A parallel Douglas–Rachford Algorithm for Data on Hadamard Manifolds.

Ronny Bergmann

Technische Universität Chemnitz

Riemannian Geometry in Optimization for Learning,
23rd International Symposium on Mathematical Programming,

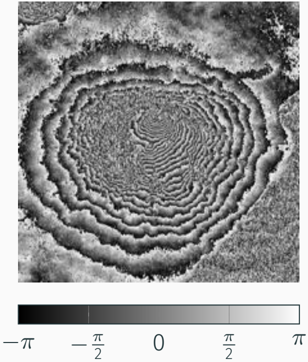
Bordeaux, July 2nd, 2018

Manifold-valued Image Processing

Manifold-valued Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,…
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)



InSAR-Data of Mt. Vesuvius

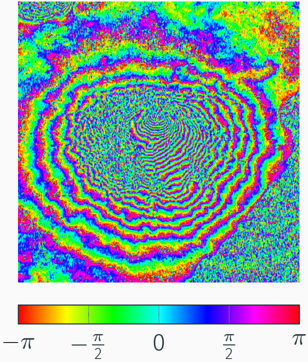
[Rocca, Prati, Guarnieri 1997]

phase-valued data, $\mathcal{M} = \mathbb{S}^1$

Manifold-valued Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)



InSAR-Data of Mt. Vesuvius

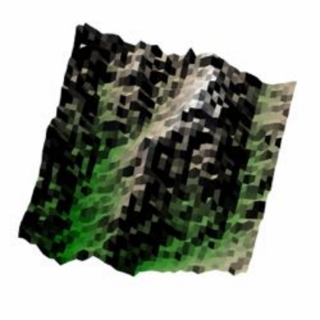
[Rocca, Prati, Guarnieri 1997]

phase-valued data, $\mathcal{M} = \mathbb{S}^1$

Manifold-valued Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)



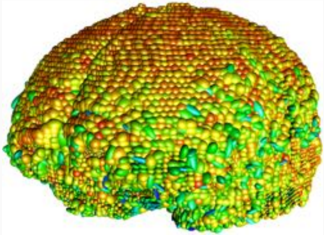
National elevation dataset [Gesch, Evans, Mauck, 2009]

directional data, $\mathcal{M} = \mathbb{S}^2$

Manifold-valued Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)



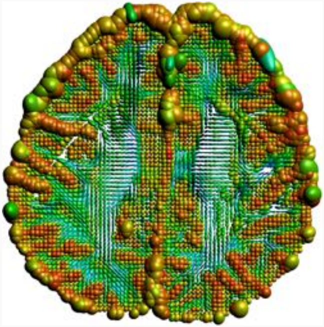
diffusion tensors in human brain
from the Camino dataset
http://cmic.cs.ucl.ac.uk/camino

sym. pos. def. matrices, $\mathcal{M} = \mathrm{SPD}(3)$

Manifold-valued Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,…
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)



horizontal slice #28
from the Camino dataset
http://cmic.cs.ucl.ac.uk/camino

sym. pos. def. matrices, $\mathcal{M} = \mathrm{SPD}(3)$

Manifold-valued Images

New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals, GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)

EBSD example from the MTEX toolbox
[Bachmann, Hielscher, since 2005]

Rotations (mod. symmetry),
$\mathcal{M} = \mathrm{SO}(3)(/\mathcal{S})$.

Manifold-valued Images

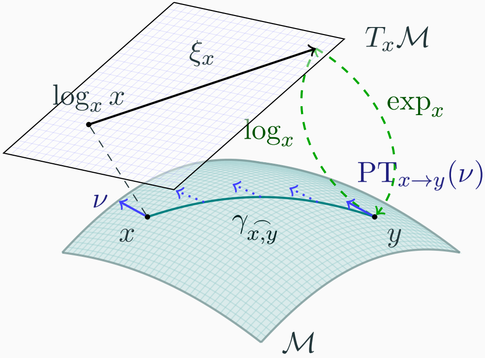
New data acquisition modalities lead to non-Euclidean range

- Interferometric synthetic aperture radar (InSAR)
- Surface normals,
  GPS data, wind, flow,...
- Diffusion tensors in magnetic resonance imaging (DT-MRI), covariance matrices
- Electron backscattered diffraction (EBSD)

Common properties

- Range of values is a Riemannian manifold
- Tasks from “classical” image processing

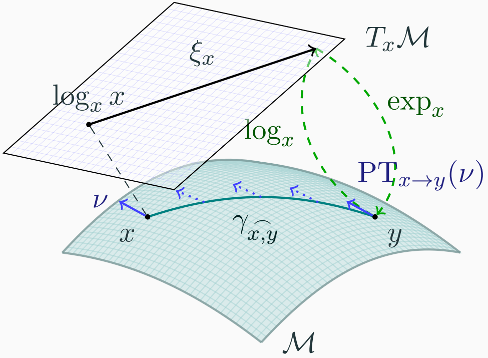
A d-dimensional Riemannian Manifold $\mathcal{M}$



A d-dimensional Riemannian manifold can be informally defined as a set $\mathcal{M}$ covered with a 'suitable' collection of charts, that identify subsets of $\mathcal{M}$ with open subsets of $\mathbb{R}^d$ and a continously varying inner product on the tangential spaces.

[Absil, Mahony, Sepulchre, 2008]

A d-dimensional Riemannian Manifold $\mathcal{M}$



Geodesic $\gamma_{\widehat{x,y}}$ shortest connection (on $\mathcal{M}$) between $x, y \in \mathcal{M}$
Tangent space $\mathrm{T}_x\mathcal{M}$ at x, with inner product $\langle\cdot,\cdot\rangle_x$
Logarithmic map $\log_x y = \dot\gamma_{\widehat{x,y}}(0)$ "speed towards y"
Exponential map $\exp_x \xi_x = \gamma(1)$, where $\gamma(0) = x$, $\dot\gamma(0) = \xi_x$
Parallel transport $\mathrm{PT}_{x\to y}(\nu)$ of $\nu \in \mathrm{T}_x\mathcal{M}$ along $\gamma_{\widehat{x,y}}$

