implies $\mathrm{Arc}^-(z,\bar{z}) \subseteq \mathcal{G}(\mathcal{A})$ and $\mathrm{Arc}^+(z,\bar{z}) \subseteq \mathcal{G}(\mathcal{A})$, respectively. $\mathcal{A}$ satisfies *an* arc property if the left or right-arc property is satisfied.

Composition of operators

Theorem 26.
Let $\mathcal{A}$ and $\mathcal{B}$ be SRG-full classes such that $\infty \notin \mathcal{G}(\mathcal{A})$, $\emptyset \neq \mathcal{G}(\mathcal{A})$, $\infty \notin \mathcal{G}(\mathcal{B})$, and $\emptyset \neq \mathcal{G}(\mathcal{B})$. Then

$$\mathcal{G}(\mathcal{A}\mathcal{B}) \supseteq \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B}).$$

If $\mathcal{A}$ or $\mathcal{B}$ furthermore satisfies an arc property, then

$$\mathcal{G}(\mathcal{A}\mathcal{B}) = \mathcal{G}(\mathcal{B}\mathcal{A}) = \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B}).$$

Outline

Basic definitions

Scaled relative graphs

Operator and SRG transformations

Averagedness coefficients

Proofs of SRG theorems

Composition of averaged operators

Theorem 27.
Let $\mathbb{T}_1$ *and* $\mathbb{T}_2$ *be* θ_1*- and* θ_2*-averaged operators on* $\mathbb{R}^n$ *with* $\theta_1, \theta_2 \in (0,1)$*. Then* $\mathbb{T}_1\mathbb{T}_2$ *is* θ*-averaged with*

$$\theta = \frac{\theta_1 + \theta_2 - 2\theta_1\theta_2}{1 - \theta_1\theta_2}.$$

Proof of Theorem 27

Note

$$z \in \mathcal{G}(\mathcal{N}_\theta) \quad \Leftrightarrow \quad |z-(1-\theta)|^2 \le \theta^2 \quad \Leftrightarrow \quad |z|^2 \le 1 - \frac{1-\theta}{\theta}|1-z|^2$$

by Theorem 19 and

$$\theta^2 - |z-(1-\theta)|^2 = \theta\left(1 - \frac{1-\theta}{\theta}|1-z|^2 - |z|^2\right).$$

Let $z_1 \in \mathcal{G}(\mathcal{N}_{\theta_1})$ and $z_2 \in \mathcal{G}(\mathcal{N}_{\theta_2})$. Then

$$\begin{aligned}
|z_1 z_2|^2 &\le |z_2|^2\left(1 - \frac{1-\theta_1}{\theta_1}|1-z_1|^2\right) \\
&\le 1 - \frac{1-\theta_2}{\theta_1}|1-z_2|^2 - \frac{1-\theta_1}{\theta_1}|1-z_1|^2|z_2|^2 \\
&= 1 - \frac{1-\theta}{\theta}|1-z_1z_2|^2 - \frac{\theta_1\theta_2}{\theta_1+\theta_2-2\theta_1\theta_2}\left|\frac{1-\theta_1}{\theta_1}(1-z_1)z_2 - \frac{1-\theta_2}{\theta_2}(1-z_2)\right|^2 \\
&\le 1 - \frac{1-\theta}{\theta}|1-z_1z_2|^2
\end{aligned}$$

and $z_1 z_2 \in \mathcal{G}(\mathcal{N}_\theta)$, i.e., $\mathcal{G}(\mathcal{N}_{\theta_1})\mathcal{G}(\mathcal{N}_{\theta_2}) \subseteq \mathcal{G}(\mathcal{N}(\mathcal{G}_\theta))$.

