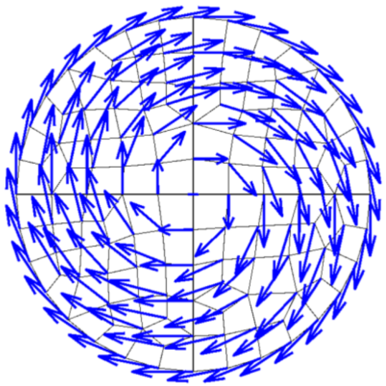
University of Stuttgart
Institute for Structural Mechanics

Optimization on manifolds: methods and applications - an engineering perspective

Alexander Müller

KAAS seminar,
Karlstad University

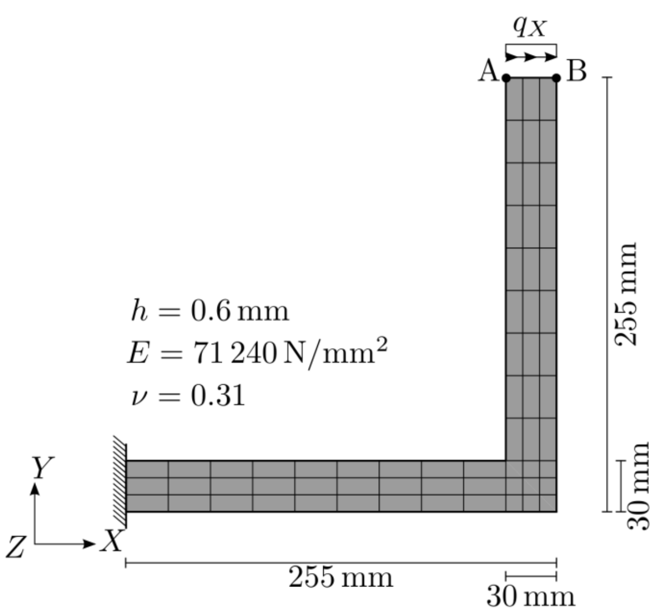
26.01.2022

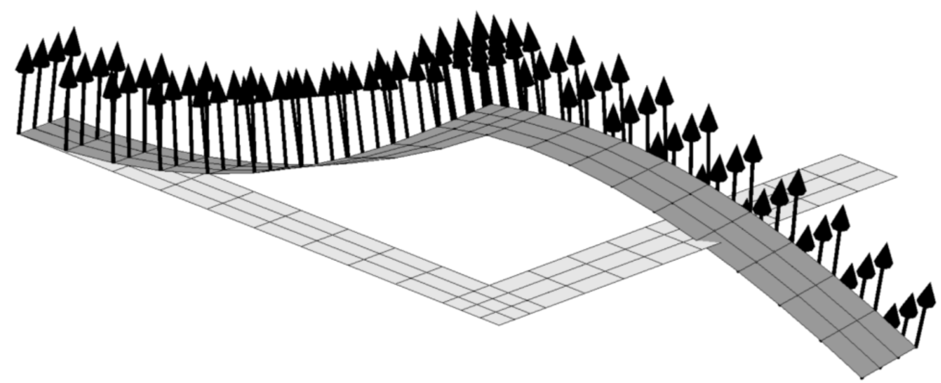


Baustatik und Baudynamik

Bio

- Studied civil engineering in Stuttgart
- PhD student at the institute for structural mechanics in Stuttgart (since 2018)
- Supervised by Prof. Bischoff (Head of the institute)





Motivation

Toy problem, Problem statement

Cost function:

$$E : \mathbb{R}^n \to \mathbb{R}, \quad E(\mathbf{x}) = (\mathbf{x} + \mathbf{b})^T \mathbf{A} (\mathbf{x} + \mathbf{b})$$

$$\mathbf{b} \in \mathbb{R}^n$$

$$\mathbf{A} \in \mathbb{R}^{n \times n} \text{ and } \mathbf{A}^T = \mathbf{A}$$

Constraints:

$$\mathbf{x} \in \mathbb{R}^n \quad \text{s.t.} \quad \mathbf{x}^T \mathbf{x} = 1$$

Solution: $\mathbf{x}^* = \arg\{\min_{\mathbf{x} \in \mathbb{R}^n} E(\mathbf{x}), \text{ where } \mathbf{x} \in \mathbb{R}^n \text{ s.t. } \mathbf{x}^T \mathbf{x} = 1\}$

„Constrained optimization on an unconstrained space“

Motivation

Toy problem: Lagrange multiplier approach

Cost function:

$$\hat{E} : (\mathbb{R}^n, \mathbb{R}) \rightarrow \mathbb{R}, \quad \hat{E}(\mathbf{x},\lambda) = E(\mathbf{x}) + \lambda(\mathbf{x}^T\mathbf{x} - 1)$$

Solution: $\{\mathbf{x}^*,\lambda^*\} = \arg\{\max_{\lambda\in\mathbb{R}} \min_{\mathbf{x}\in\mathbb{R}^n} \hat{E}(\mathbf{x},\lambda)\}$

- Dimension increase of the search space
- Minimization problem → Saddle-point problem
- Allows general constraints
- Simple linearization

Baustatik und Baudynamik

Toy problem: Penalty

Cost function:

$$\breve{E} : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \breve{E}(\mathbf{x}) = E(\mathbf{x}) + c(\mathbf{x}^T\mathbf{x} - 1)^2$$

Solution: $\mathbf{x}^* = \arg\{\min\limits_{\mathbf{x}\in\mathbb{R}^n} \breve{E}(\mathbf{x})\}$

- Dependence on artificial parameter
- Simple linearization
- General constraints
- Problem structure retained

Motivation

Baustatik und Baudynamik

Toy problem, Parametrization of the manifold

Cost function:

$$\bar{E} : \mathbb{R}^{n-1} \to \mathbb{R}, \quad \bar{E}(\boldsymbol{\varphi}) = E(\mathbf{x}(\boldsymbol{\varphi}))$$

Solution: $$\boldsymbol{\varphi}^* = \arg\{ \min_{\boldsymbol{\varphi} \in \mathbb{R}^{n-1}} \bar{E}(\boldsymbol{\varphi}) \}$$

First order conditions:

…lengthy…

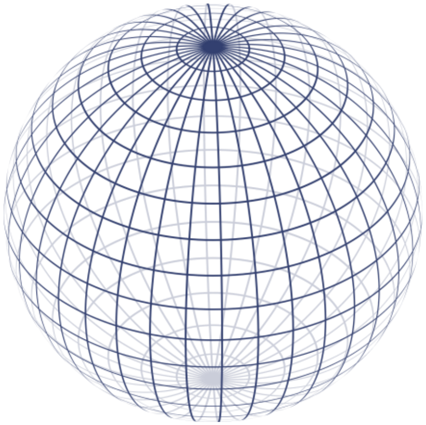
- Involved linearization
- Possibility of introducing singularities
- Problem structure retained
- Solves the problem on the „correct“ design space

Spherical coordinates [edit]

We may define a coordinate system in an n-dimensional Euclidean space which is ar
of a radial coordinate r, and $n-1$ angular coordinates $\varphi_1, \varphi_2, \dots \varphi_{n-1}$, where the an
$[0,360)$ degrees). If x_i are the Cartesian coordinates, then we may compute $x_1, \dots x_n$

$$\begin{aligned}
x_1 &= r\cos(\varphi_1) \\
x_2 &= r\sin(\varphi_1)\cos(\varphi_2) \\
x_3 &= r\sin(\varphi_1)\sin(\varphi_2)\cos(\varphi_3) \\
&\vdots \\
x_{n-1} &= r\sin(\varphi_1)\cdots\sin(\varphi_{n-2})\cos(\varphi_{n-1}) \\
x_n &= r\sin(\varphi_1)\cdots\sin(\varphi_{n-2})\sin(\varphi_{n-1}).
\end{aligned}$$

WIKIPEDIA: N-SPHERE



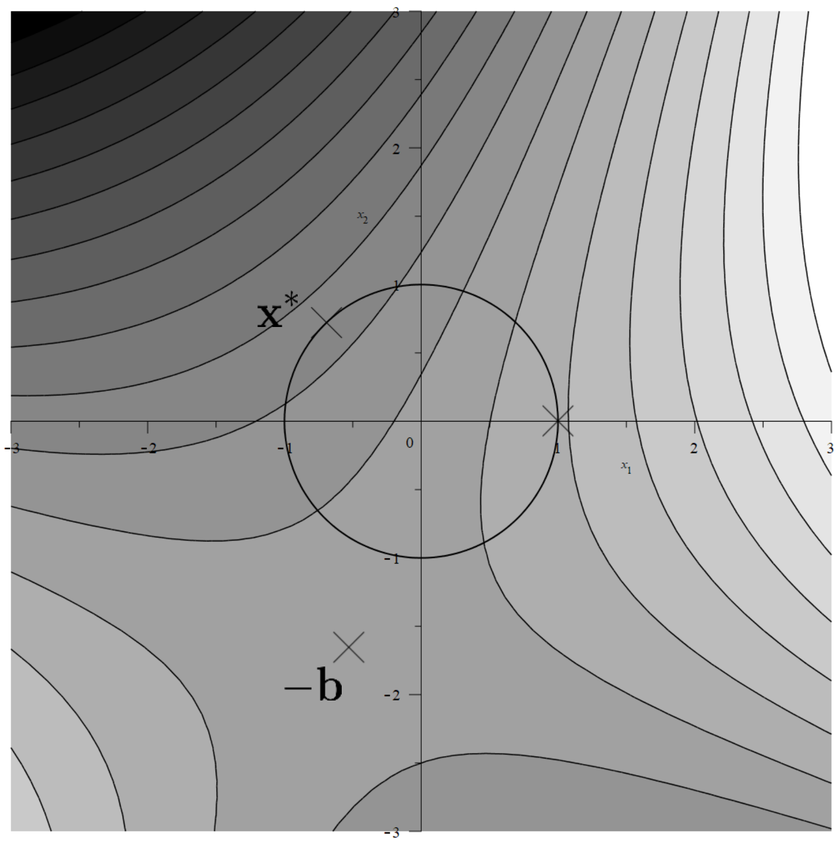
WIKIPEDIA: HAIRY_BALL_THEOREM

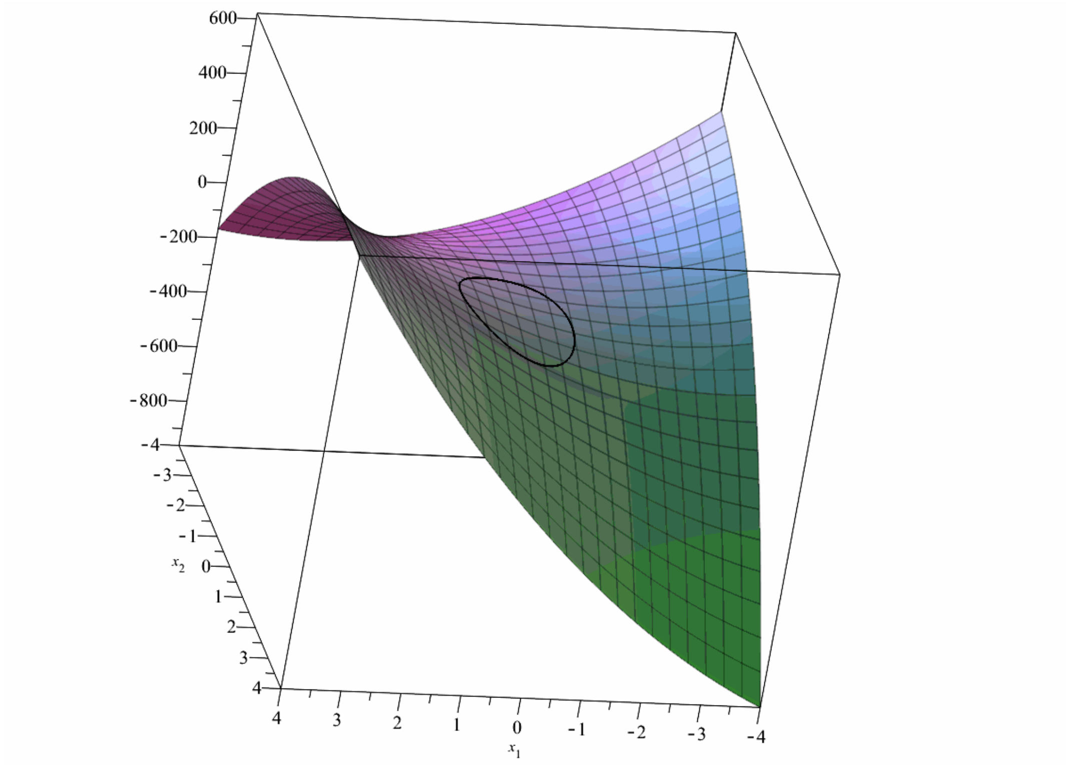
Motivation

Baustatik und Baudynamik

Toy problem $\mathbf{x} \in \mathbb{R}^2 \quad E : \mathbb{R}^n \to \mathbb{R}, \quad E(\mathbf{x}) = (\mathbf{x} + \mathbf{b})^T \mathbf{A} (\mathbf{x} + \mathbf{b})$

$$\mathbf{A} = \begin{bmatrix} 12.6 & 16.23 \\ 16.23 & -14.92 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0.53 \\ 1.65 \end{bmatrix}$$





Motivation

Baustatik und Baudynamik

Toy problem: Newton's method

Algorithm 1: Classic Newton in vector spaces

Goal : Find stationary point of $E(\mathbf{x}) \in \mathbb{R}^n$, i.e. find $\mathbf{x}$ such that $\operatorname{grad} E(\mathbf{x}) = \mathbf{0}$.

input : Initial iterate $\mathbf{x}_0$

output: Converged solution: $\mathbf{x}^*$

while $||\operatorname{grad} E(\mathbf{x}_k)|| > tol$ **do**

$\quad \operatorname{Hess} E(\mathbf{x}_k)\Delta\mathbf{x}_k = -\operatorname{grad} E(\mathbf{x}_k)$