Variational Methods for Euclidean Data

- Let $\mathcal{V} \subseteq \mathcal{G} = \{1, \ldots, N\} \times \{1, \ldots, M\}$
- **Task:** Given noisy, possibly lossy, data $f \colon \mathcal{V} \to \mathbb{R}^m$:
  Reconsrtuct the original image $u_0 \colon \mathcal{G} \to \mathbb{R}^m$
- **Approach:** Compute minimizer u^* of a variational model

$$\mathcal{E}(u) \coloneqq \underbrace{\mathcal{D}(u; f)}_{\text{data term}} + \underbrace{\alpha\, \mathcal{R}(u)}_{\text{regularizer/prior}}, \qquad \alpha > 0.$$

Variational Methods for Euclidean Data

- Let $\mathcal{V} \subseteq \mathcal{G} = \{1, \ldots, N\} \times \{1, \ldots, M\}$
- **Task:** Given noisy, possibly lossy, data $f\colon \mathcal{V} \to \mathbb{R}^m$:
  Reconsrtuct the original image $u_0\colon \mathcal{G} \to \mathbb{R}^m$
- **Approach:** Compute minimizer u^* of a variational model

$$\mathcal{E}(u) := \underbrace{\mathcal{D}(u;f)}_{\text{data term}} + \underbrace{\alpha\,\mathcal{R}(u)}_{\text{regularizer/prior}}, \qquad \alpha > 0.$$

- L^2 data term: $D(u;f) = \sum_{i\in\mathcal{V}} \lVert u_i - f_i \rVert^2$

Variational Methods for Euclidean Data

- Let $\mathcal{V} \subseteq \mathcal{G} = \{1, \ldots, N\} \times \{1, \ldots, M\}$
- **Task:** Given noisy, possibly lossy, data $f \colon \mathcal{V} \to \mathbb{R}^m$:
  Reconsrtuct the original image $u_0 \colon \mathcal{G} \to \mathbb{R}^m$
- **Approach:** Compute minimizer u^* of a variational model

$$\mathcal{E}(u) \coloneqq \underbrace{\mathcal{D}(u; f)}_{\text{data term}} + \underbrace{\alpha\,\mathcal{R}(u)}_{\text{regularizer/prior}}, \qquad \alpha > 0.$$

- L^2 data term: $D(u; f) = \sum_{i \in \mathcal{V}} \|u_i - f_i\|^2$
- regularizer: Total Variation (TV) [Rudin, Osher, Fatemi, 1992]
  known to be edge preserving, for example:

  anisotropic TV: $\mathcal{R}(u) \coloneqq \sum_{i,j} \left(\|u_{i+1,j} - u_{i,j}\| + \|u_{i,j+1} - u_{i,j}\|\right)$

Variational Methods for Euclidean Data

- Let $\mathcal{V} \subseteq \mathcal{G} = \{1, \ldots, N\} \times \{1, \ldots, M\}$
- **Task:** Given noisy, possibly lossy, data $f\colon \mathcal{V} \to \mathbb{R}^m$:
  Reconsrtuct the original image $u_0\colon \mathcal{G} \to \mathbb{R}^m$
- **Approach:** Compute minimizer u^* of a variational model

$$\mathcal{E}(u) := \underbrace{\mathcal{D}(u; f)}_{\text{data term}} + \underbrace{\alpha\, \mathcal{R}(u)}_{\text{regularizer/prior}}, \qquad \alpha > 0.$$

- L^2 data term: $D(u; f) = \sum_{i \in \mathcal{V}} \|u_i - f_i\|^2$
- regularizer: Total Variation (TV) [Rudin, Osher, Fatemi, 1992]
  known to be edge preserving, for example:

  anisotropic TV: $\mathcal{R}(u) := \sum_{i,j} \bigl(\|u_{i+1,j} - u_{i,j}\| + \|u_{i,j+1} - u_{i,j}\|\bigr)$
- high dimensional, non-differentiable, convex

Total Variation of Manifold-valued Data

Note. All summands of the ROF model are (squared) distances.

Let $d_{\mathcal{M}} \colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ denote the geodesic distance on $\mathcal{M}$. Then the TV model for manifold-valued data $f \colon \mathcal{V} \to \mathcal{M}$ reads

$$\mathcal{E}(u) = \sum_{i \in \mathcal{V}} d^2_{\mathcal{M}}(u_i, f_i) + \alpha \sum_{i,j} \Big(d_{\mathcal{M}}(u_{i+1,j}, u_{i,j}) + d_{\mathcal{M}}(u_{i,j+1}, u_{i,j}) \Big).$$

This can be minmized with

- functional Lifting
  [Cremers,Strekalovski, 2011/13; Lellmann, Kötters, Strekalovski, Cremers, 2013]
- Cyclic Proximal Point-Algorithm
  [Bačák, 2013; Weinmann, Storath, Demaret, 2014]
- discrete Gradient, Gradient descent, quasi-Newton
  [RB, Fitschen, Persch, Steidl, 2017; Celledoni, Eidnes, Owren, Ringholm, 2018; RB, Chan, Hielscher, Persch, Steidl, 2016]

The Douglas-Rachford Algorithm

Proximum and Reflection

For $\varphi\colon \mathcal{M}^n \to (-\infty, +\infty]$ and $\lambda > 0$ the Proximum is defined by [Moreau, 1962; Rockafellar, 1976; Ferreira, Oliveira, 2002]

$$\operatorname{prox}_{\lambda\varphi}(g) \coloneqq \operatorname*{arg\,min}_{u\in\mathcal{M}^n} \frac{1}{2}\sum_{i=1}^{n} d_{\mathcal{M}}(u_i, g_i)^2 + \lambda\varphi(u).$$

! For a minimizer u^* of φ it holds: $\operatorname{prox}_{\lambda\varphi}(u^*) = u^*$.

Proximum and Reflection

For $\varphi\colon \mathcal{M}^n \to (-\infty, +\infty]$ and $\lambda > 0$ the Proximum is defined by [Moreau, 1962; Rockafellar, 1976; Ferreira, Oliveira, 2002]

$$\operatorname{prox}_{\lambda\varphi}(g) := \operatorname*{arg\,min}_{u \in \mathcal{M}^n} \frac{1}{2} \sum_{i=1}^{n} d_{\mathcal{M}}(u_i, g_i)^2 + \lambda\varphi(u).$$

! For a minimizer u^* of φ it holds: $\operatorname{prox}_{\lambda\varphi}(u^*) = u^*$.