Proof of Theorem 27

Since $\mathcal{N}_{\theta_1}$ satisfies an arc property, $\mathcal{G}(\mathcal{N}_{\theta_1})\mathcal{G}(\mathcal{N}_{\theta_2}) = \mathcal{G}(\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2})$ by Theorem 26. So

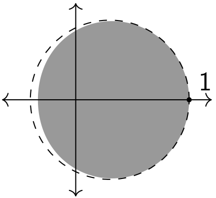
$$\mathcal{G}(\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2}) = \mathcal{G}(\mathcal{N}_{\theta_1})\mathcal{G}(\mathcal{N}_{\theta_2}) \subseteq \mathcal{G}(\mathcal{N}_{\theta}),$$

implies $\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2} \subseteq \mathcal{N}_{\theta}$ by SRG-fullness of $\mathcal{N}_{\theta}$. $\square$

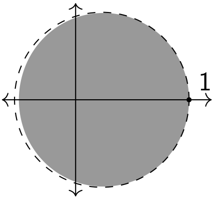
Alternate proof outline of Theorem 27

$\mathcal{G}(\mathcal{N}_{\theta_1})\mathcal{G}(\mathcal{N}_{\theta_2})$ is enclosed by the outer curve defined by

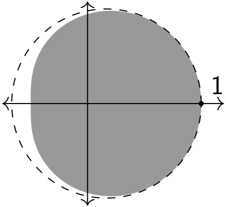
$$r(\varphi)^2 - 2r(\varphi)(\cos(\varphi)(1-\theta_1)(1-\theta_2) + \theta_1\theta_2) + (1-2\theta_1)(1-2\theta_2) = 0.$$



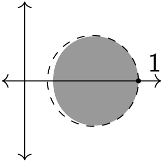
$\theta_1 = \frac{2}{3}$, $\theta_2 = \frac{1}{4}$



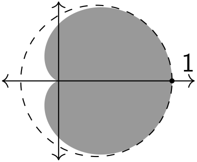
$\theta_1 = \frac{1}{4}$, $\theta_2 = \frac{3}{4}$



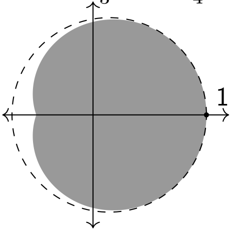
$\theta_1 = \frac{2}{3}$, $\theta_2 = \frac{3}{4}$



$\theta_1 = \theta_2 = \frac{1}{4}$



$\theta_1 = \theta_2 = \frac{1}{2}$



$\theta_1 = \theta_2 = \frac{3}{4}$

Alternate proof outline of Theorem 27

The disk $\mathcal{G}(\mathcal{N}_\theta)$ with $\theta = \frac{\theta_1+\theta_2-2\theta_1\theta_2}{1-\theta_1\theta_2}$ contains this region so $\mathcal{G}(\mathcal{N}_{\theta_1})\mathcal{G}(\mathcal{N}_{\theta_2}) \subseteq \mathcal{G}(\mathcal{N}_\theta)$. On the other hand, the two shapes have matching curvature at 1, the inclusion does not hold with smaller θ. $\square$

Davis–Yin splitting

Theorem 28.
Assume $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{C}$ are maximal monotone. Assume $\mathbb{C}$ is β-cocoercive and $\alpha \in (0, 2\beta)$. The DYS operator $\mathbb{I} - \mathbb{J}_{\alpha\mathbb{B}} + \mathbb{J}_{\alpha\mathbb{A}}(\mathbb{R}_{\alpha\mathbb{B}} - \alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}})$ is θ-averaged with

$$\theta = \frac{2\beta}{4\beta - \alpha}.$$

Proof of Theorem 28

Lemma 5.