$\quad \mathbf{x}^{k+1} = \mathbf{x}^k + \Delta\mathbf{x}_k$

$\quad k \leftarrow k + 1$

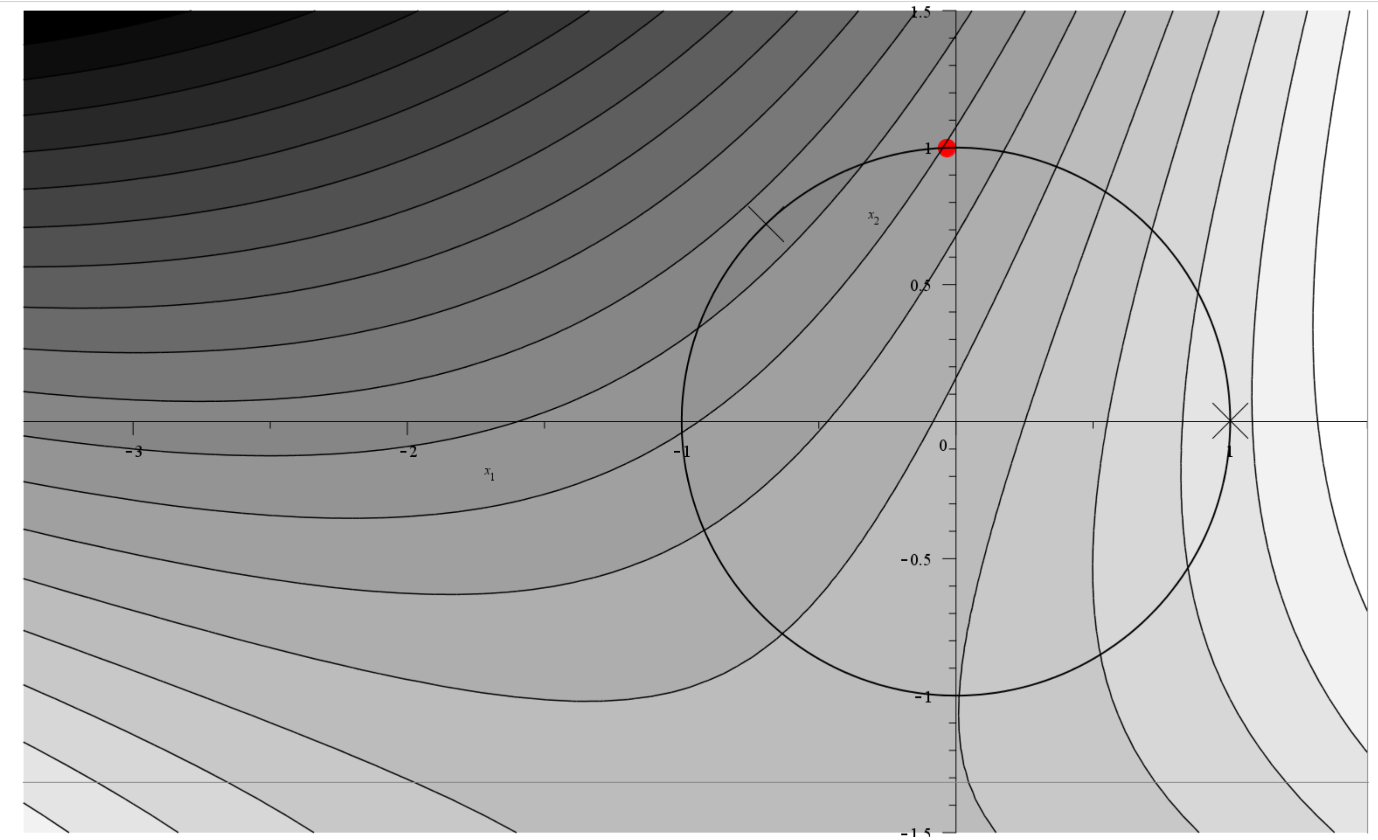
end

$$\operatorname{Hess} E(\mathbf{x}) = \frac{\partial^2 E(\mathbf{x})}{\partial \mathbf{x}^2}$$

$$\operatorname{grad} E(\mathbf{x}) = \frac{\partial E(\mathbf{x})}{\partial \mathbf{x}}$$

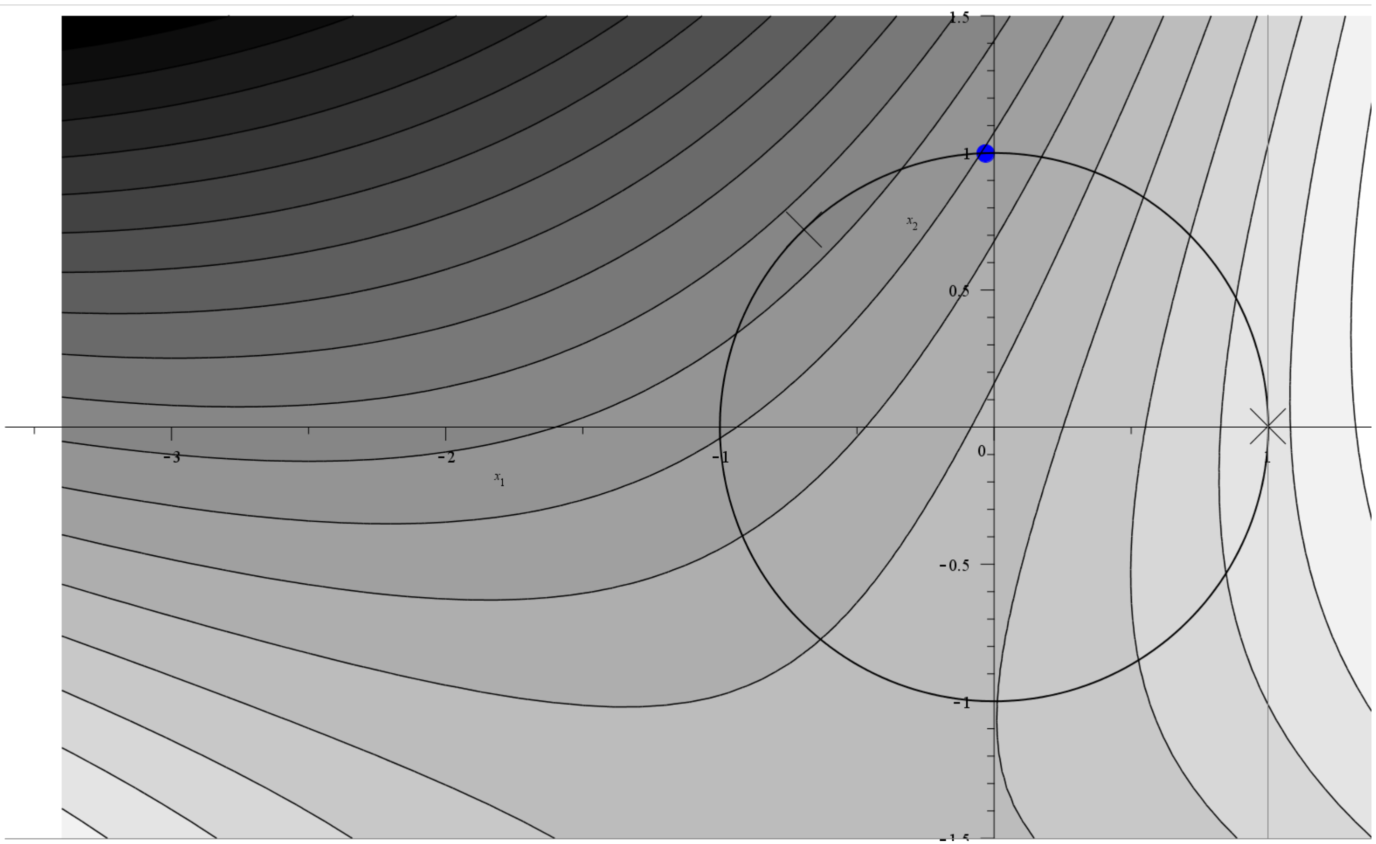
Motivation

Toy problem: iterations of the Lagrange multiplier approach



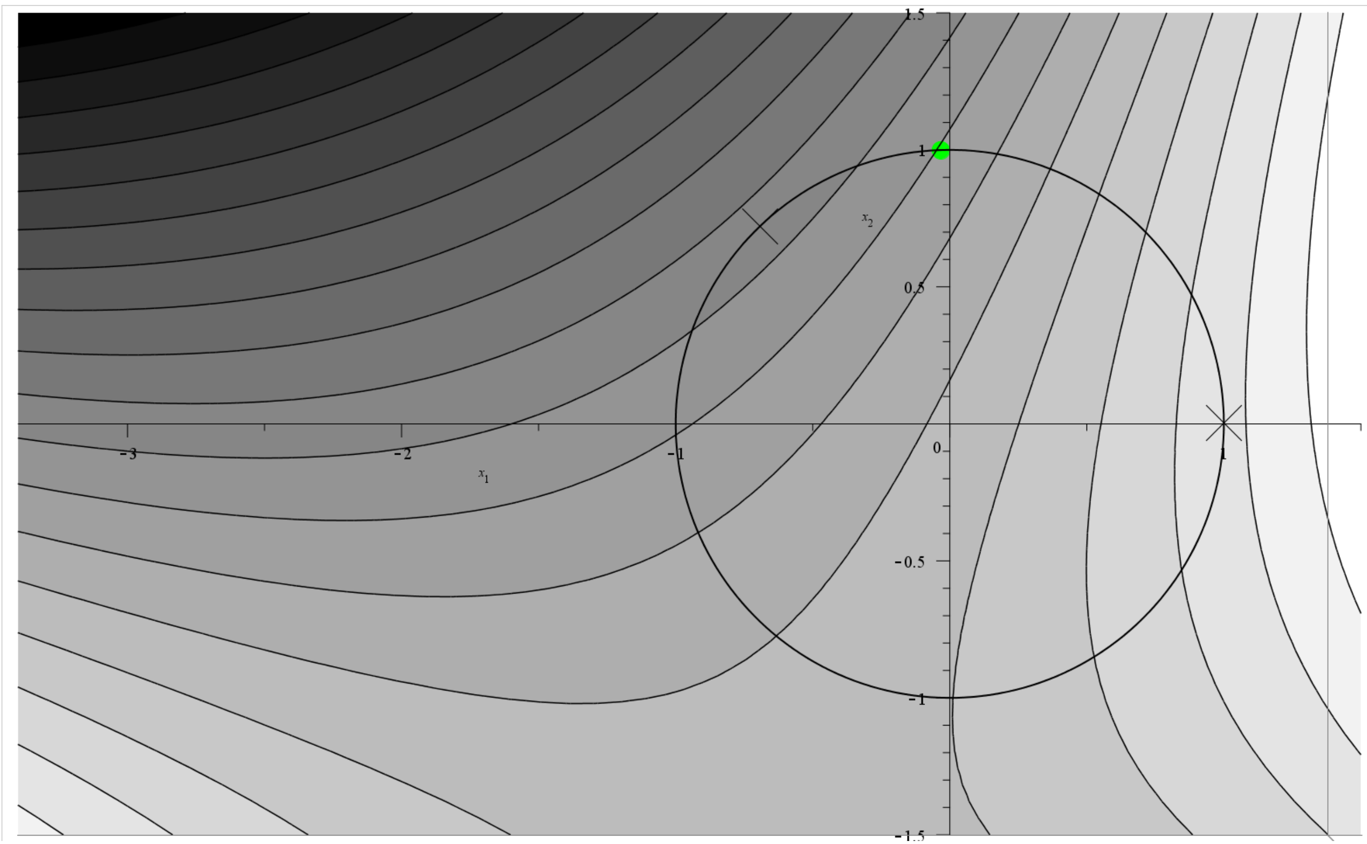
Motivation

Toy problem: iterations of the penalty approach



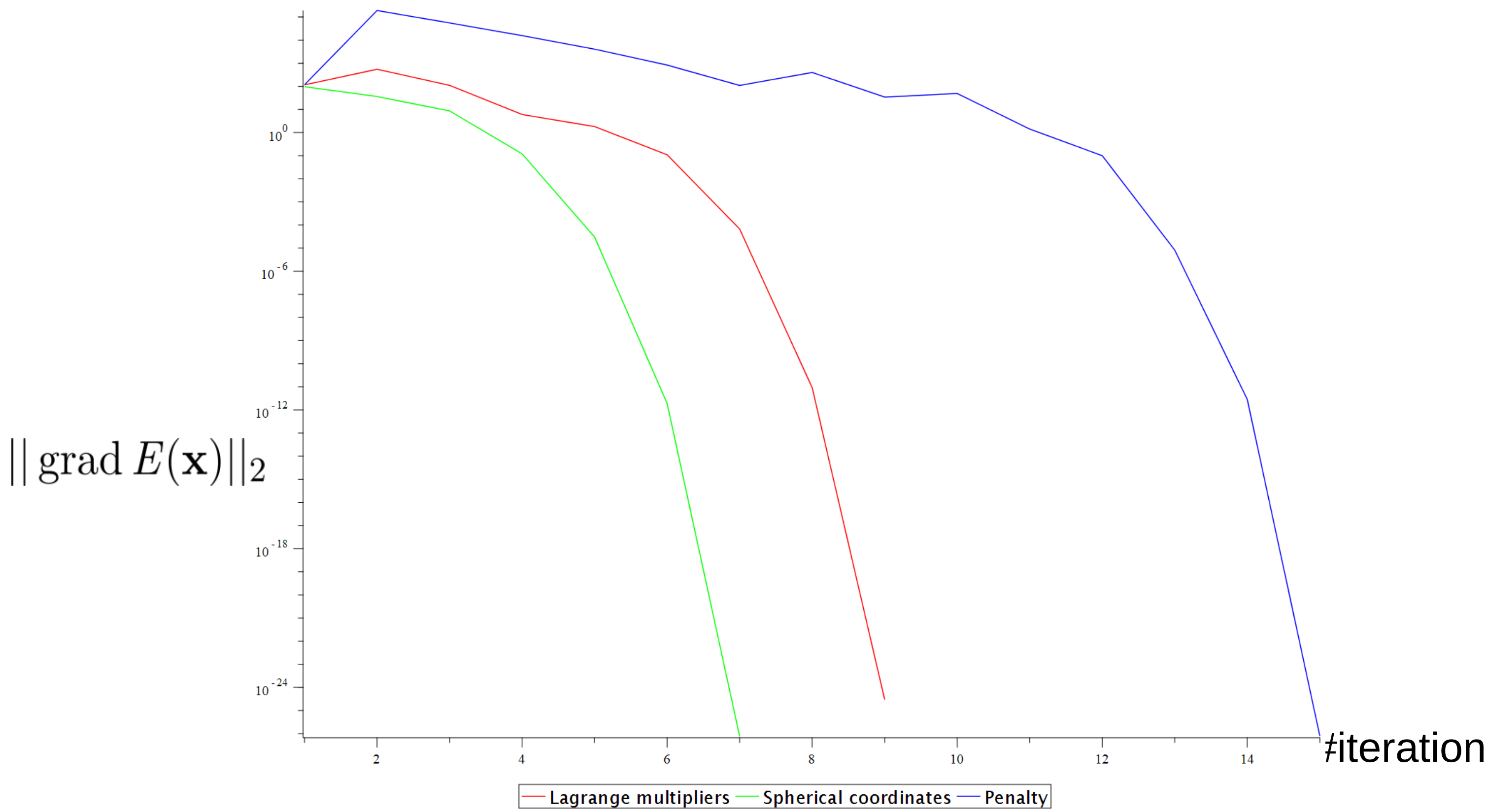
Motivation

Toy problem: iterations of the coordinate approach



Motivation

Toy problem: number of iterations vs. gradient norm comparison



Motivation

Toy problem: Presented method

	LAM	Penalty	Coordinates	This talk
Linearization	☺	☺	☹	☺
Singularities	☺	☺	☹	☺
Search space	☹	☹	☺	☺
Minimization	☹	☺	☺	☺
Iterations	☹	☹	☺	☺