A map $\mathcal{R}_p$ is called Reflection on $\mathcal{M}$, if

$$\mathcal{R}_p(p) = p \quad \text{and} \quad D_p\mathcal{R}_p = -I \quad \text{hold.}$$

Analogous: Reflection with respect to $\lambda\varphi$

$$\mathcal{R}_{\lambda\varphi}(x) = \mathcal{R}_{\operatorname{prox}_{\lambda\varphi}(x)}(x)$$

Proximum and Reflection

For $\varphi\colon \mathcal{M}^n \to (-\infty, +\infty]$ and $\lambda > 0$ the Proximum is defined by

[Moreau, 1962; Rockafellar, 1976; Ferreira, Oliveira, 2002]

$$\operatorname{prox}_{\lambda\varphi}(g) := \operatorname*{arg\,min}_{u \in \mathcal{M}^n} \frac{1}{2} \sum_{i=1}^{n} d_{\mathcal{M}}(u_i, g_i)^2 + \lambda\varphi(u).$$

! For a minimizer u^* of φ it holds: $\operatorname{prox}_{\lambda\varphi}(u^*) = u^*$.

A map $\mathcal{R}_p$ is called Reflection on $\mathcal{M}$, if

$$\mathcal{R}_p(p) = p \quad \text{and} \quad D_p\mathcal{R}_p = -I \quad \text{hold.}$$

Analogous: Reflection with respect to $\lambda\varphi$

$$\mathcal{R}_{\lambda\varphi}(x) = \mathcal{R}_{\operatorname{prox}_{\lambda\varphi}(x)}(x)$$

Example. On $\mathbb{R}^m$ we have $\mathcal{R}_p(x) = 2p - x = p - (x - p)$.

Proximal Maps for the Summands of the ROF Model

Theorem **[Oliveira, Ferreira, 2002]**
For $f \in \mathcal{M}$, $\varphi\colon \mathcal{M} \to \mathbb{R}, \varphi(x) = d^2_{\mathcal{M}}(x, f)$ and $\lambda > 0$.
the proximal map is given by

$$\operatorname{prox}_{\lambda\varphi}(x) = \gamma_{\overset{\frown}{x,f}}\left(\tfrac{\lambda}{1+\lambda}\right)$$

Theorem **[Weinmann, Storath, Demaret, 2014]**
For $\varphi\colon \mathcal{M} \times \mathcal{M}, \varphi(x, y) = d_{\mathcal{M}}(x, y)$, $\lambda > 0$, the proximal map is given by

$$\operatorname{prox}_{\lambda\varphi}(x, y) = \begin{cases} \left(\gamma_{\overset{\frown}{x,y}}\left(\tfrac{\lambda}{d_{\mathcal{M}}(x,y)}\right), \gamma_{\overset{\frown}{x,y}}\left(1 - \tfrac{\lambda}{d_{\mathcal{M}}(x,y)}\right)\right) & \text{if } \lambda < \frac{d_{\mathcal{M}}(x,y)}{2}, \\ \left(\gamma_{\overset{\frown}{x,y}}\left(\tfrac{1}{2}\right), \gamma_{\overset{\frown}{x,y}}\left(\tfrac{1}{2}\right)\right) & \text{else.} \end{cases}$$

The Douglas-Rachford Splitting in Euclidean Space

Goal: Minimize

$$\operatorname*{arg\,min}_{x \in \mathbb{R}^n} \varphi(x) + \psi(x)$$

using a splitting approach

- for linear operators and PDEs [Douglas, Rachford, 1956]
- for monotone inclusion problems [Lions, Mercier, 1979, Eckstein, 1989]
- applications to image processing [Combettes, Pesquet, 2007]

The iteration of the Douglas-Rachford Algorithm reads

$$t^{(k+1)} := \frac{1}{2} t^{(k)} + \frac{1}{2} \mathcal{R}_{\lambda\varphi}\big(\mathcal{R}_{\lambda\psi}(t^{(k)})\big), \quad k \in \mathbb{N}_0, t^{(0)} \in \mathbb{R}^n,$$

and is related to the minimizer by $x^\star = \operatorname{prox}_{\lambda\psi}(\hat{t})$.

The Douglas-Rachford Splitting in Euclidean Space

Goal: Minimize

$$\operatorname*{arg\,min}_{x \in \mathbb{R}^n} \varphi(x) + \psi(x)$$

using a splitting approach

- for linear operators and PDEs [Douglas, Rachford, 1956]
- for monotone inclusion problems [Lions, Mercier, 1979, Eckstein, 1989]
- applications to image processing [Combettes, Pesquet, 2007]

The iteration of the Douglas-Rachford Algorithm reads

$$t^{(k+1)} := \beta t^{(k)} + (1-\beta)\mathcal{R}_{\lambda\varphi}\big(\mathcal{R}_{\lambda\psi}(t^{(k)})\big), \quad k \in \mathbb{N}_0, t^{(0)} \in \mathbb{R}^n, \beta \in (0,1),$$

and is related to the minimizer by $x^\star = \operatorname{prox}_{\lambda\psi}(\hat{t})$.

Hadamard Manifolds and Symmetric Spaces

A manifold $\mathcal{H}$ is called Hadamard manifold, if

$$d^2_{\mathcal{M}}(x,v) + d^2_{\mathcal{M}}(y,w) \leq d^2_{\mathcal{M}}(x,w) + d^2_{\mathcal{M}}(y,v) + 2d_{\mathcal{M}}(x,y)^2_{\mathcal{M}}(v,w)$$

holds for all $x,y,v,w \in \mathcal{H}$, i.e. we have a nonpositive sectional curvature. Then

- geodesics $\gamma_{\widehat{x,y}}\colon [0,1] \to \mathcal{H}$ are unique
- $\mathcal{C} \subset \mathcal{H}$ is convex, if $\gamma_{\widehat{x,y}} \subset \mathcal{C}$ for all $x,y \in \mathcal{C}$
- $\varphi\colon \mathcal{H} \to (-\infty,\infty]$ is convex on $\mathcal{C}$ if $\varphi \circ \gamma_{\widehat{x,y}}$ is convex
- The reflection reads $\mathcal{R}_p(x) = \exp_p(-\log_p x)$

$\mathcal{H}$ is called symmetric, if $\mathcal{R}_p$ is a isometry for all p.