For $\theta \in (0,1)$, $\mathbb{T}$ is θ-averaged if and only if

$$\|\mathbb{T}x - \mathbb{T}y\|^2 \leq \|x-y\|^2 - \frac{1-\theta}{\theta}\|\mathbb{T}x - x - \mathbb{T}y + y\|^2 \qquad \forall,\ x, y \in \mathbb{R}^n.$$

Proof. Note $\mathbb{T}$ is θ-averaged if and only if $\frac{1}{\theta}\mathbb{T} - \left(\frac{1}{\theta} - 1\right)\mathbb{I}$ is nonexpansive. The claim follows from

$$\begin{aligned} 0 &\geq \left\|\frac{1}{\theta}\mathbb{T}x - \left(\frac{1}{\theta}-1\right)x - \frac{1}{\theta}\mathbb{T}y + \left(\frac{1}{\theta}-1\right)y\right\|^2 - \|x-y\|^2 \\ &= \frac{1}{\theta}\left(\|\mathbb{T}x - \mathbb{T}y\|^2 + \frac{1-\theta}{\theta}\|\mathbb{T}x - x - \mathbb{T}y + y\|^2 - \|x-y\|^2\right). \end{aligned}$$

□

Proof of Theorem 28

For any $z^0, \hat{z}^0 \in \mathbb{R}^n$, let

$$\begin{aligned} x^{1/2} &= \mathbb{J}_{\alpha\mathbb{B}}(z^0) \\ x^1 &= \mathbb{J}_{\alpha\mathbb{A}}(2x^{1/2} - z^k - \alpha\mathbb{C}x^{1/2}) \\ z^1 &= z^0 + x^1 - x^{1/2} \end{aligned} \quad \Bigg| \quad \begin{aligned} \hat{x}^{1/2} &= \mathbb{J}_{\alpha\mathbb{B}}(\hat{z}^0) \\ \hat{x}^1 &= \mathbb{J}_{\alpha\mathbb{A}}(2\hat{x}^{1/2} - \hat{z}^k - \alpha\mathbb{C}\hat{x}^{1/2}) \\ \hat{z}^1 &= \hat{z}^0 + \hat{x}^1 - \hat{x}^{1/2}. \end{aligned}$$

Define

$$\begin{aligned} \tilde{\mathbb{B}}x^{1/2} &= \frac{1}{\alpha}(z^0 - x^{1/2}) \\ \tilde{\mathbb{A}}x^1 &= \frac{1}{\alpha}(2x^{1/2} - z^k - \alpha\mathbb{C}x^{1/2} - x^1) \end{aligned} \quad \Bigg| \quad \begin{aligned} \tilde{\mathbb{B}}\hat{x}^{1/2} &= \frac{1}{\alpha}(\hat{z}^0 - \hat{x}^{1/2}) \\ \tilde{\mathbb{A}}\hat{x}^1 &= \frac{1}{\alpha}(2\hat{x}^{1/2} - \hat{z}^k - \alpha\mathbb{C}\hat{x}^{1/2} - \hat{x}^1), \end{aligned}$$

Proof of Theorem 28

which implies

$$\tilde{\mathbb{B}}x^{1/2} \in \mathbb{B}x^{1/2} \qquad \tilde{\mathbb{B}}\hat{x}^{1/2} \in \mathbb{B}\hat{x}^{1/2}$$
$$\tilde{\mathbb{A}}x^{1} \in \mathbb{A}x^{1} \qquad \tilde{\mathbb{A}}\hat{x}^{1} \in \mathbb{A}\hat{x}^{1}.$$

Then

$$\begin{aligned}
\|z^1 - \hat{z}^1\|^2 &= \|z^0 - \hat{z}^0\|^2 - \frac{1-\theta}{\theta}\|z^1 - z^0 - \hat{z}^1 + \hat{z}^0\|^2 \\
&\quad - 2\alpha\langle \tilde{\mathbb{A}}x^1 - \tilde{\mathbb{A}}\hat{x}^1, x^1 - \hat{x}^1\rangle - 2\alpha\langle \tilde{\mathbb{B}}x^{1/2} - \tilde{\mathbb{B}}\hat{x}^{1/2}, x^{1/2} - \hat{x}^{1/2}\rangle \\
&\quad - 2\alpha\left(\langle \mathbb{C}x^{1/2} - \mathbb{C}\hat{x}^{1/2}, x^{1/2} - \hat{x}^{1/2}\rangle - \beta\|\mathbb{C}x^{1/2} - \mathbb{C}\hat{x}^{1/2}\|^2\right) \\
&\quad - \frac{\alpha^2}{2\beta}\left\|\tilde{\mathbb{A}}x^1 - \tilde{\mathbb{A}}\hat{x}^1 + \tilde{\mathbb{B}}x^{1/2} - \tilde{\mathbb{B}}\hat{x}^{1/2} - \frac{2\beta-\alpha}{\alpha}(\mathbb{C}x^{1/2} - \mathbb{C}\hat{x}^{1/2})\right\|^2 \\
&\le \|z^0 - \hat{z}^0\|^2 - \frac{1-\theta}{\theta}\|z^1 - z^0 - \hat{z}^1 + \hat{z}^0\|^2,
\end{aligned}$$