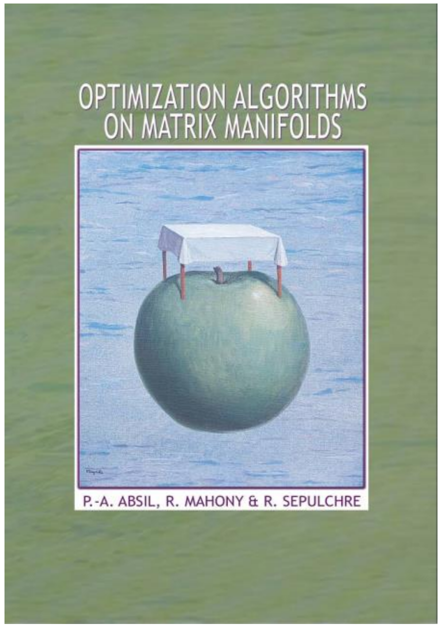
Cost function: $\tilde{E}:\mathcal{S}^{n-1}\rightarrow\mathbb{R},\quad \tilde{E}(\mathbf{x})=(\mathbf{x}+\mathbf{b})^T\mathbf{A}(\mathbf{x}+\mathbf{b})\qquad \mathbf{x}\in\mathcal{S}^{n-1}$

Solution: $\mathbf{x}^*=\arg\{\min_{\mathbf{x}\in\mathcal{S}^{n-1}}\tilde{E}(\mathbf{x})\}$ *„Unconstrained optimization on a constrained space“*

- Not appropriate for all constraints
- More theory, but nice theory ☺

Optimization on Manifolds

Literature



ABSIL PA, MAHONY R, SEPULCHRE R (2008) OPTIMIZATION ALGORITHMS ON MATRIX MANIFOLDS. PRINCETON UNIVERSITY PRESS, DOI:10.1515/9781400830244

BOUMAL N (2020) AN INTRODUCTION TO OPTIMIZATION ON SMOOTH MANIFOLDS. AVAILABLE ONLINE, LINK

Optimization on manifolds

Update of design variables

Riemannian gradient / Hessian

Riemannian Newton

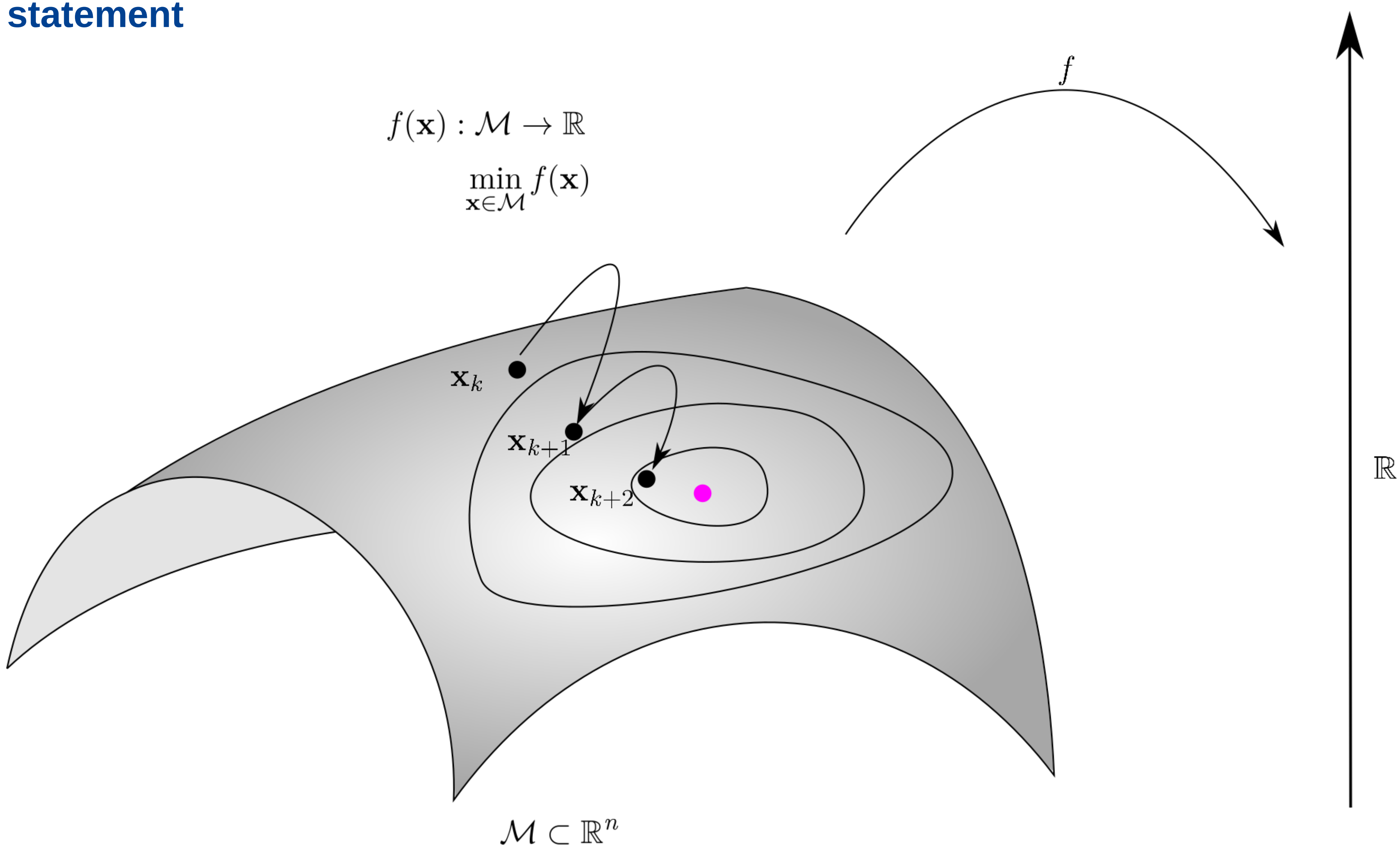
Simulations

Other manifolds

Optimization on Manifolds

Baustatik und Baudynamik

Problem statement



Baustatik und Baudynamik

Steepest descent in vector spaces

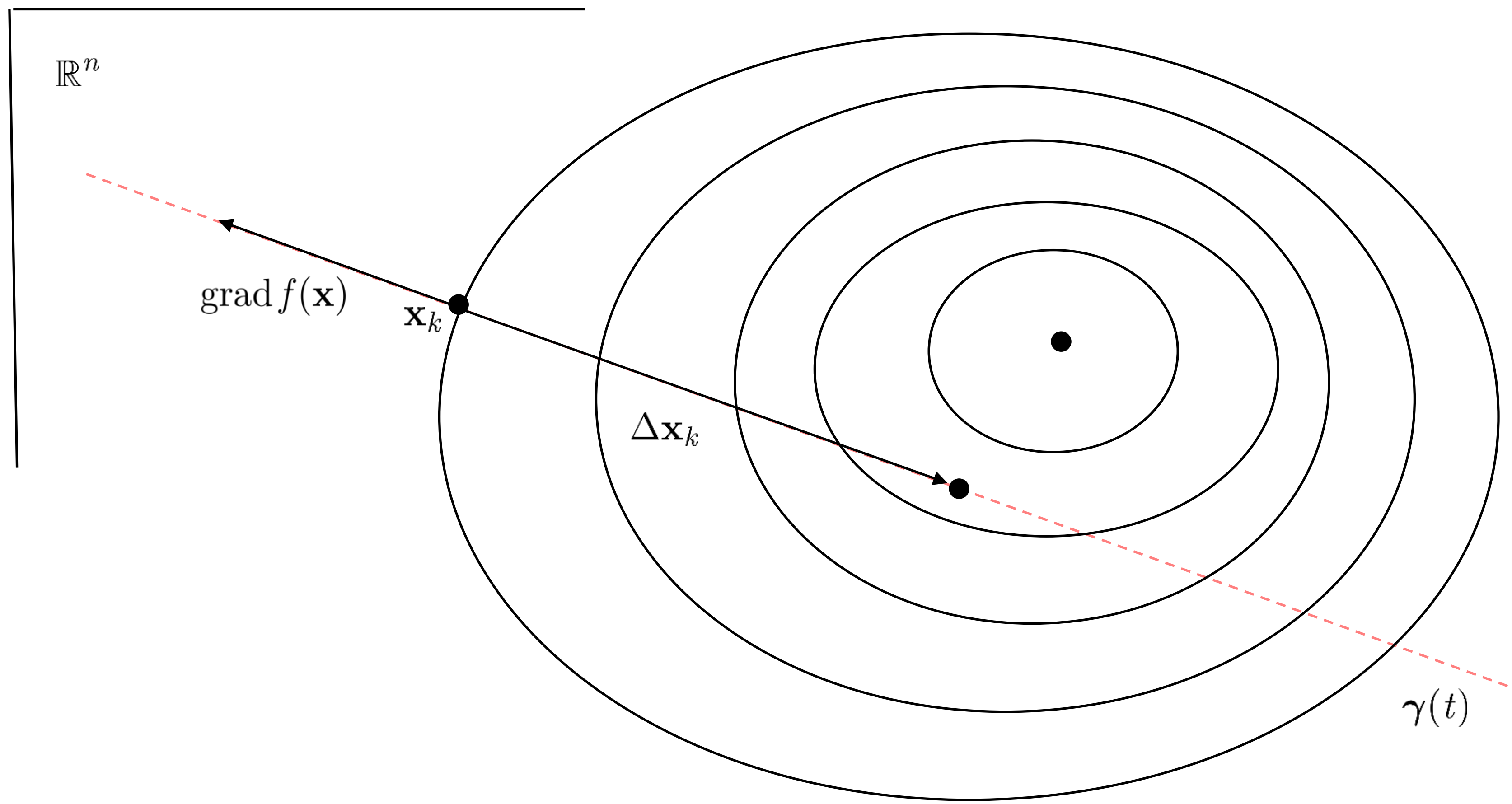
$$f : \mathbb{R}^n \to \mathbb{R} \qquad \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

$$\boldsymbol{\gamma} : \mathbb{R} \to \mathbb{R}^n$$

$$\boldsymbol{\gamma}(t) = \mathbf{x}_k - t \operatorname{grad} f(\mathbf{x})$$

$$\Delta \mathbf{x}_k = -\alpha \operatorname{grad} f(\mathbf{x})$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{x}_k$$