The Douglas-Rachford Algorithm (DRA)

For $\varphi, \psi \in \Gamma^0(\mathcal{H})$ (proper, convex, lsc.)

Goal: Find minimizer

$$x^\star \in \operatorname*{arg\,min}_{x \in \mathcal{H}} \varphi(x) + \psi(x)$$

Iteration: For some $t^{(0)} \in \mathcal{H}$ compute the Krasnoselskii-Mann-iteration, $k \in \mathbb{N}_0$,

[RB, Persch, Steidl, 2016]

$$\begin{aligned} s^{(k)} &= \mathcal{R}_{\lambda\varphi}(\mathcal{R}_{\lambda\psi}(t^{(k)})) \\ t^{(k+1)} &= \gamma_{\overset{\frown}{t^{(k)}, s^{(k)}}}(\beta_k) \end{aligned}$$

with $\beta_k \in (0,1)$ and $\sum_{k \in \mathbb{N}} \beta_k(1 - \beta_k) = \infty$

Convergence of the DRA

Theorem **[Kakavandi 2013]**

Let $\mathcal{R}_{\lambda\varphi}$,$\mathcal{R}_{\lambda\psi}$ are nonexpansive and hence $T = \mathcal{R}_{\lambda\varphi} \circ \mathcal{R}_{\lambda\psi}$ is nonexpansive. Let T possess a fix point $\hat{t}$.

Then the DRA converges for every start point $t^{(0)} \in \mathcal{H}$ to a fix point $\hat{t}$ of T.

Convergence of the DRA

Theorem **[Kakavandi 2013]**

Let $\mathcal{R}_{\lambda\varphi}$,$\mathcal{R}_{\lambda\psi}$ are nonexpansive and hence $T = \mathcal{R}_{\lambda\varphi} \circ \mathcal{R}_{\lambda\psi}$ is nonexpansive. Let T possess a fix point $\hat{t}$.

Then the DRA converges for every start point $t^{(0)} \in \mathcal{H}$ to a fix point $\hat{t}$ of T.

Theoreom **[RB, Persch, Steidl, 2016]**

Let $\varphi, \psi \in \Gamma^0(\mathcal{H})$, let there be a minimizer $x^\star$ of $\varphi + \psi$, and let $T = \mathcal{R}_{\lambda\varphi} \circ \mathcal{R}_{\lambda\psi}$ be nonexpansive.

Then there exists for every $x^\star$ a fix point $\hat{t}$ of T, such that

$$x^\star = \operatorname{prox}_{\lambda\psi}(\hat{t})$$

holds. Further, for every $\hat{t}$, the point $\operatorname{prox}_{\lambda\psi}(\hat{t})$ is a minimizer of $\varphi + \psi$.

Parallelization

Given: $\varphi_i \in \Gamma^0(\mathcal{H}^m),\ i = 1, \ldots, c$

Goal: Find $x^\star \in \operatorname*{arg\,min}_{x \in \mathcal{H}^m} \sum_{i=1}^{c} \varphi_i(x)$

Employ: $\Phi(\mathbf{x}) \coloneqq \sum_{i=1}^{c} \varphi_i(x_i), \qquad \mathbf{x} = (x_1, \ldots, x_c)^{\mathrm{T}} \in \mathcal{H}^{mc}$

and D $\coloneqq \left\{\mathbf{x} \in \mathcal{H}^{mc} \colon x_1 = \ldots = x_c \in \mathcal{H}^m\right\} \subset \mathcal{H}^{mc}$

Parallelization

Given: $\varphi_i \in \Gamma^0(\mathcal{H}^m),\ i = 1, \dots, c$

Goal: Find $x^\star \in \operatorname*{arg\,min}_{x \in \mathcal{H}^m} \sum_{i=1}^{c} \varphi_i(x) \ \Leftrightarrow\ \mathbf{x}^\star \in \operatorname*{arg\,min}_{\mathbf{x} \in \mathcal{H}^{mc}} \Phi(\mathbf{x}) + \iota_{\mathsf{D}}(\mathbf{x})$

Employ: $\Phi(\mathbf{x}) \coloneqq \sum_{i=1}^{c} \varphi_i(x_i), \qquad \mathbf{x} = (x_1, \dots, x_c)^{\mathrm{T}} \in \mathcal{H}^{mc}$

and D $\coloneqq \left\{\mathbf{x} \in \mathcal{H}^{mc} \colon x_1 = \ldots = x_c \in \mathcal{H}^m\right\} \subset \mathcal{H}^{mc}$

Parallelization

Given: $\varphi_i \in \Gamma^0(\mathcal{H}^m),\ i = 1, \ldots, c$

Goal: Find $x^\star \in \operatorname*{arg\,min}_{x \in \mathcal{H}^m} \sum_{i=1}^{c} \varphi_i(x) \;\Leftrightarrow\; \mathbf{x}^\star \in \operatorname*{arg\,min}_{\mathbf{x} \in \mathcal{H}^{mc}} \Phi(\mathbf{x}) + \iota_{\mathrm{D}}(\mathbf{x})$

Employ: $\Phi(\mathbf{x}) := \sum_{i=1}^{c} \varphi_i(x_i), \qquad \mathbf{x} = (x_1, \ldots, x_c)^{\mathrm{T}} \in \mathcal{H}^{mc}$

and $\mathrm{D} := \{\mathbf{x} \in \mathcal{H}^{mc} \colon x_1 = \ldots = x_c \in \mathcal{H}^m\} \subset \mathcal{H}^{mc}$

We obtain the parallel Douglas-Rachford algorithm (PDRA):
For a start point $\mathbf{t}^{(0)} \in \mathcal{H}^{mc}$, and $k = 0, \ldots,$ compute

[RB, Persch, Steidl, 2016]

$$\begin{aligned} \mathbf{s}^{(k)} &= \mathcal{R}_{\lambda\Phi}\mathcal{R}_{\iota_{\mathrm{D}}}\big(\mathbf{t}^{(k)}\big), \\ \mathbf{t}^{(k+1)} &= \gamma_{\overset{\frown}{\mathbf{t}^{(k)},\mathbf{s}^{(k)}}}\big(\beta_k\big) \end{aligned}$$

$$\Rightarrow x^\star = \operatorname{prox}_{\iota_{\mathrm{D}}}(\hat{\mathbf{t}})_1 = \operatorname*{arg\,min}_{x \in \mathcal{H}^m} \sum_{i=1}^{c} d^2_{\mathcal{H}}(\hat{t}_k, x) \quad \text{(Fréchet mean)}$$

Convergence of PDRA for the ROF model

Are the reflections of the ROF model non-expansive?