where the inequality follows from monotonicity of $\mathbb{A}$ and $\mathbb{B}$ and β-cocoercivity of $\mathbb{C}$. Finally, the claim follows from Lemma 5. □

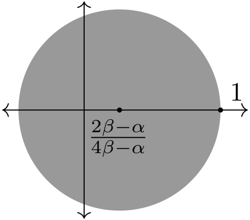
SRG of DYS

Let

$$\mathcal{T}_{\alpha,\beta} = \{\mathbb{I} - \mathbb{J}_{\alpha\mathbb{B}} + \mathbb{J}_{\alpha\mathbb{A}}(\mathbb{R}_{\alpha\mathbb{B}} - \alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}) \,|\, \mathbb{A}, \mathbb{B} \in \mathcal{M},\, \mathbb{C} \in \mathcal{C}_\beta\}.$$

be the class of DYS operators. Theorem 28 states $\mathcal{G}(\mathcal{T}_{\alpha,\beta}) \subseteq \mathcal{G}\left(\mathcal{N}_{\frac{2\beta}{4\beta-\alpha}}\right)$ for $\alpha \in (0, 2\beta)$ One can furthermore show

$$\mathcal{G}(\mathcal{T}_{\alpha,\beta}) = \mathcal{G}\left(\mathcal{N}_{\frac{2\beta}{4\beta-\alpha}}\right) =$$



Outline

Basic definitions

Scaled relative graphs

Operator and SRG transformations

Averagedness coefficients

Proofs of SRG theorems

Operator from SRG

$\mathcal{G}(\cdot)$ maps $\mathbb{A}$ to a subset of $\overline{\mathbb{C}}$. Below conversely maps z to $\mathbb{A}_z$ such that $z \in \mathcal{G}(\mathbb{A}_z)$.

Lemma 4.

Take any $z = z_r + z_i i \in \mathbb{C}$. Define $\mathbb{A}_z : \mathbb{R}^2 \to \mathbb{R}^2$ and $\mathbb{A}_\infty : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ as

$$\mathbb{A}_z \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \begin{bmatrix} z_r\zeta_1 - z_i\zeta_2 \\ z_r\zeta_2 + z_i\zeta_1 \end{bmatrix} \qquad \mathbb{A}_\infty(x) = \begin{cases} \mathbb{R}^2 & \text{if } x = 0 \\ \emptyset & \text{otherwise.} \end{cases}$$

Then,

$$\mathcal{G}(\mathbb{A}_z) = \{z, \bar{z}\}, \qquad \mathcal{G}(\mathbb{A}_\infty) = \{\infty\}.$$

If $\cong$ identifies $\mathbb{R}^2$ with $\mathbb{C}$,

$$\begin{bmatrix} x \\ y \end{bmatrix} \cong x + yi,$$

then $\mathbb{A}_z$ is complex multiplication with z:

$$\mathbb{A}_z \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} \cong z(\zeta_1 + \zeta_2 i).$$