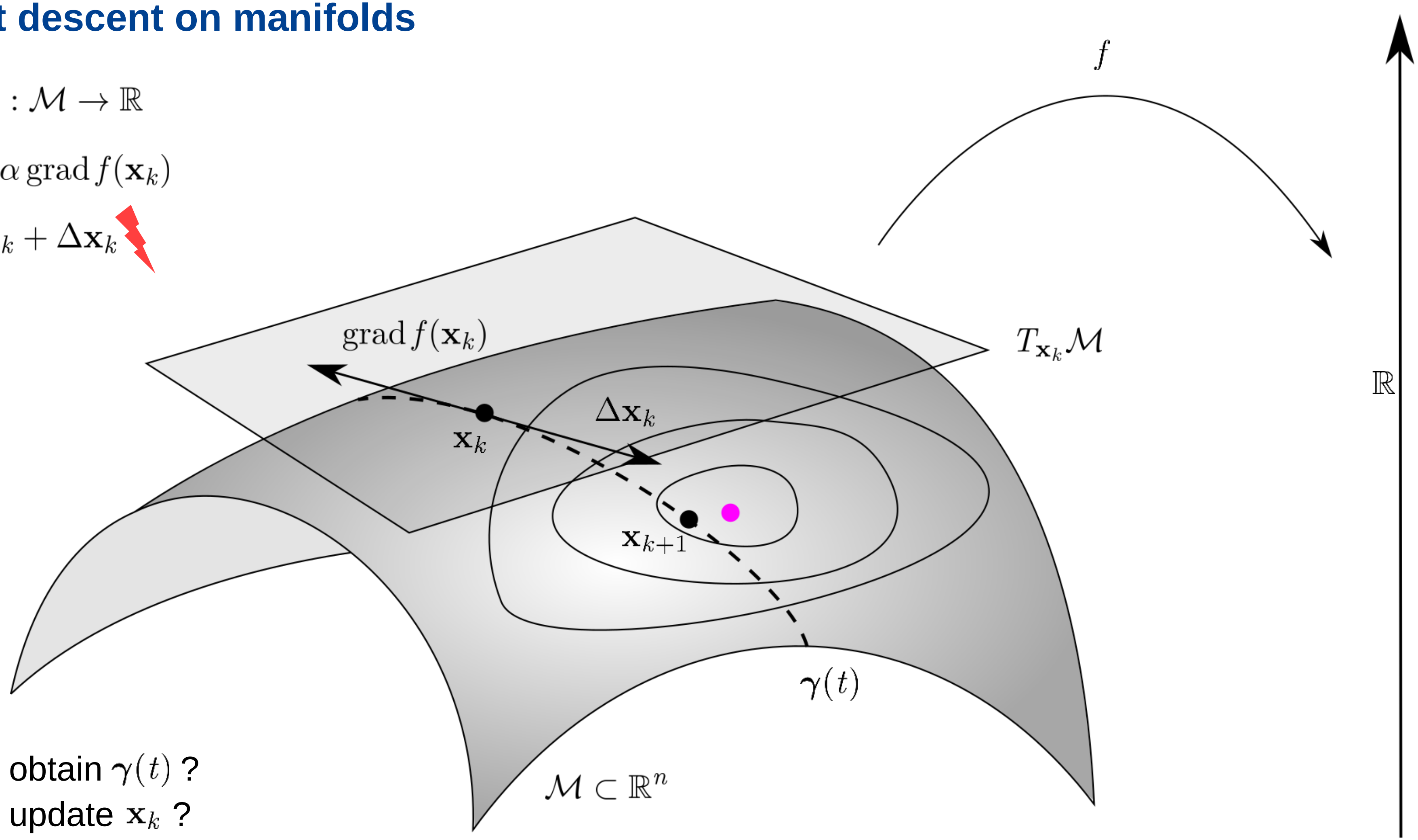
Optimization on Manifolds

Steepest descent on manifolds

$$f(\mathbf{x}) : \mathcal{M} \to \mathbb{R}$$

$$\Delta\mathbf{x}_k = -\alpha \operatorname{grad} f(\mathbf{x}_k)$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k$$



1. How to obtain $\gamma(t)$?
2. How to update $\mathbf{x}_k$?
3. How to obtain $\operatorname{grad} f(\mathbf{x}_k)$?

Update of design variables

Baustatik und Baudynamik

Update of design variables in vector spaces

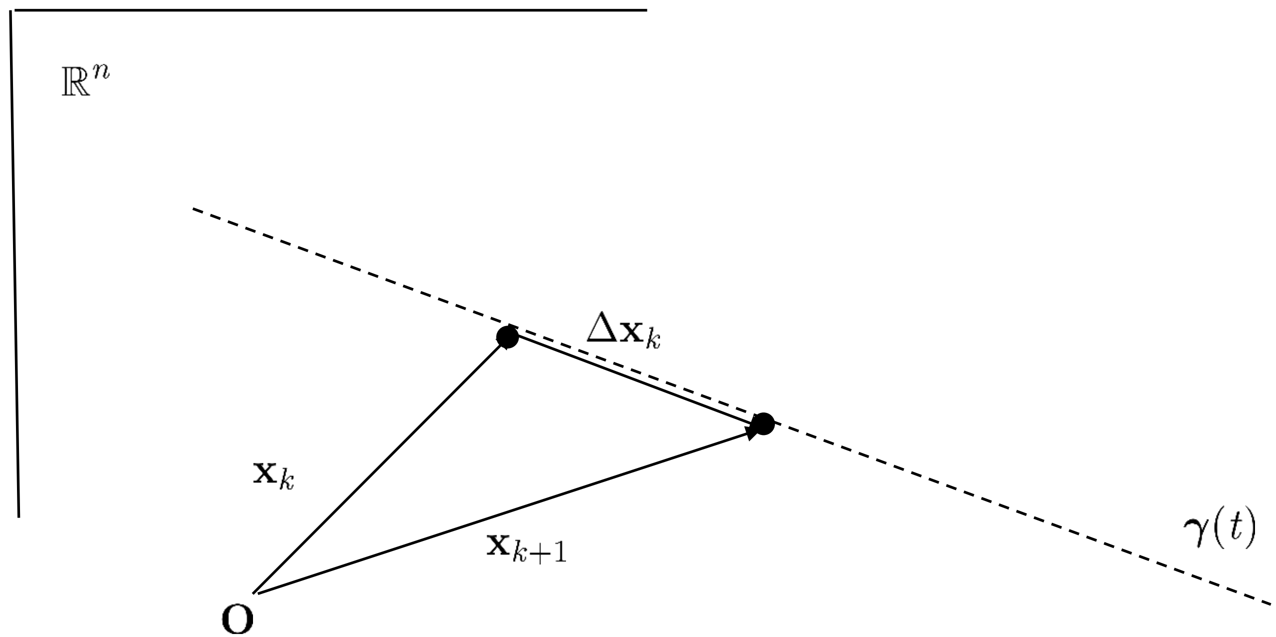
$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k$$

$$\boldsymbol{\gamma} : \mathbb{R} \to \mathbb{R}^n$$

$$\boldsymbol{\gamma}(t) = \mathbf{x}_k + t\Delta\mathbf{x}_k$$

$$\frac{\partial\boldsymbol{\gamma}(t)}{\partial t} = \Delta\mathbf{x}_k$$

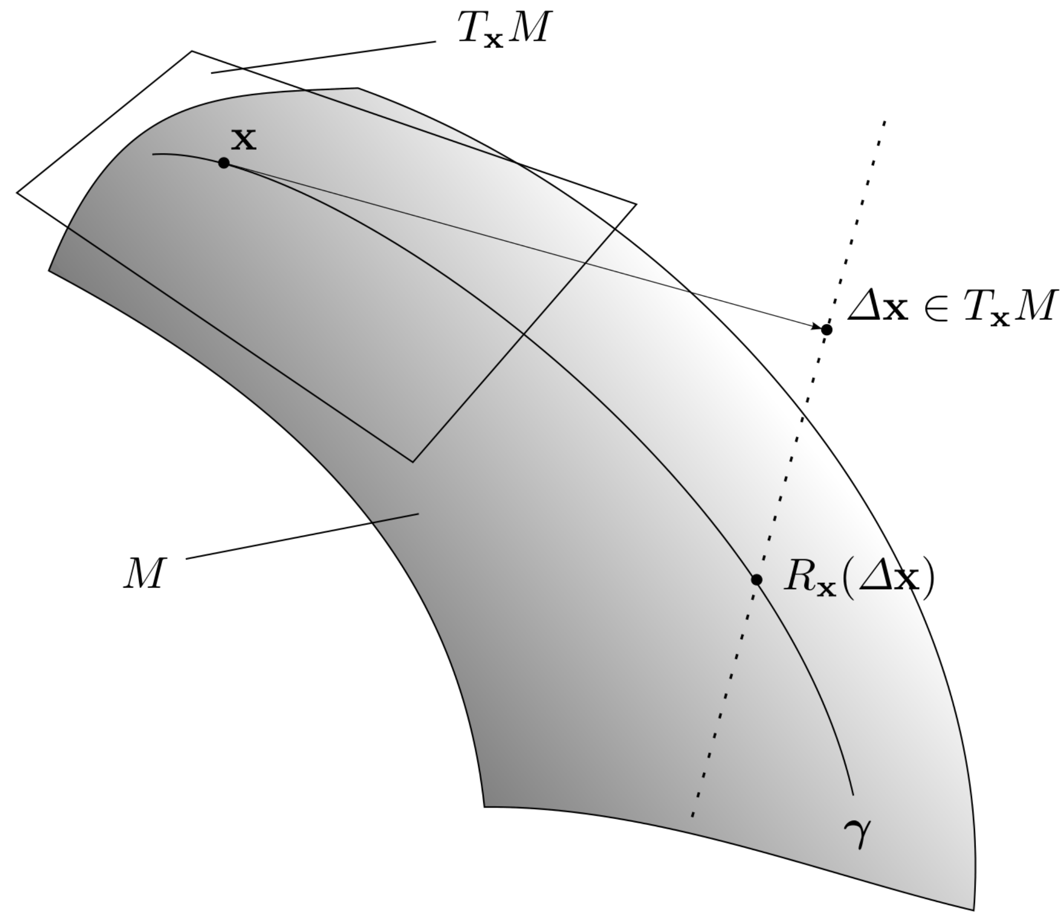
$$\frac{\partial^2\boldsymbol{\gamma}(t)}{\partial t^2} = \mathbf{0}$$



Optimization on Manifolds

Update of design variables

$\mathbf{x}_k + \Delta\mathbf{x}_k \notin \mathcal{M}$



Geodesics generalize the concept of straight lines

$$\gamma : \mathbb{R} \to \mathcal{M}$$

$$\gamma(0) = \mathbf{x}, \quad \dot{\gamma}(t) = \left.\frac{\partial \gamma(t)}{\partial t}\right|_{t=0} = \Delta\mathbf{x}$$

$$\ddot{\gamma}(t) = \nabla_{\dot{\gamma}}\dot{\gamma}(t) = \mathbf{0}$$

The exponential map creates the *unique* geodesic curve starting at $\mathbf{x}_k$ in direction $\Delta\mathbf{x}_k$ with constant speed

$$\gamma(t) = \exp_{\mathbf{x}_k}(t\Delta\mathbf{x}_k) \quad \Rightarrow \quad \mathbf{x}_{k+1} = \exp_{\mathbf{x}_k}(\Delta\mathbf{x}_k)$$

Optimization on Manifolds

Update of design variables

ABSIL PA , "OPTIMIZATION ON MANIFOLDS: METHODS AND APPLICATIONS", LEUVEN, 18 SEP 2009.

Luenberger (1973), Introduction to linear and nonlinear programming.
Luenberger mentions the idea of performing line search along geodesics, "*which we would use if it were computationally feasible (which it definitely is not)*".

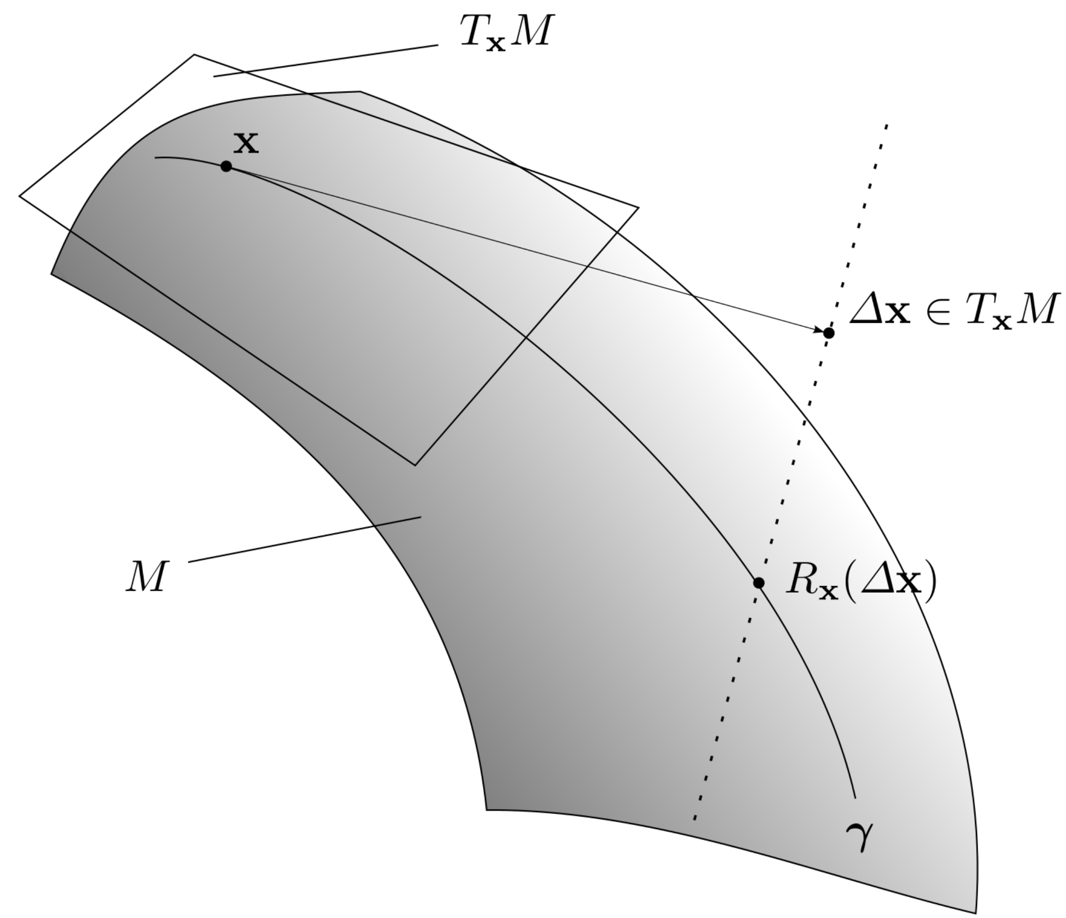
Generalize the concept of the exponential map → Retractions

Optimization on Manifolds

Update of design variables

Retraction

$$\mathbf{x}_k + \Delta\mathbf{x}_k \notin \mathcal{M}$$



Minimal requirements

$$\mathbf{R}_{\mathbf{x}_k} : T_{\mathbf{x}_k}\mathcal{M} \to \mathcal{M}$$

$$\mathbf{R}_{\mathbf{x}_k}(\mathbf{0}) = \mathbf{x}_k$$

$$\left.\frac{\partial \mathbf{R}_{\mathbf{x}_k}(t\Delta\mathbf{x}_k)}{\partial t}\right|_{t=0} = \Delta\mathbf{x}_k, \quad \Delta\mathbf{x}_k \in T_{\mathbf{x}_k}\mathcal{M}$$

$$\xcancel{\ddot{\gamma}(t) = \nabla_{\dot{\gamma}}\dot{\gamma}(t) = \mathbf{0}}$$