Convergence of PDRA for the ROF model

Are the reflections of the ROF model non-expansive?

Theorem: Nonexpansiveness of $\mathcal{R}_{\lambda\Phi}$ **[RB, Persch, Steidl, 2016]**
Let $\mathcal{H}$ be a symmetric Hadamard manifold, $a \in \mathcal{H}$ and $\lambda > 0$.
For $g(x) := d^2_{\mathcal{H}}(a, x)$ and $G(x_0, x_1) := d_{\mathcal{H}}(x_0, x_1)$
the reflections $\mathcal{R}_{\lambda g}$ und $\mathcal{R}_{\lambda G}$ are nonexpansive.

Convergence of PDRA for the ROF model

Are the reflections of the ROF model non-expansive?

Theorem: Nonexpansiveness of $\mathcal{R}_{\lambda\Phi}$ **[RB, Persch, Steidl, 2016]**
Let $\mathcal{H}$ be a symmetric Hadamard manifold, $a \in \mathcal{H}$ and $\lambda > 0$.
For $g(x) := d^2_{\mathcal{H}}(a, x)$ and $G(x_0, x_1) := d_{\mathcal{H}}(x_0, x_1)$
the reflections $\mathcal{R}_{\lambda g}$ und $\mathcal{R}_{\lambda G}$ are nonexpansive.

Theorem: Nonexpansiveness of $\mathcal{R}_{\iota_{\mathrm{D}}}$ **[Fernández-León, Nicolae, 2013]**
Let $\mathcal{H}$ be a symmetric Hadamard manifold with constant sectional curvature and $\mathcal{C}$ be a nonempty, convex subset of $\mathcal{H}$.
Then $\mathcal{R}_{\iota_{\mathcal{C}}}$ is nonexpansive.

Convergence of PDRA for the ROF model

Are the reflections of the ROF model non-expansive?

Theorem: Nonexpansiveness of $\mathcal{R}_{\lambda\Phi}$ **[RB, Persch, Steidl, 2016]**
Let $\mathcal{H}$ be a symmetric Hadamard manifold, $a \in \mathcal{H}$ and $\lambda > 0$.
For $g(x) := d^2_{\mathcal{H}}(a, x)$ and $G(x_0, x_1) := d_{\mathcal{H}}(x_0, x_1)$
the reflections $\mathcal{R}_{\lambda g}$ und $\mathcal{R}_{\lambda G}$ are nonexpansive.

Theorem: Nonexpansiveness of $\mathcal{R}_{\iota_D}$ **[Fernández-León, Nicolae, 2013]**
Let $\mathcal{H}$ be a symmetric Hadamard manifold with constant sectional curvature and $\mathcal{C}$ be a nonempty, convex subset of $\mathcal{H}$.
Then $\mathcal{R}_{\iota_{\mathcal{C}}}$ is nonexpansive.

$\Rightarrow$ Convergence of PDRA

Numerical Examples

The Manifold-valued Image Restoration Toolbox

- inspired by Manopt[1]; focus on image processing
- implemented in Matlab & C++ (`mex`); Julia in preparation
- easy access to manifold-valued image processing
  - Documentation http://ronnybergmann.net/mvirt/
  - Code github.com/kellertuer/mvirt/
- manifolds; object with `exp`,`log`,`dist`,`parallelTransport`
  - symmetric positive definite $d \times d$ matrices $\mathcal{P}(d)$
  - special orthogonal group $\mathrm{SO}(3)$
  - spheres $\mathbb{S}^n$
  - hyperbolic spaces $\mathcal{H}^n$
  - ...
- algorithms implemented on the abstract `manifold` object
- plot functions and exports to TikZ/Asymptote.

[1] manopt.org - Optimization on Manifolds in Matlab

An Algorithm to compare to: CPPA

Decomposing $\mathcal{E} = \sum_{i=1}^{c} \varphi_i$ wo obtain the

Cyclic Proximal Point Algorithmus (CPPA) [Bertsekas, 2011; Bačák, 2014]

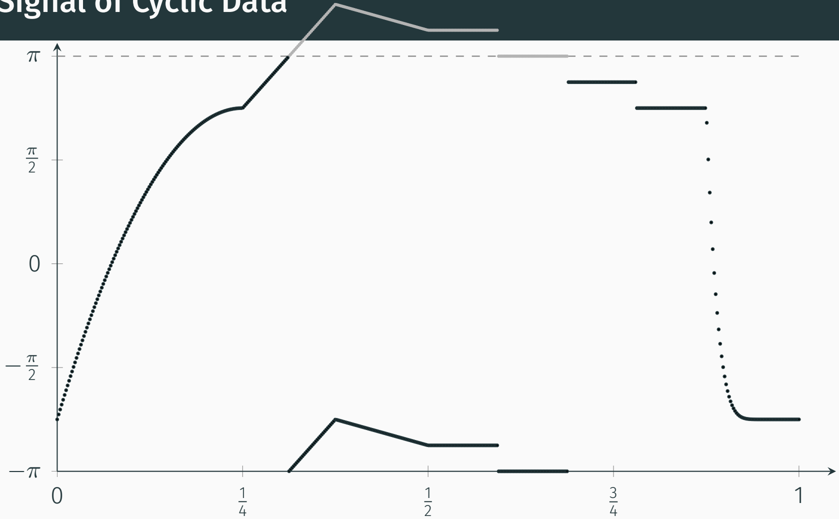
defined for a starting point $x^{(0)} \in \mathcal{M}$ by

$$x^{(k+\frac{i+1}{c})} = \operatorname{prox}_{\lambda_k \varphi_i}(x^{(k+\frac{i}{c})}), \quad i = 0, \ldots, c-1,\ k \geq 0.$$