Proof of Lemma 4

Write $z = r_z e^{i\theta_z}$. Consider any $x, y \in \mathbb{R}^2$ where $x \neq y$ and define $u = \mathbb{A}_z x$ and $v = \mathbb{A}_z y$. Then we can write

$$x - y = r_w \begin{bmatrix} \cos(\theta_w) \\ \sin(\theta_w) \end{bmatrix}$$

where $r_w > 0$, and

$$u - v = \mathbb{A}_z(x - y) \cong r_z r_w e^{i(\theta_z + \theta_w)}.$$

This gives us

$$\frac{\|u - v\|}{\|x - y\|} = r_z, \qquad \angle(u - v, x - y) = |\theta_z|, \quad \mathcal{G}(\mathbb{A}_z) = \left\{ r_z e^{i\theta_z}, r_z e^{-i\theta_z} \right\}.$$

Now consider $\mathbb{A}_\infty$. By definition, $\infty \in \mathcal{G}(\mathbb{A}_\infty)$. For any $u \in \mathbb{A}_\infty x$ and $v \in \mathbb{A}_\infty y$, we have $x = y = 0$, and therefore $\mathcal{G}(\mathbb{A}_\infty)$ contains no finite $z \in \mathbb{C}$. We conclude $\mathcal{G}(\mathbb{A}_\infty) = \{\infty\}$. □

Proof of Theorem 19

Characterize $\mathcal{G}(\mathcal{M})$. For any $\mathbb{A} \in \mathcal{M}$,

$$\operatorname{Re} z = \frac{\langle u - v, x - y \rangle}{\|x - y\|^2} \geq 0, \quad \forall\, u \in \mathbb{A}x,\ v \in \mathbb{A}y,\ x \neq y.$$

(Cf. page 10.) Therefore, $\mathcal{G}(\mathbb{A}) \backslash \{\infty\} \subseteq \{z \mid \operatorname{Re} z \geq 0\}$. On the other hand, given any $z \in \{z \mid \operatorname{Re} z \geq 0\}$, the operator $\mathbb{A}_z$ of Lemma 4 satisfies $\langle \mathbb{A}_z x - \mathbb{A}_z y, x - y \rangle \geq 0$ for any $x, y \in \mathbb{R}^2$, i.e., $\mathbb{A}_z \in \mathcal{M}$, and $\mathcal{G}(\mathbb{A}_z) = \{z, \bar{z}\}$. Therefore, $z \in \mathcal{G}(\mathbb{A}_z) \subset \mathcal{G}(\mathcal{M})$, and we conclude $\{z \mid \operatorname{Re} z \geq 0\} \subseteq \mathcal{G}(\mathcal{M})$. Finally, note that $\infty \in \mathcal{G}(\mathcal{M})$ is equivalent to saying that there exists a multi-valued operator in $\mathcal{M}$. The $\mathbb{A}_\infty$ of Lemma 4 is one such example.

The other SRGs follow from a similar reasoning. □

Proof outline of Theorem 20

$\partial \mathcal{F}_{0,\infty} \subset \mathcal{M}$ and Theorem 19 implies
$\mathcal{G}(\partial \mathcal{F}_{0,\infty}) \subseteq \mathcal{G}(\mathcal{M}) = \{z \in \mathbb{C} \mid \operatorname{Re} z \geq 0\} \cup \{\infty\}$.

With basic computation, we can verify $f \colon \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = |x|$ satisfies $\mathcal{G}(\partial f) = \{z \in \mathbb{C} \mid \operatorname{Re} z \geq 0\} \cup \{\infty\}$. This tells us $\{z \in \mathbb{C} \mid \operatorname{Re} z \geq 0\} \cup \{\infty\} \subseteq \mathcal{G}(\partial \mathcal{F}_{0,\infty})$.

The other SRGs follow from a similar reasoning. □

Proof of Theorem 21

Proof of Theorem 22

Since $\mathcal{A}$ and $\mathcal{B}$ are SRG-full

$$\begin{aligned}
\mathcal{G}(\mathbb{C}) \subseteq \mathcal{G}(\mathcal{A} \cap \mathcal{B}) \subseteq \mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B}) \quad &\Rightarrow \quad \mathcal{G}(\mathbb{C}) \subseteq \mathcal{G}(\mathcal{A}) \text{ and } \mathcal{G}(\mathbb{C}) \subseteq \mathcal{G}(\mathcal{B}) \\
&\Rightarrow \quad \mathbb{C} \in \mathcal{A} \text{ and } \mathbb{C} \in \mathcal{B} \\
&\Rightarrow \quad \mathbb{C} \in \mathcal{A} \cap \mathcal{B}
\end{aligned}$$

for an operator $\mathbb{C}$, and we conclude $\mathcal{A} \cap \mathcal{B}$ is SRG-full.