Optimization on Manifolds

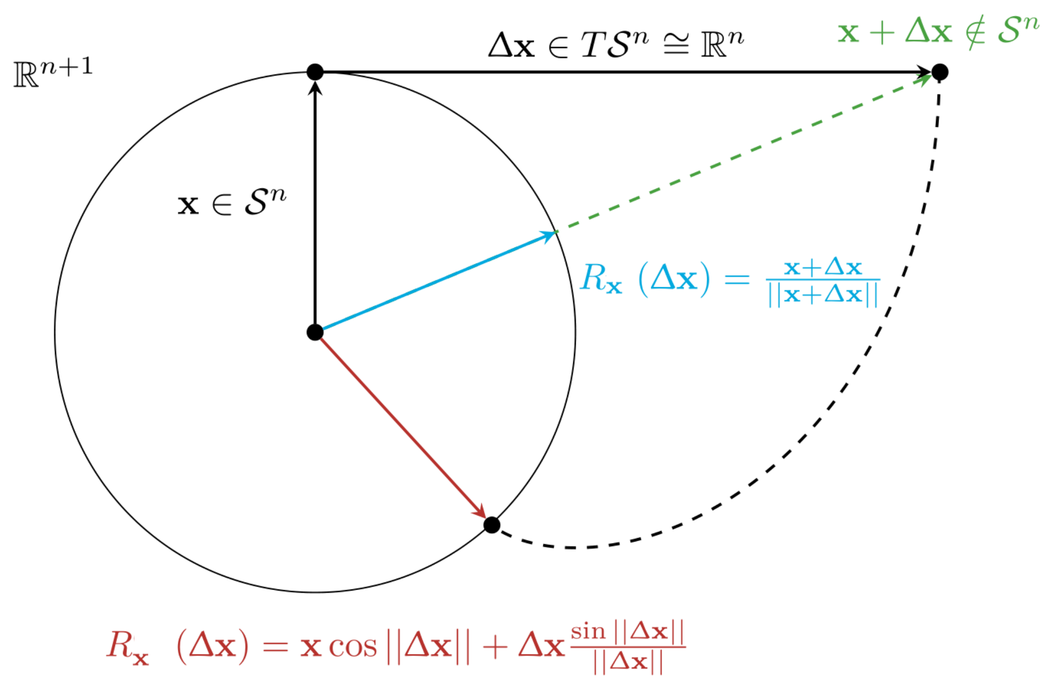
Baustatik und Baudynamik

Update of design variables

Retractions for the unit sphere

$$R_{\mathbf{x}}^{\exp}(\Delta\mathbf{x}) = \exp_{\mathbf{x}}(\Delta\mathbf{x}) = \cos(||\Delta\mathbf{x}||)\mathbf{x} + \frac{\sin(||\Delta\mathbf{x}||)}{||\Delta\mathbf{x}||}\Delta\mathbf{x}$$

$$R_{\mathbf{x}}^{\mathrm{rrn}}(\Delta\mathbf{x}) = \frac{\mathbf{x}+\Delta\mathbf{x}}{||\mathbf{x}+\Delta\mathbf{x}||}$$



Update of design variables

Retractions for the unit sphere

Taylor expansions coincide up to 2nd order

$$R_{\mathbf{x}}^{\exp}(\Delta\mathbf{x}) = \cos(||\Delta\mathbf{x}||)\mathbf{x} + \frac{\sin(||\Delta\mathbf{x}||)}{||\Delta\mathbf{x}||}\Delta\mathbf{x} = \mathbf{x} + \Delta\mathbf{x} - \frac{||\mathbf{x}||^2}{2}\mathbf{x} + \dots$$

$$R_{\mathbf{x}}^{\mathrm{rrn}}(\Delta\mathbf{x}) = \frac{\mathbf{x} + \Delta\mathbf{x}}{||\mathbf{x} + \Delta\mathbf{x}||} = \mathbf{x} + \Delta\mathbf{x} - \frac{||\mathbf{x}||^2}{2}\mathbf{x} + \dots$$

Riemannian gradient

Baustatik und Baudynamik

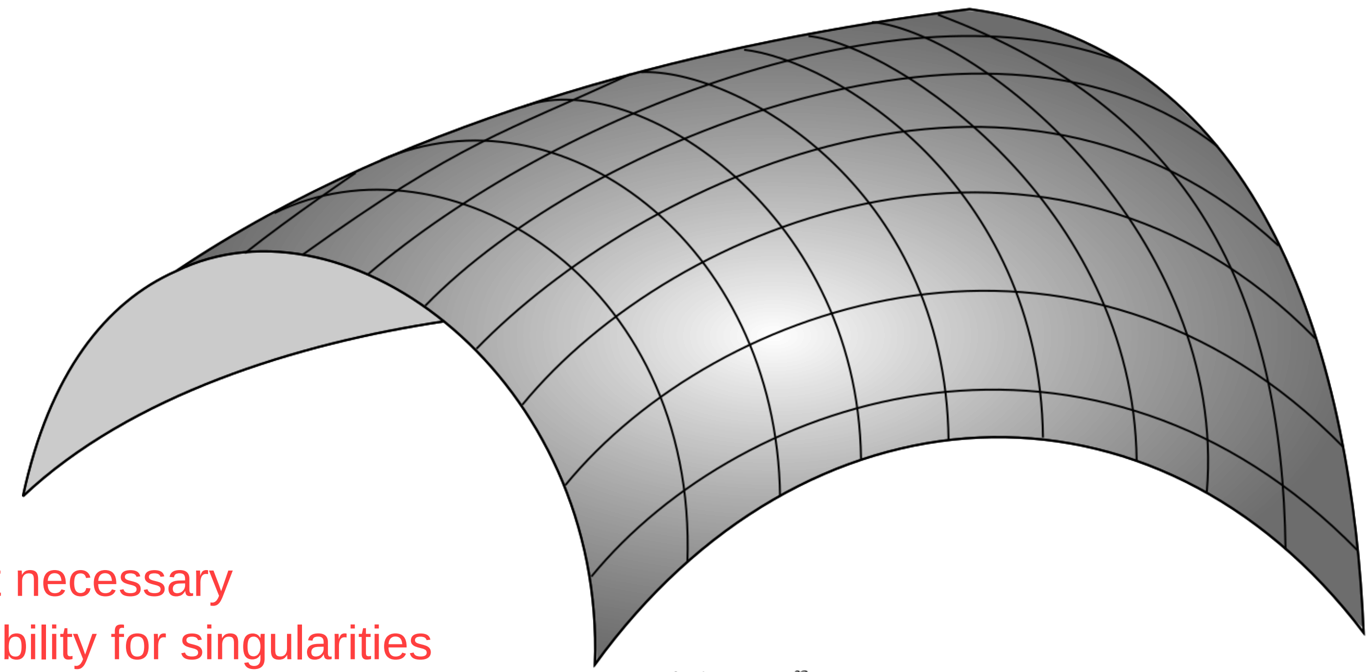
Riemannian Gradient

Introduce chart

$$f(\mathbf{x}) : \mathbb{R}^n \supset \mathcal{M} \to \mathbb{R}$$
$$\hat{f}(\boldsymbol{\varphi}) : \mathbb{R}^m \to \mathbb{R}$$
$$f(\mathbf{x}(\boldsymbol{\varphi})) = \hat{f}(\boldsymbol{\varphi})$$

$$\operatorname{grad} f(\mathbf{x}(\boldsymbol{\varphi})) = \frac{\partial f(\mathbf{x})}{\partial x^j}\frac{\partial x^j}{\partial \varphi^i}\mathbf{e}_i, \quad i = [1,\dots,m], j = [1,\dots,n]$$

$$\operatorname{grad} \hat{f}(\boldsymbol{\varphi}) = \frac{\partial \hat{f}(\boldsymbol{\varphi})}{\partial \varphi^i}\mathbf{e}_i, \quad i = [1,\dots,m]$$



$\mathcal{M} \subset \mathbb{R}^n$

Example:

$$\mathbf{x} \in \mathcal{S}^2, n = 3, m = 2$$
$$x^1 = \cos\varphi^1 \,\sin\varphi^2$$
$$x^2 = \sin\varphi^1 \,\sin\varphi^2$$
$$x^3 = \cos\varphi^2$$

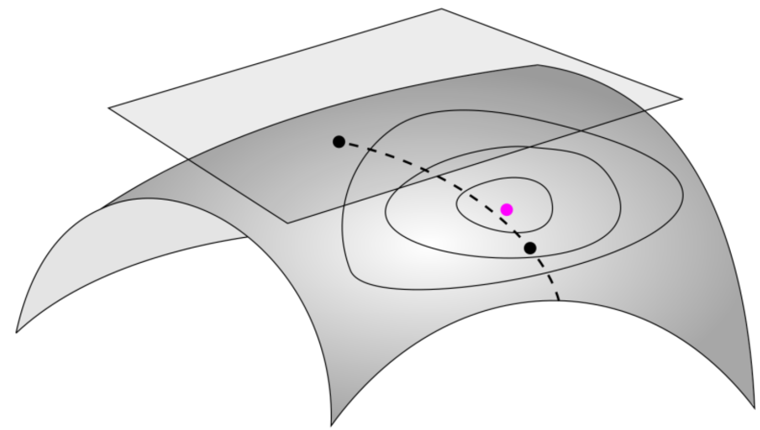
- Chart necessary
- Possibility for singularities
- Involved linearization

Riemannian Gradient

Definition of the Riemannian gradient

$$\mathrm{D}f(\mathbf{x})[\mathbf{v}] = \lim_{t=0} \frac{f(\mathbf{x}+t\mathbf{v}) - f(\mathbf{x})}{t} \qquad \mathbf{x},\mathbf{v} \in \mathbb{R}^n$$

$$\mathrm{D}f(\mathbf{x})[\mathbf{v}] = \lim_{t=0} \frac{f(\underline{\mathbf{x}+t\mathbf{v}}) - f(\mathbf{x})}{t} \qquad \mathbf{x} \in \mathcal{M},\mathbf{v} \in T_{\mathbf{x}}\mathcal{M}$$



Exploit retraction to lift the function

$$f : \mathcal{M} \to \mathbb{R} \qquad \hat{f} = f \circ R_{\mathbf{x}} : T_{\mathbf{x}}\mathcal{M} \to \mathbb{R}$$

$$\hat{f}(\boldsymbol{\eta}) = \bar{f}(R_{\mathbf{x}}(\boldsymbol{\eta})), \quad \boldsymbol{\eta} \in T_{\mathbf{x}}\mathcal{M}$$

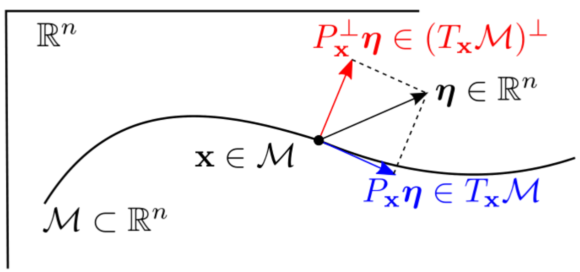
Absil PA, Mahony R, Sepulchre R (2008) Optimization Algorithms on Matrix Manifolds

Boumal N (2020) An introduction to optimization on smooth manifolds.

Optimization on Manifolds

Optimization on Riemannian submanifolds

Exploit embedding of the manifold



$$P_{\mathbf{x}} : \mathbb{R}^n \to T_{\mathbf{x}}\mathcal{M}$$

$$P_{\mathbf{x}}^{\perp} : \mathbb{R}^n \to (T_{\mathbf{x}}\mathcal{M})^{\perp}$$

Decompose using projections $\boldsymbol{\eta} = P_{\mathbf{x}}\boldsymbol{\eta} + P_{\mathbf{x}}^{\perp}\boldsymbol{\eta}$

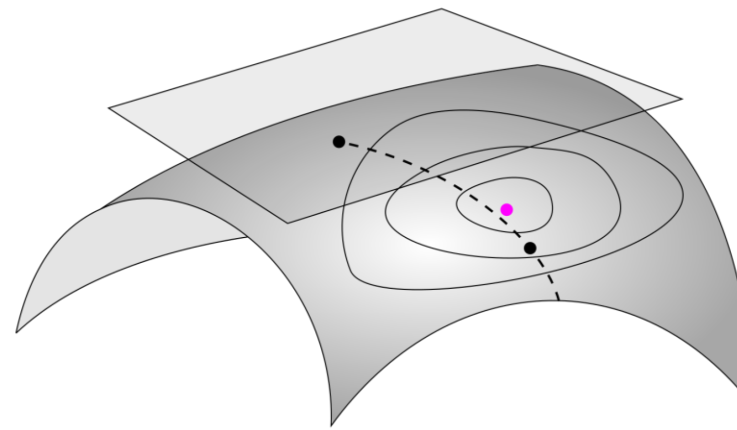
Optimization on Manifolds

Riemannian Gradient: submanifolds

Exploit embedding of the manifold

$f(\mathbf{x}) : \mathbb{R}^n \supset \mathcal{M} \to \mathbb{R}$ $\qquad \bar{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$

$\boldsymbol{\eta} \in T_{\mathbf{x}}\mathcal{M}$



$$\langle \operatorname{grad} \bar{f}(\mathbf{x}), \boldsymbol{\eta} \rangle = \langle \operatorname{grad} f(\mathbf{x}), \boldsymbol{\eta} \rangle$$

$$\langle P_{\mathbf{x}}(\operatorname{grad} \bar{f}(\mathbf{x})), \boldsymbol{\eta} \rangle + \langle P_{\mathbf{x}}^{\perp}(\operatorname{grad} \bar{f}(\mathbf{x})), \boldsymbol{\eta} \rangle = \langle \operatorname{grad} f(\mathbf{x}), \boldsymbol{\eta} \rangle$$

$$\langle P_{\mathbf{x}}(\operatorname{grad} \bar{f}(\mathbf{x})), \boldsymbol{\eta} \rangle = \langle \operatorname{grad} f(\mathbf{x}), \boldsymbol{\eta} \rangle$$

$$P_{\mathbf{x}}(\operatorname{grad} \bar{f}(\mathbf{x})) = \operatorname{grad} f(\mathbf{x})$$