Convergence of the ROF model

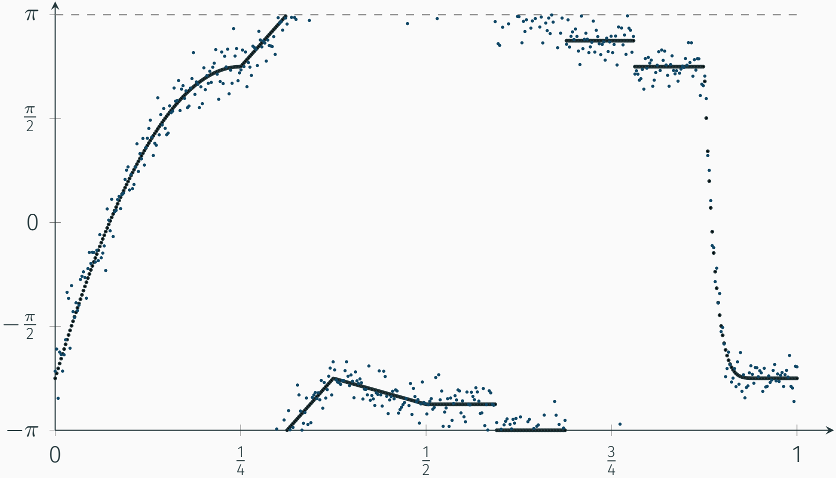
- in Euclidean space $\mathcal{M} = \mathbb{R}^n$ if $\{\lambda_k\}_{k\in\mathbb{N}} \in \ell_2(\mathbb{Z})\backslash\ell_1(\mathbb{Z})$
- On Hadamard manifolds, if φ_i add. Lipschitz [Bačák, 2013; Weinmann, Storath, Demaret, 2014]
- with locality restrictions also on $\mathbb{S}^1$, $(\mathbb{S}^1)^m\mathbb{R}^n$ [RB, Laus, Weinmann, Steidl, 2014; RB, Weinmann, 2016]
- can be extended to second order differences [RB, Laus, Weinmann, Steidl, 2014; RB, Weinmann, 2016; Bačák, RB, Weinmann, Steidl, 2016]

A Signal of Cyclic Data



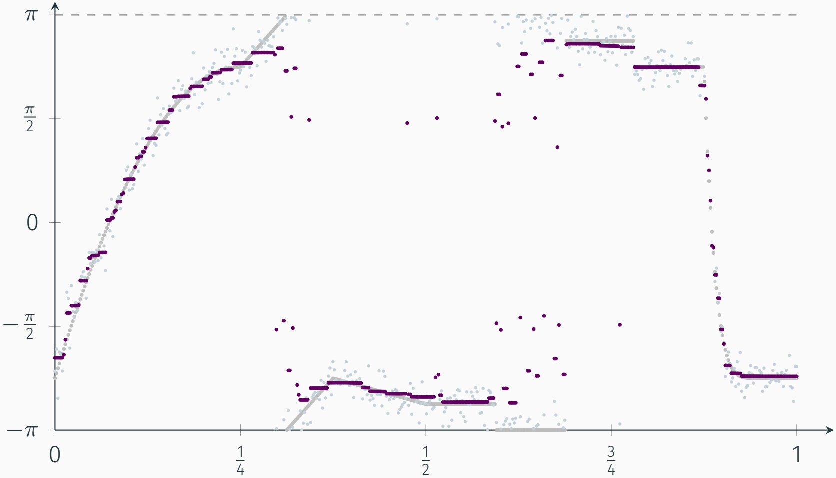
- A function $f\colon [0,1] \to \mathbb{S}^1$ is sampled $\Rightarrow f_o = (f_{o,i})_{i=1}^{500}$
- Data f stems from the gray plot via modulo
- Jumps $> \pi$ at $\frac{5}{16}$ and $\frac{11}{16}$ just from choice of representation

A Signal of Cyclic Data



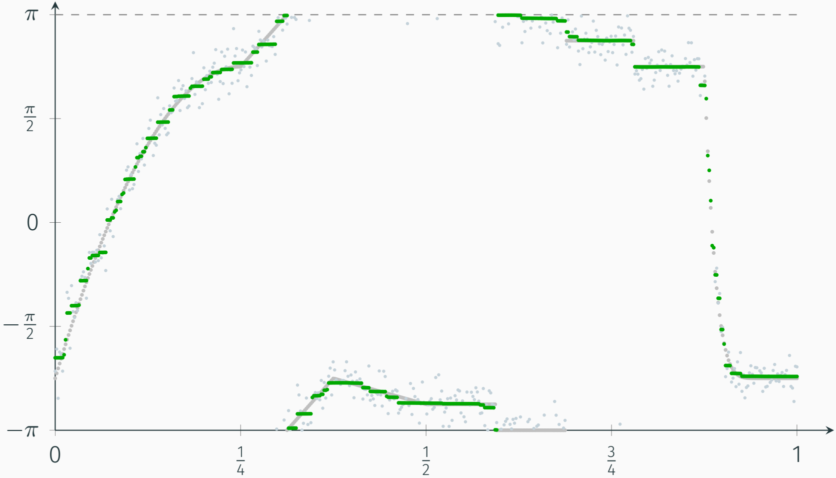
- A function $f\colon [0,1] \to \mathbb{S}^1$ is sampled $\Rightarrow f_o = (f_{o,i})_{i=1}^{500}$
- Noise: wrapped Gaussian, $\sigma = 0.2$
- noisy $f_n = (f_o + \eta)_{2\pi}$