Assume $z \in \mathbb{C}$ satisfies $\{z, \bar{z}\} \subseteq \mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B})$. Then $\mathbb{A}_z$ of Lemma 4 satisfies $\mathcal{G}(\mathbb{A}_z) = \{z, \bar{z}\} \subseteq \mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B})$. Since $\mathcal{A}$ and $\mathcal{B}$ are SRG-full, $\mathbb{A}_z \in \mathcal{A}$ and $\mathbb{A}_z \in \mathcal{B}$ and $\{z, \bar{z}\} = \mathcal{G}(\mathbb{A}_z) \subseteq \mathcal{G}(\mathcal{A} \cap \mathcal{B})$. If $\infty \in \mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B})$, then a similar argument using $\mathbb{A}_\infty$ of Lemma 4 proves $\infty \in \mathcal{G}(\mathcal{A} \cap \mathcal{B})$. Therefore $\mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B}) \subseteq \mathcal{G}(\mathcal{A} \cap \mathcal{B})$. Since the other containment $\mathcal{G}(\mathcal{A} \cap \mathcal{B}) \subseteq \mathcal{G}(\mathcal{A}) \cap \mathcal{G}(\mathcal{B})$ holds by definition, we have the equality. □

Proof of Theorem 23

$\mathcal{G}(\alpha\mathbb{A}) = \alpha\mathcal{G}(\mathbb{A})$ follows from the definition of the SRG, and $\mathcal{G}(\mathbb{I} + \mathbb{A}) = 1 + \mathcal{G}(\mathbb{A})$ follows from

$$\operatorname{Re} z = \frac{\langle u - v, x - y\rangle}{\|x - y\|^2}, \qquad \operatorname{Im} z = \pm\frac{\|P_{\{x-y\}^\perp}(u - v)\|}{\|x - y\|}.$$

The scaling and translation operations are reversible and $\mathcal{G}((1/\alpha)\mathcal{A}) = (1/\alpha)\mathcal{G}(\mathcal{A})$ and $\mathcal{G}(\mathcal{A} - \mathbb{I}) = \mathcal{G}(A) - 1$. For any $\mathbb{B}\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$,

$$\mathcal{G}(\mathbb{B}) \subseteq \mathcal{G}(\alpha\mathcal{A}) \;\Rightarrow\; \mathcal{G}((1/\alpha)\mathbb{B}) \subseteq \mathcal{G}(\mathcal{A}) \;\Rightarrow\; (1/\alpha)\mathbb{B} \in \mathcal{A} \;\Rightarrow\; \mathbb{B} \in \alpha\mathcal{A},$$

and we conclude $\alpha\mathcal{A}$ is SRG-full. By a similar reasoning, $\mathbb{I} + \mathcal{A}$ is SRG-full. □

Proof of Theorem 24

The equivalence of non-zero finite points, i.e.,

$$\mathcal{G}(\mathbb{A}^{-1})\backslash\{0,\infty\} = (\mathcal{G}(\mathbb{A})\backslash\,\{0,\infty\})^{-1},$$

follows from

$$\mathcal{G}(\mathbb{A})\backslash\{0,\infty\} = \left\{ \frac{\|u-v\|}{\|x-y\|} \exp\left[\pm i\angle(u-v, x-y)\right] \,\middle|\, (x,u),(y,v) \in \mathbb{A},\, x \neq y,\, u \neq v \right\}$$