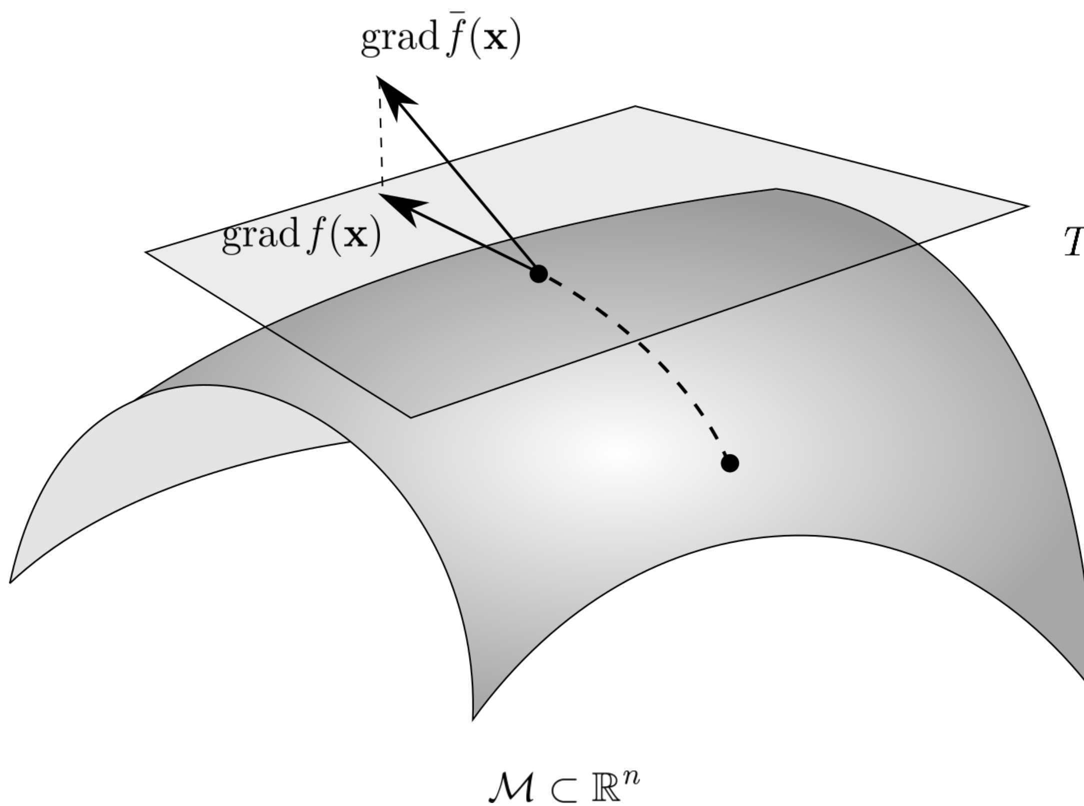
Details: Metric of embedding space

Riemannian Gradient: submanifolds

"For Riemannian submanifolds, the Riemannian gradient is the orthogonal projection of the "classical" gradient to the tangent spaces."

BOUMAL N (2020) AN INTRODUCTION TO OPTIMIZATION ON SMOOTH MANIFOLDS.

$$\operatorname{grad} f(\mathbf{x}) = P_{\mathbf{x}} \operatorname{grad} \bar{f}(\mathbf{x})$$



- No charts
- No artificial singularities
- Simple linearization

Optimization on Manifolds

Baustatik und Baudynamik

Toy problem: gradient and Riemannian gradient

$$E : \mathcal{S}^1 \to \mathbb{R}, \quad E(\mathbf{x}) = (\mathbf{x}+\mathbf{b})^T \mathbf{A} (\mathbf{x}+\mathbf{b})$$

$$\bar{E} : \mathbb{R}^2 \to \mathbb{R}, \quad \bar{E}(\mathbf{x}) = (\mathbf{x}+\mathbf{b})^T \mathbf{A} (\mathbf{x}+\mathbf{b})$$

$$\operatorname{grad} \bar{E}(\mathbf{x}) = \frac{\partial \bar{E}}{\partial \mathbf{x}} = 2\mathbf{A}(\mathbf{x}+\mathbf{b})$$

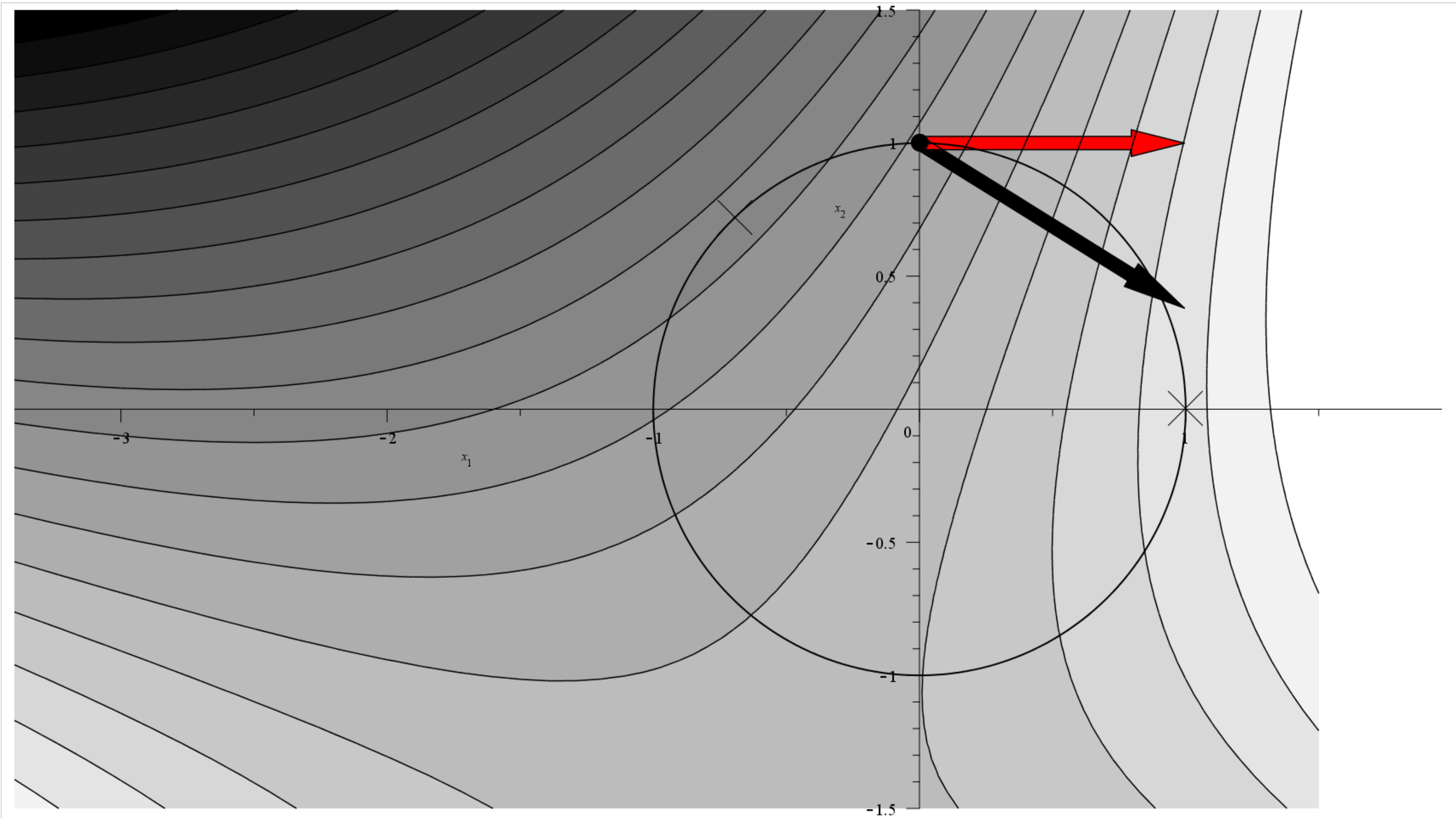
$$\mathbf{P}_{\mathbf{x}} = \mathbf{I} - \mathbf{x}\mathbf{x}^T$$

$$\Rightarrow$$

$$\operatorname{grad} E(\mathbf{x}) = 2\mathbf{P}_{\mathbf{x}} \mathbf{A} (\mathbf{x}+\mathbf{b})$$

Optimization on Manifolds

Toy problem: gradient and Riemannian gradient



Riemannian Hessian

Optimization on Manifolds

Riemannian Hessian: submanifolds

Introduce chart

$$f(\mathbf{x}) : \mathbb{R}^n \supset \mathcal{M} \to \mathbb{R}$$

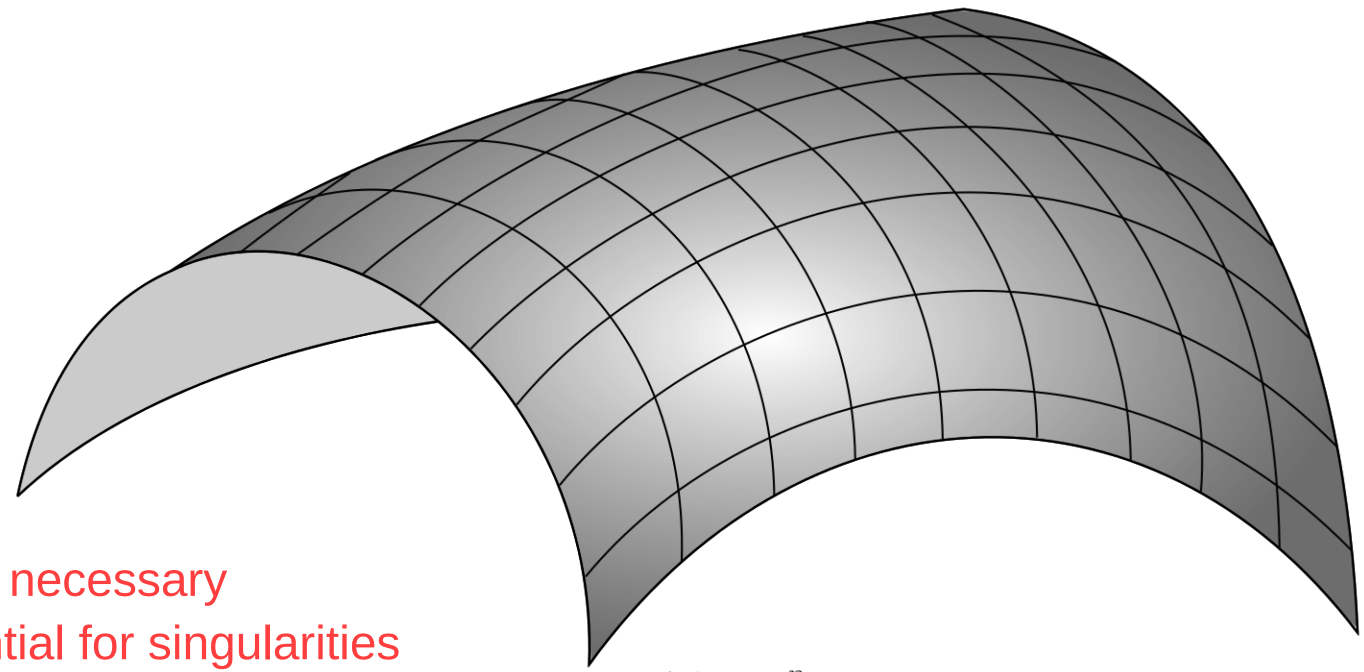
$$\hat{f}(\boldsymbol{\varphi}) : \mathbb{R}^m \to \mathbb{R}$$

$$f(\mathbf{x}(\boldsymbol{\varphi})) = \hat{f}(\boldsymbol{\varphi})$$

$$\operatorname{grad} f(\mathbf{x}(\boldsymbol{\varphi})) = \frac{\partial f(\mathbf{x})}{\partial x^j} \frac{\partial x^j}{\partial \varphi^i} \mathbf{e}_i$$

Christoffel symbols

$$\operatorname{Hess} f(\mathbf{x}(\boldsymbol{\varphi})) = \frac{\partial^2 f(\mathbf{x})}{\partial x^j \partial x^k} \frac{\partial x^j}{\partial \varphi^i} \frac{\partial x^k}{\partial \varphi^l} \mathbf{e}_i \otimes \mathbf{e}_l + \frac{\partial f(\mathbf{x})}{\partial x^j} \frac{\partial^2 x^j}{\partial \varphi^i \partial \varphi^l} \mathbf{e}_i \otimes \mathbf{e}_l$$



$\mathcal{M} \subset \mathbb{R}^n$

Example:

$$\mathbf{x} \in \mathcal{S}^2, n = 3, m = 2$$

$$x^1 = \cos \varphi^1 \ \sin \varphi^2$$

$$x^2 = \sin \varphi^1 \ \sin \varphi^2$$

$$x^3 = \cos \varphi^2$$

- Chart necessary
- Potential for singularities
- Involved linearization
- Bad conditioning of the Hessian