A Signal of Cyclic Data



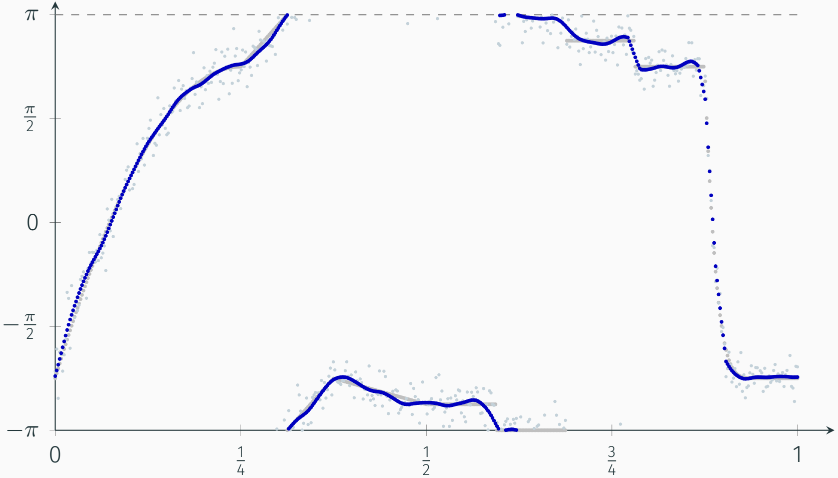
- Comparison of f_o & f_n width f_R
- Denoised with CPPA and realvalued TV_1, ($\alpha = \frac{3}{4}$, $\beta = 0$)
- Artefacts at the “jumps that are none“ from representation

A Signal of Cyclic Data



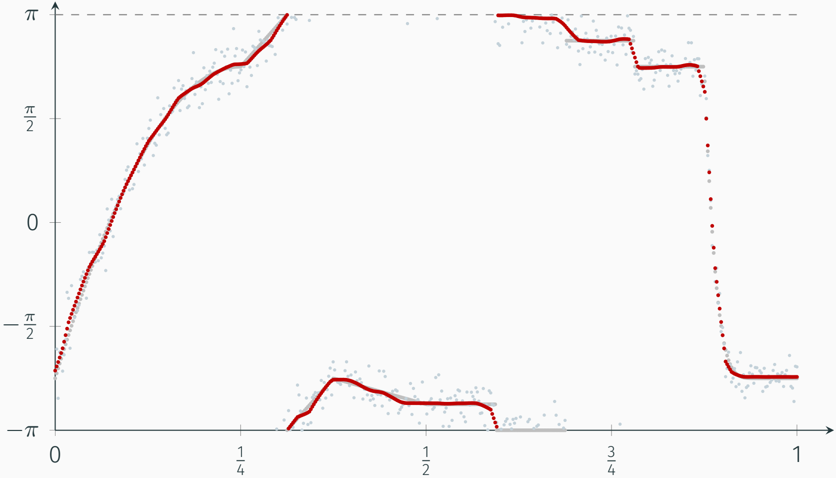
- Comparison of f_o & f_n width f_1
- Denoised with CPPA and TV_1 $(\alpha = \frac{3}{4}, \beta = 0)$
- but: stair caising

A Signal of Cyclic Data



- Comparison of f_o & f_n width f_2
- Denoised with CPPA and TV_2 ($\alpha = 0$, $\beta = \frac{3}{2}$)
- but: problems in constant areas

A Signal of Cyclic Data

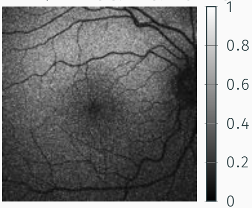


- Comparison of f_o & f_n width f_3
- Denoised with CPPA and TV_1 & TV_2 ($\alpha = \frac{1}{4}$, $\beta = \frac{3}{4}$)
- combined: smallest mean squarred error.

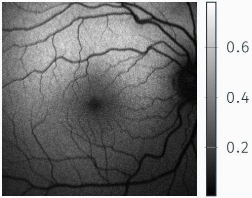
An Image of Gaussian Distributions

Given: 20 Retina images [Angulo, Velasco-Forero, 2014]

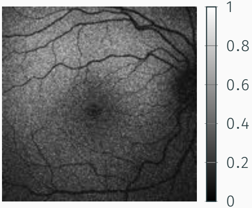
$\Rightarrow$ mean and variance per pixel, $\mathcal{M} = (\mathbb{H}^2)^{384^2}$



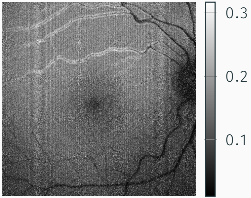
First Image



Measured mean.



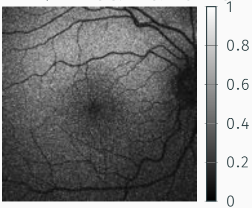
Last image.



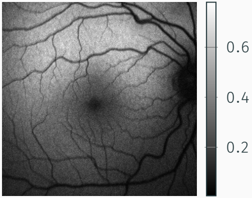
measured variance.

Given: 20 Retina images [Angulo, Velasco-Forero, 2014]

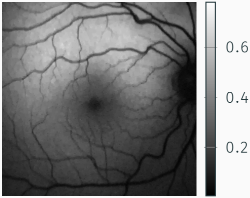
$\Rightarrow$ mean and variance per pixel, $\mathcal{M} = (\mathbb{H}^2)^{384^2}$



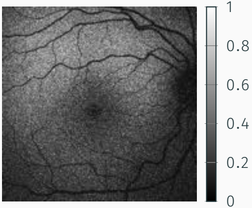
First Image



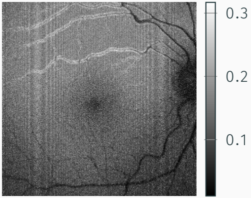
Measured mean.



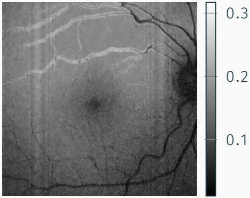
Denoised mean.



Last image.



measured variance.



denoised variance.