and

$$\begin{aligned}
&\mathcal{G}(\mathbb{A}^{-1})\backslash\{0,\infty\} \\
&\quad = \left\{ \frac{\|x-y\|}{\|u-v\|} \exp\left[\pm i\angle(x-y, u-v)\right] \,\middle|\, (u,x),(v,y) \in \mathbb{A}^{-1},\, x \neq y,\, u \neq v \right\} \\
&\quad = \left\{ \frac{\|x-y\|}{\|u-v\|} \exp\left[\pm i\angle(u-v, x-y)\right] \,\middle|\, (x,u),(y,v) \in \mathbb{A},\, x \neq y,\, u \neq v \right\} \\
&\quad = \left(\mathcal{G}(\mathbb{A})\backslash\{0,\infty\}\right)^{-1}.
\end{aligned}$$

Proof of Theorem 24

The equivalence of the zero and infinite points follow from

$$\begin{aligned} \infty \in \mathcal{G}(\mathbb{A}) \quad &\Leftrightarrow \quad \exists\,(x,u),(x,v) \in \mathbb{A},\, u \neq v \\ &\Leftrightarrow \quad \exists\,(u,x),(v,x) \in \mathbb{A}^{-1},\, u \neq v \\ &\Leftrightarrow \quad 0 \in \mathcal{G}(\mathbb{A}^{-1}). \end{aligned}$$

With the same argument, we have $0 \in \mathcal{G}(\mathbb{A}) \Leftrightarrow \infty \in \mathcal{G}(\mathbb{A}^{-1})$.

The inversion operation is reversible. Therefore, for any $\mathbb{B} \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^n$,

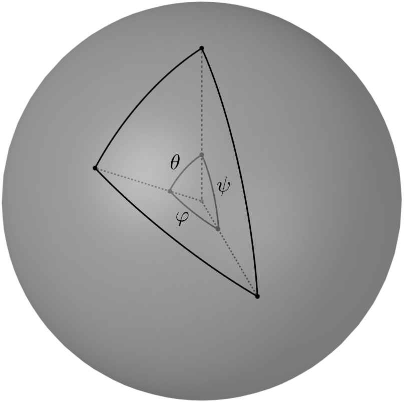
$$\mathcal{G}(\mathbb{B}) \subseteq \mathcal{G}(\mathcal{A}^{-1}) \;\Rightarrow\; \mathcal{G}(\mathbb{B}^{-1}) \subseteq \mathcal{G}(\mathcal{A}) \;\Rightarrow\; \mathbb{B}^{-1} \in \mathcal{A} \;\Rightarrow\; \mathbb{B} \in \mathcal{A}^{-1},$$

and we conclude $\mathcal{A}^{-1}$ is SRG-full. □

Proof of Theorem 25

Spherical triangle inequality

Spherical triangle inequality: $|\theta - \varphi| \leq \psi \leq \theta + \varphi$



For any nonzero $a, b, c \in \mathbb{R}^n$,

$$|\angle(a,b) - \angle(b,c)| \leq \angle(a,c) \leq \angle(a,b) + \angle(b,c).$$

Proof of Theorem 26

We first show $\mathcal{G}(\mathcal{A}\mathcal{B}) \supseteq \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$. Assume $\mathcal{G}(\mathcal{A}) \neq \emptyset$ and $\mathcal{G}(\mathcal{B}) \neq \emptyset$ as otherwise there is nothing to show. Let $z \in \mathcal{G}(\mathcal{A})$ and $w \in \mathcal{G}(\mathcal{B})$ and let $\mathbb{A}_z$ and $\mathbb{A}_w$ be their corresponding operators as defined in Lemma 4. Then it is straightforward to see that $\mathbb{A}_z\mathbb{A}_w$ corresponds to complex multiplication with respect to zw, and $zw \in \mathcal{G}(\mathbb{A}_z\mathbb{A}_w) \subseteq \mathcal{G}(\mathcal{A}\mathcal{B})$.