Riemannian Hessian: submanifolds

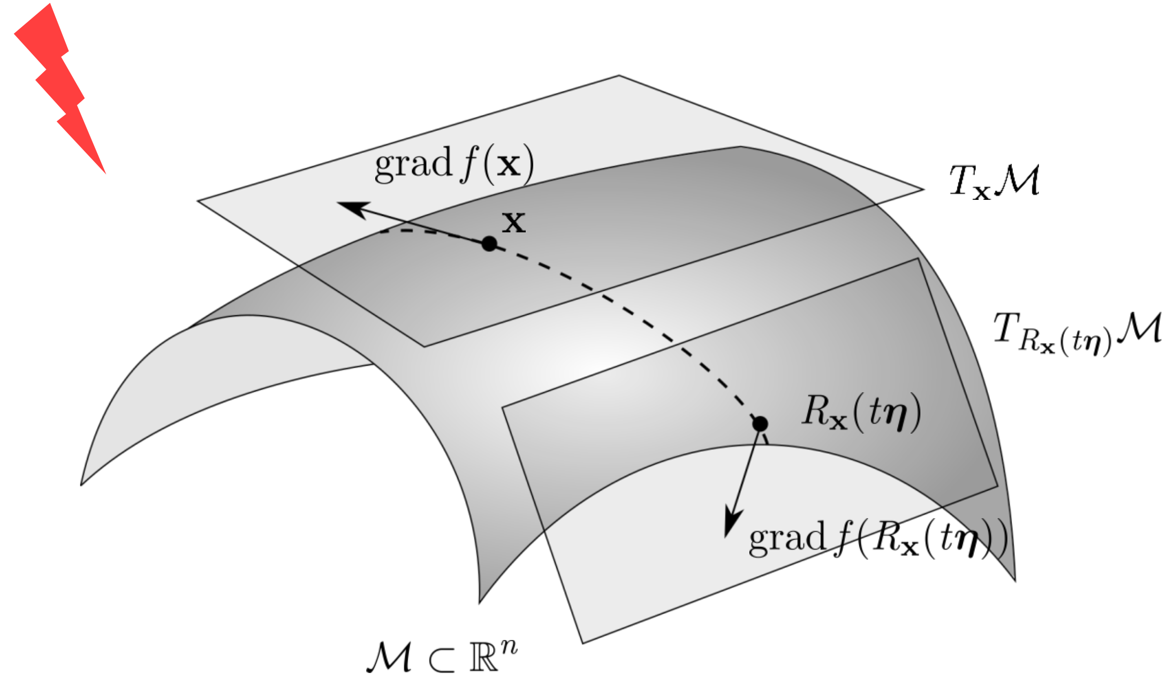
Definition of the Riemannian Hessian

$$\mathrm{D}\,\mathrm{grad}\, f(\mathbf{x})[\boldsymbol{\eta}] = \lim_{t=0} \frac{\mathrm{grad}\, f(R_{\mathbf{x}}(t\boldsymbol{\eta})) - \mathrm{grad}\, f(\mathbf{x})}{t}$$

$$\boldsymbol{\eta} \in T_{\mathbf{x}}\mathcal{M}$$

Exploit retraction to lift the function

$$f : \mathcal{M} \to \mathbb{R} \qquad \hat{f} = f \circ R_{\mathbf{x}} : T_{\mathbf{x}}\mathcal{M} \to \mathbb{R}$$



ABSIL PA, MAHONY R, SEPULCHRE R (2008) OPTIMIZATION ALGORITHMS ON MATRIX MANIFOLDS
BOUMAL N (2020) AN INTRODUCTION TO OPTIMIZATION ON SMOOTH MANIFOLDS.

Optimization on Manifolds

Baustatik und Baudynamik

Riemannian Hessian: submanifolds

1. Exploit embedding of the manifold to lift function $f(R_{\mathbf{x}}(\boldsymbol{\eta}))$
2. Taylor expansion $\hat{f}(t\boldsymbol{\eta}) = \bar{f}(\mathbf{x}) + \langle \operatorname{grad} \bar{f}(\mathbf{x}), \boldsymbol{\gamma}' \rangle + \langle \operatorname{Hess} \bar{f}(\mathbf{x}) \boldsymbol{\gamma}', \boldsymbol{\gamma}' \rangle + \langle \operatorname{grad} \bar{f}(\mathbf{x}), \boldsymbol{\gamma}'' \rangle + \mathcal{O}(t^3)$
3. Equating terms

$$\boldsymbol{\gamma}' = \mathrm{D} R_{\mathbf{x}}(\mathbf{0})[t\boldsymbol{\eta}] = t\boldsymbol{\eta}$$
$$\boldsymbol{\gamma}'' = \mathrm{D}^2 R_{\mathbf{x}}(\mathbf{0})[t\boldsymbol{\eta}]$$

$$\langle \operatorname{Hess} f(\mathbf{x})\boldsymbol{\eta}, \boldsymbol{\eta} \rangle = \langle \operatorname{Hess} \bar{f}(\mathbf{x})\boldsymbol{\eta}, \boldsymbol{\eta} \rangle + \langle \operatorname{grad} \bar{f}(\mathbf{x}), \mathrm{D}_{\boldsymbol{\eta}} R_{\mathbf{x}}(\boldsymbol{\eta}) \rangle$$

$$\operatorname{Hess} f(\mathbf{x})\boldsymbol{\eta} = P_{\mathbf{x}} \operatorname{Hess} \bar{f}(\mathbf{x}) P_{\mathbf{x}}\boldsymbol{\eta} + W_{\mathbf{x}}(\boldsymbol{\eta}, P_{\mathbf{x}}^{\perp} \operatorname{grad} \bar{f}(\mathbf{x}))$$

"[…] This shows that, for Riemannian submanifolds of Euclidean spaces, the Riemannian Hessian is the projected Euclidean Hessian plus a correction term which depends only on the normal part of the Euclidean gradient."

BOUMAL N (2020) AN INTRODUCTION TO OPTIMIZATION ON SMOOTH MANIFOLDS.

Details: Connection, Levi-Civita-Connection, Parallel transport, vector transport, Weingarten map, shape operator, second order retractions,…

Optimization on Manifolds

Baustatik und Baudynamik

Toy problem: Hessian and Riemannian Hessian

$$E : \mathcal{S}^1 \to \mathbb{R}, \quad E(\mathbf{x}) = (\mathbf{x}+\mathbf{b})^T \mathbf{A}(\mathbf{x}+\mathbf{b})$$

$$\bar{E} : \mathbb{R}^2 \to \mathbb{R}, \quad \bar{E}(\mathbf{x}) = (\mathbf{x}+\mathbf{b})^T \mathbf{A}(\mathbf{x}+\mathbf{b})$$

$$\operatorname{grad} \bar{E}(\mathbf{x}) = \frac{\partial \bar{E}}{\partial \mathbf{x}} = 2\mathbf{A}(\mathbf{x}+\mathbf{b})$$

$$\mathbf{P}_{\mathbf{x}} = \mathbf{I} - \mathbf{x}\mathbf{x}^T$$

$$\operatorname{Hess} \bar{E}(\mathbf{x}) = \frac{\partial^2 \bar{E}}{\partial \mathbf{x}^2} = 2\mathbf{A}$$

$$W_{\mathbf{x}}(\boldsymbol{\eta},\mathbf{z}) = -\boldsymbol{\eta}\mathbf{x}^T\mathbf{z}$$

$$\Rightarrow$$

$$\operatorname{grad} E(\mathbf{x}) = 2\mathbf{P}_{\mathbf{x}}\mathbf{A}(\mathbf{x}+\mathbf{b})$$

$$\begin{aligned}\operatorname{Hess} E(\mathbf{x}) &= 2\mathbf{P}_{\mathbf{x}}\mathbf{A}\mathbf{P}_{\mathbf{x}} - (\mathbf{x}^T \operatorname{grad} \bar{E}(\mathbf{x}))\mathbf{I} \\ &= 2\mathbf{P}_{\mathbf{x}}\mathbf{A}\mathbf{P}_{\mathbf{x}} - 2\mathbf{x}^T(\mathbf{A}(\mathbf{x}+\mathbf{b}))\mathbf{I}\end{aligned}$$

- Simple linearization
- No artificial singularities

Riemannian Newton

Motivation

Riemannian Newton

	Classic optimization	Optimization on Riemannian submanifolds
Update	$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k$	$\mathbf{x}_{k+1} = R_{\mathbf{x}_k}(\Delta\mathbf{x}_k)$
Gradient	$\operatorname{grad} f(\mathbf{x})$	$\operatorname{grad} f(\mathbf{x}) = P_{\mathbf{x}} \operatorname{grad} \bar{f}(\mathbf{x})$
Hessian	$\operatorname{Hess} f(\mathbf{x})$	$\operatorname{Hess} f(\mathbf{x})\boldsymbol{\eta} = P_{\mathbf{x}} \operatorname{Hess} \bar{f}(\mathbf{x}) P_{\mathbf{x}}\boldsymbol{\eta} + W_{\mathbf{x}}(\boldsymbol{\eta}, P_{\mathbf{x}}^{\perp} \operatorname{grad} \bar{f}(\mathbf{x}))$

Algorithm 1: Riemannian Newton

Goal : Find stationary point of $E : \mathcal{M} \to \mathbb{R}$, i.e. find $\mathbf{x}$ such that $\operatorname{grad} E(\mathbf{x}) = \mathbf{0}$.

input : Initial iterate $\mathbf{x}_0$

output: Converged solution: $\mathbf{x}^*$

while $||\operatorname{grad} E(\mathbf{x}_k)|| > tol$ **do**

- $\operatorname{Hess} E(\mathbf{x}_k)\Delta\mathbf{x}_k = -\operatorname{grad} E(\mathbf{x}_k)$
- $\mathbf{x}^{k+1} = R_{\mathbf{x}^k}(\Delta\mathbf{x}_k)$
- $k \leftarrow k + 1$

end

Ingredients:

$$R_{\mathbf{x}} : T_{\mathbf{x}}\mathcal{M} \to \mathcal{M}$$

$$\mathbf{P}_{\mathbf{x}} : \mathbb{R}^n \to T_{\mathbf{x}}\mathcal{M}$$

$$W_{\mathbf{x}} : T_{\mathbf{x}}\mathcal{M} \times T_{\mathbf{x}}^{\perp}\mathcal{M} \to T_{\mathbf{x}}\mathcal{M}$$

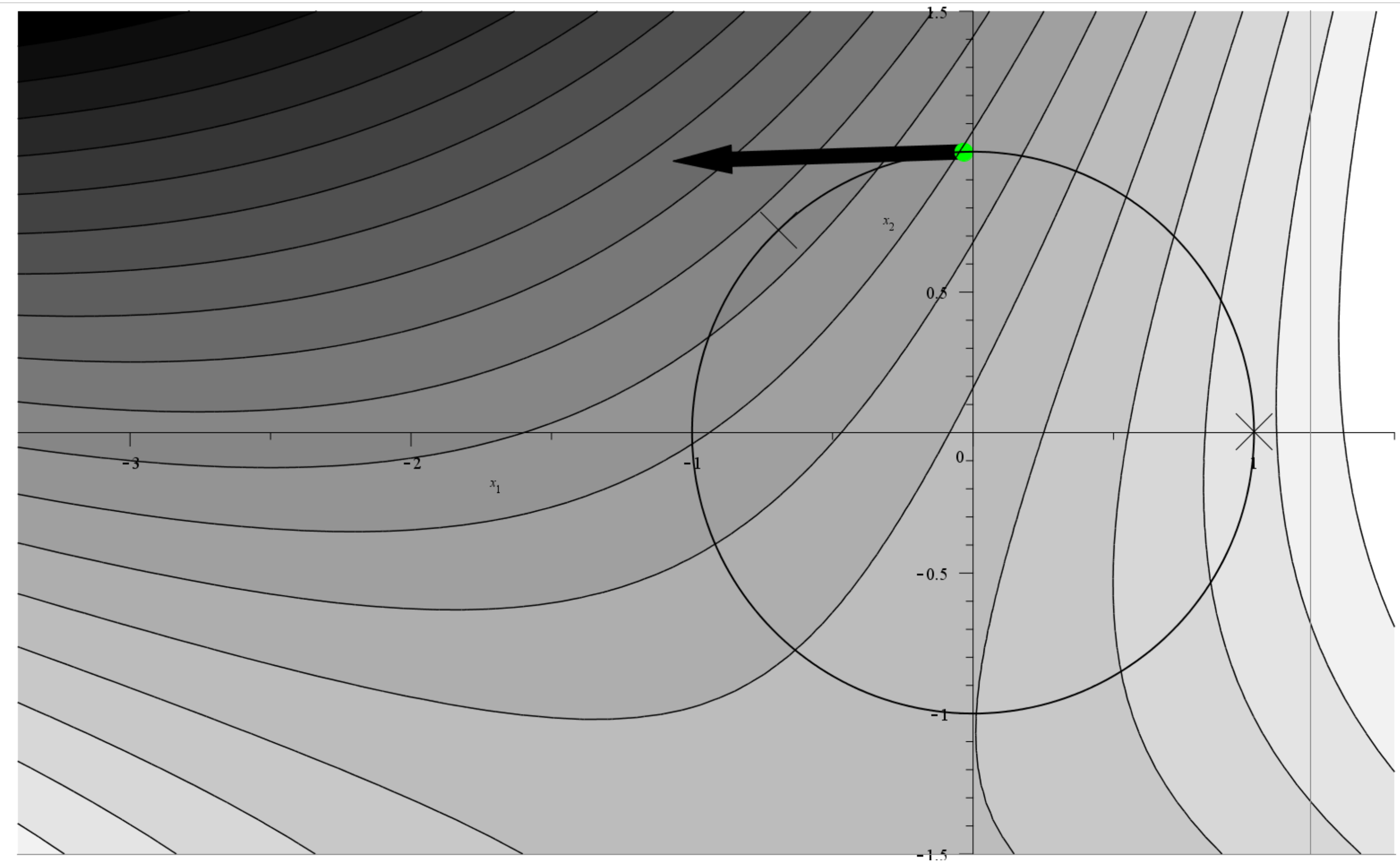
$$\operatorname{Hess} \bar{E}(\mathbf{x})$$