Comparison of CPPA & PDRA

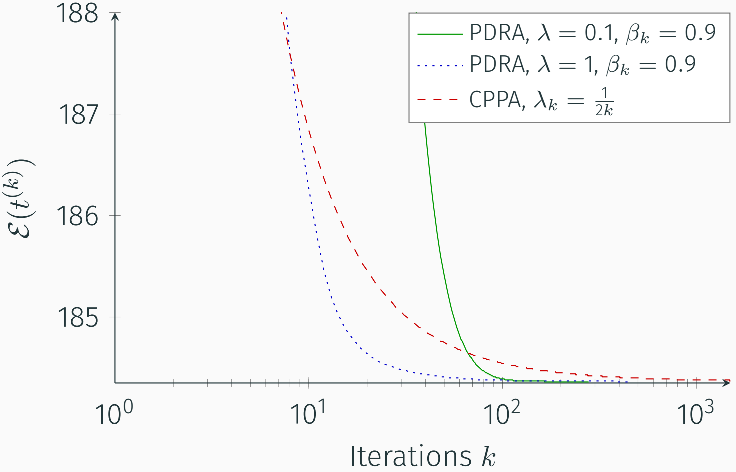
Stopping criterion:

- $\epsilon^{(k)} := \max_{(i,j)\in\mathcal{G}}\left\{d\left(t_{i,j}^{(k)}, t_{i,j}^{(k-1)}\right)\right\} < 10^{-6}$
- $k > 1500$

λ	CPPA	PDRA		
	in sec.	$\beta_k = 0.5$	$\beta_k = 0.9$	$\beta_k = 0.95$
0.05	56.85	129.26	65.21	59.84
0.1	56.54	59.21	34.32	36.67
0.5	65.17	57.41	42.06	46.07
1	57.14	93.75	63.58	58.66

Table 1: Runtimes (seconds) of the algorithms.

Comparison of CPPA & PDRA



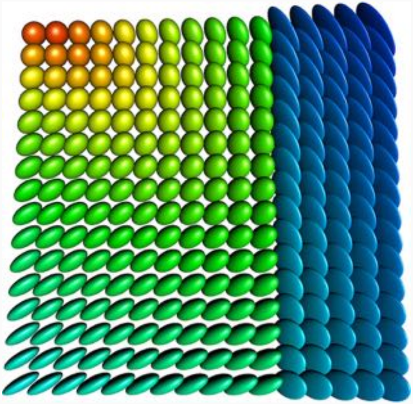
Comparison of CPPA & PDRA

λ	CPPA	PDRA		
	184.3643+...	$\beta_k = 0.5$	$\beta_k = 0.9$	$\beta_k = 0.95$
0.05	44.80	1.021×10^{-5}	1.180×10^{-5}	1.627×10^{-5}
0.1	10.65	2.514×10^{-5}	2.969×10^{-5}	3.429×10^{-5}
0.5	1.055×10^{-2}	5.082×10^{-4}	2.785×10^{-4}	2.256×10^{-4}
1	1.953×10^{-2}	8.189×10^{-4}	5.027×10^{-4}	4.992×10^{-4}

Table 1: Functional values $\mathcal{E}(t^{(k_{\text{last}})})$

Inpainting of a $\mathcal{P}(3)$-valued Image

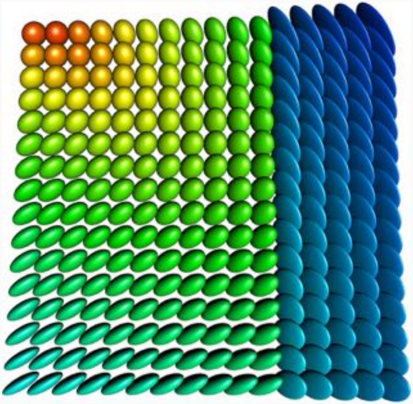
Visualization of sym. pos. def. 3×3 matrices as ellipsods



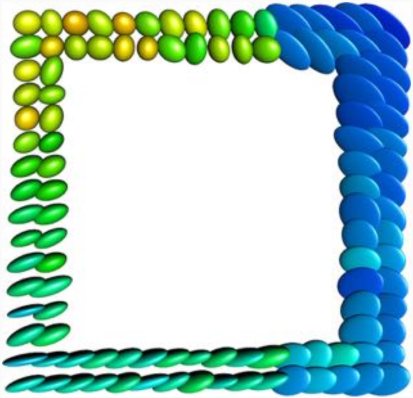
Original.

Inpainting of a $\mathcal{P}(3)$-valued Image

Visualization of sym. pos. def. 3×3 matrices as ellipsods



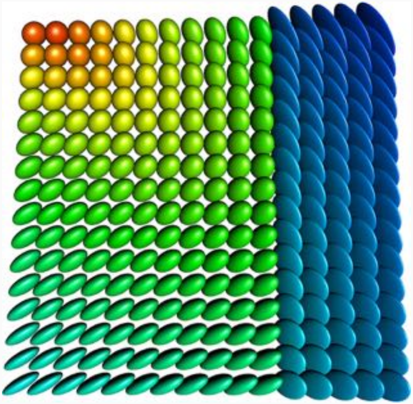
Original.



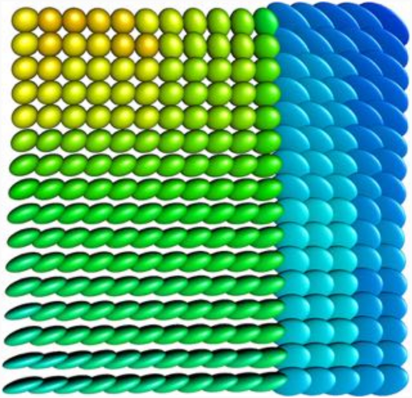
lossy, noisy data

Inpainting of a $\mathcal{P}(3)$-valued Image

Visualization of sym. pos. def. 3×3 matrices as ellipsods



Original.



Reconstruction, $\alpha = 0.1$

Conclusion

Conclusion

- Variational methods can be generalized to manifold valued data
- Douglas-Rachford Algorithm for efficient minimization (eqiv. to ADMM on $\mathbb{R}^n$)

Numerical examples implemented in MVIRT

http://ronnybergmann.net/mvirt/

which also serves as an easy starting point for manifold-valued image processing. A port to Julia is in progress.

RB, J. Persch, and G. Steidl. "A Parallel Douglas–Rachford Algorithm for Minimizing ROF-like Functionals on Images with Values in Symmetric Hadamard Manifolds". In: *SIAM J. Imag. Sci.* 9.3 (2016), pp. 901–937. arXiv: 1512.02814.

ronnybergmann.net/talks/2018-ISMP18-DouglasRachford.pdf