Proof of Theorem 26

Next, we show $\mathcal{G}(\mathcal{AB}) \subseteq \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$. Let $\mathbb{A} \in \mathcal{A}$ and $\mathbb{B} \in \mathcal{B}$. Consider $(u, s), (v, t) \in \mathbb{A}$ and $(x, u), (y, v) \in \mathbb{B}$, where $x \neq y$. This implies $(x, s), (y, t) \in \mathbb{AB}$. Define

$$z = \frac{\|s - t\|}{\|x - y\|} \exp\left[i\angle(s - t, x - y)\right].$$

Consider the case $u = v$. Then $0 \in \mathcal{G}(\mathcal{B})$. Moreover, $s = t$, since $\mathbb{A}$ is single-valued (by the assumption $\infty \notin \mathcal{G}(\mathcal{A})$), and $z = 0$. Therefore, $z = 0 \in \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$.

Next, consider the case $u \neq v$. Define

$$z_A = \frac{\|s - t\|}{\|u - v\|} e^{i\varphi_A}, \quad z_B = \frac{\|u - v\|}{\|x - y\|} e^{i\varphi_B},$$

where $\varphi_A = \angle(s - t, u - v)$ and $\varphi_B = \angle(u - v, x - y)$.

Proof of Theorem 26

Consider the case where $\mathcal{A}$ satisfies the right-arc property. Using the spherical triangle inequality (further discussed in the appendix) we see that either $\varphi_A \geq \varphi_B$ and

$$\begin{aligned} z &\in \frac{\|s-t\|}{\|u-v\|}\frac{\|u-v\|}{\|x-y\|}\exp\left[i[\varphi_A-\varphi_B,\varphi_A+\varphi_B]\right] \\ &\subseteq \frac{\|s-t\|}{\|u-v\|}\frac{\|u-v\|}{\|x-y\|}\exp\left[i[\varphi_B-\varphi_A,\varphi_B+\varphi_A]\right] \\ &= z_B \mathrm{Arc}^+\left(z_A,\overline{z_A}\right) \end{aligned}$$

or $\varphi_A < \varphi_B$ and

$$\begin{aligned} z &\in \frac{\|s-t\|}{\|u-v\|}\frac{\|u-v\|}{\|x-y\|}\exp\left[i[\varphi_B-\varphi_A,\varphi_B+\varphi_A]\right] \\ &= z_B \mathrm{Arc}^+\left(z_A,\overline{z_A}\right). \end{aligned}$$

This gives us

$$z \in \underbrace{z_B}_{\in\mathcal{G}(\mathcal{B})}\underbrace{\mathrm{Arc}^+\left(z_A,\overline{z_A}\right)}_{\subseteq\mathcal{G}(\mathcal{A})} \subseteq \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B}).$$

That $\bar{z} \in \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$ follows from the same argument. That $z, \bar{z} \in \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$ when instead $\mathcal{B}$ satisfies the right-arc property follows from the same argument.

Proof of Theorem 26

Putting everything together, we conclude $\mathcal{G}(\mathcal{A}\mathcal{B}) = \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$ when $\mathcal{A}$ or $\mathcal{B}$ satisfies the right-arc property. When $\mathcal{A}$ satisfies the left-arc property, $-\mathcal{A}$ satisfies the right-arc property. So

$$-\mathcal{G}(\mathcal{A}\mathcal{B}) = \mathcal{G}(-\mathcal{A}\mathcal{B}) = \mathcal{G}(-\mathcal{A})\mathcal{G}(\mathcal{B}) - \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$$

by Theorem 23, and we conclude $\mathcal{G}(\mathcal{A}\mathcal{B}) = \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$. When $\mathcal{B}$ satisfies the left-arc property, $\mathcal{B} \circ (-\mathbb{I})$ satisfies the right-arc property. So

$$-\mathcal{G}(\mathcal{A}\mathcal{B}) = \mathcal{G}(\mathcal{A}\mathcal{B} \circ (-I)) = \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B} \circ (-\mathbb{I})) = -\mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$$

by Theorem 23, and we conclude $\mathcal{G}(\mathcal{A}\mathcal{B}) = \mathcal{G}(\mathcal{A})\mathcal{G}(\mathcal{B})$. $\square$