$$\operatorname{grad} \bar{E}(\mathbf{x})$$

Optimization on Manifolds

Baustatik und Baudynamik

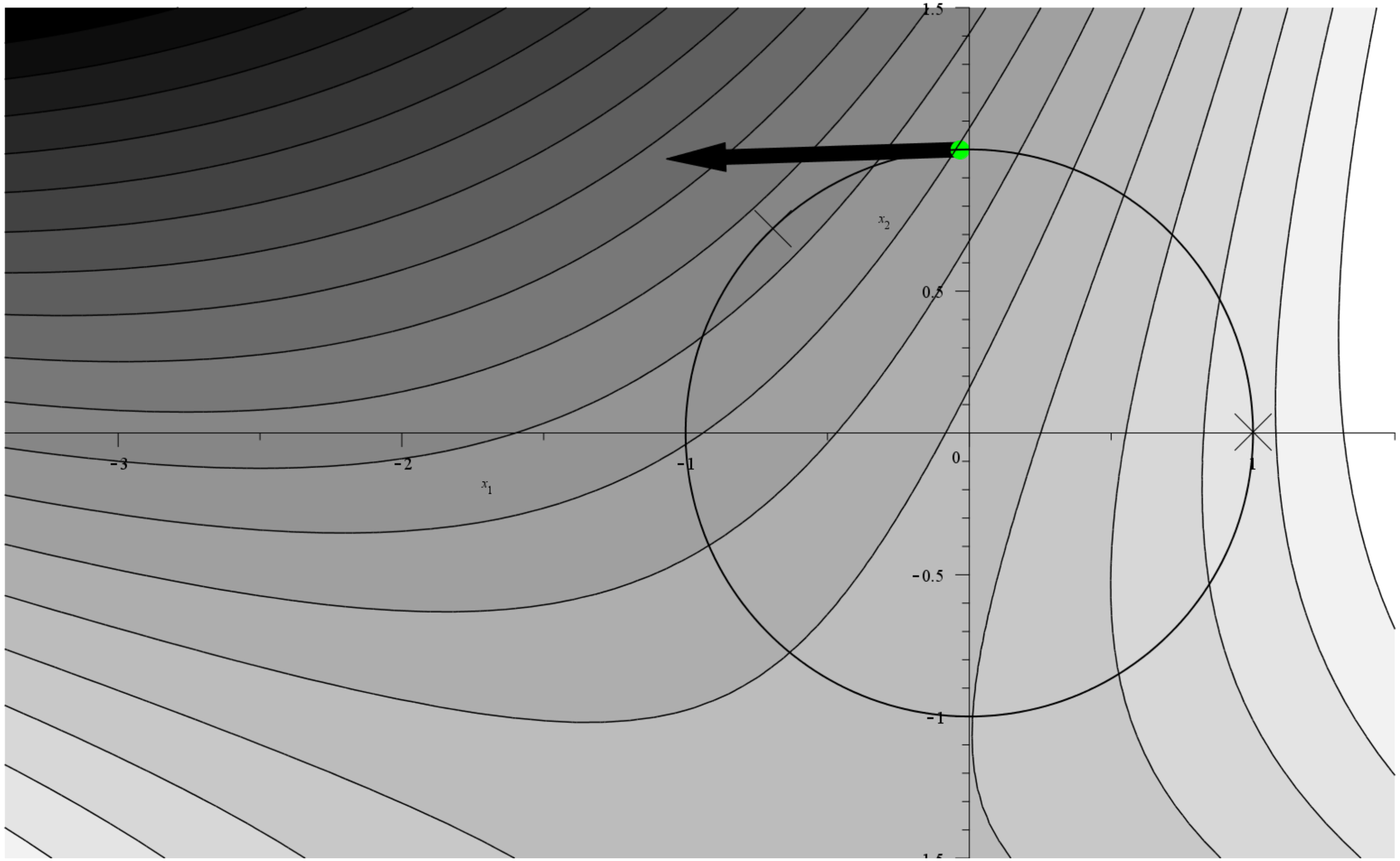
Toy problem: Exponential map



Optimization on Manifolds

Baustatik und Baudynamik

Toy problem: Radial return

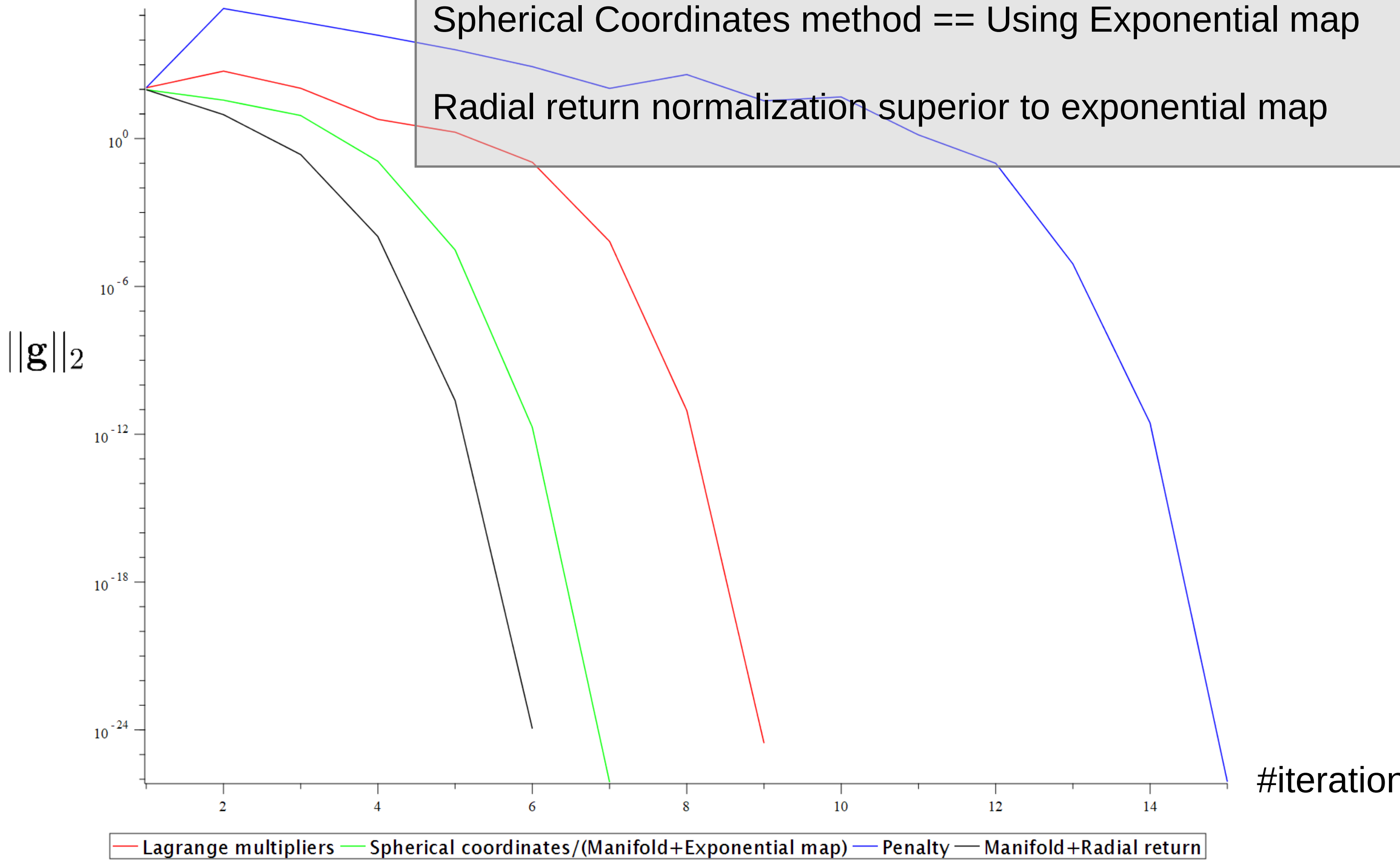


Optimization on Manifolds

Toy problem: Newton’s method, iteration count vs. gradient norm

Spherical Coordinates method == Using Exponential map

Radial return normalization superior to exponential map



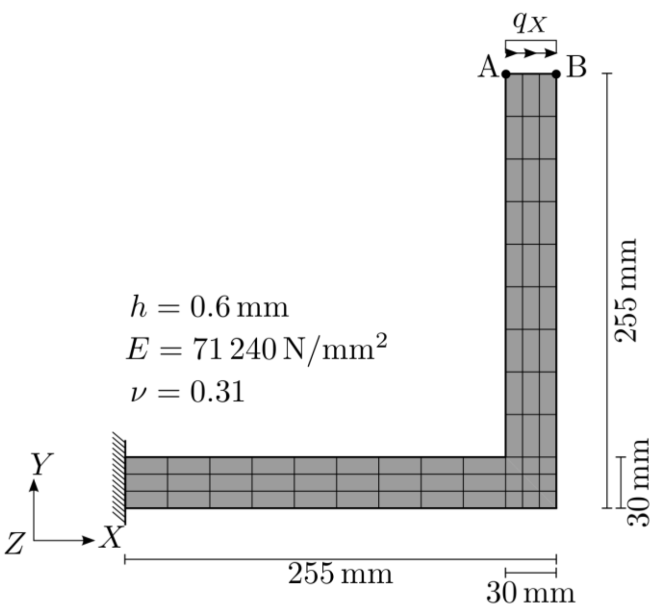
Simulations

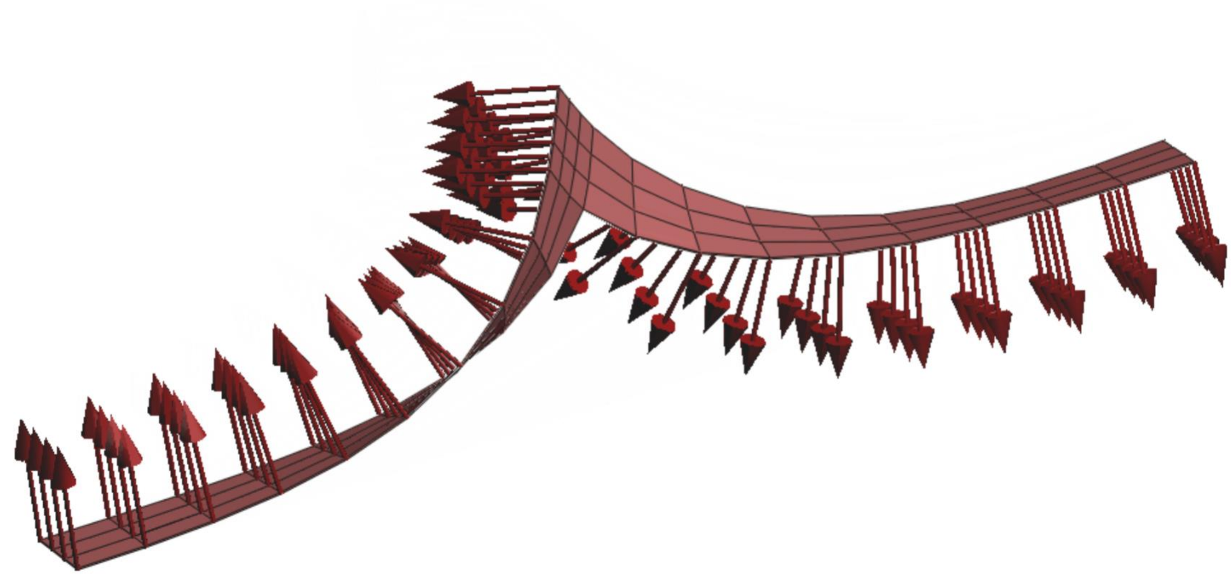
Simulation of Reissner-Mindlin shells

Simulation elastic deformation of shells

$$\Pi(\boldsymbol{\varphi}, \mathbf{t}) = \int_{\mathcal{B}} \psi_{\text{strain}}(\boldsymbol{\varphi}, \mathbf{t}) + \mathrm{d}V + L(\boldsymbol{\varphi}, \mathbf{t})$$

$$\{\boldsymbol{\varphi}^*, \mathbf{t}^*\} = \arg\left\{ \min_{\boldsymbol{\varphi}\in\mathbb{R}^d} \min_{\mathbf{t}\in\mathcal{S}^{d-1}} \Pi(\boldsymbol{\varphi}, \mathbf{t}) \right\}.$$



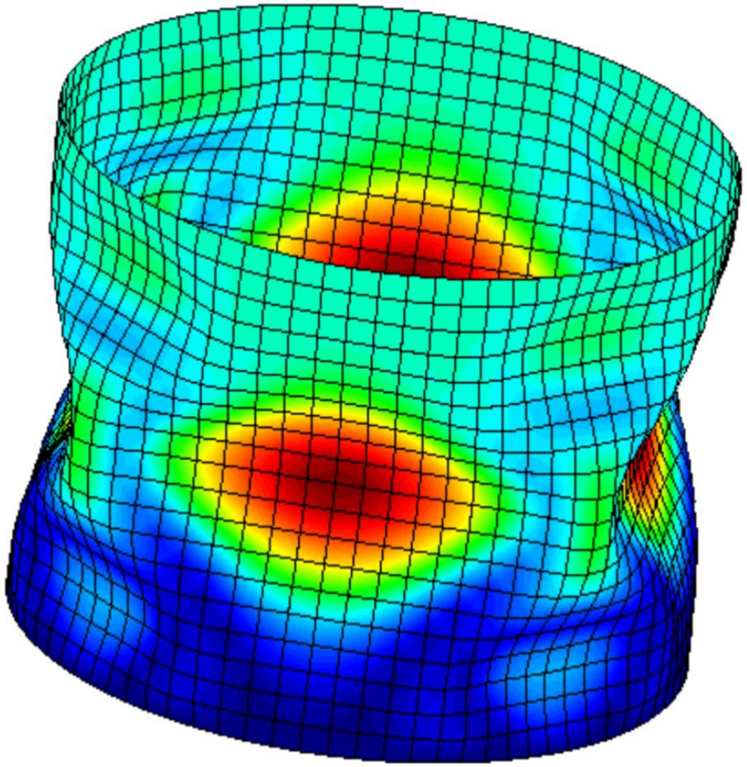


AM, BISCHOFF (2022): A CONSISTENT FINITE ELEMENT FORMULATION OF THE GEOMETRICALLY NON-LINEAR REISSNER-MINDLIN SHELL MODEL, DOI

Simulation of Reissner-Mindlin shells

Simulation elastic deformation of shells

Riemannian **Trust-Region** method



Minimizers for cylinder buckling

Simulation of magnetic vorticies

Experiments

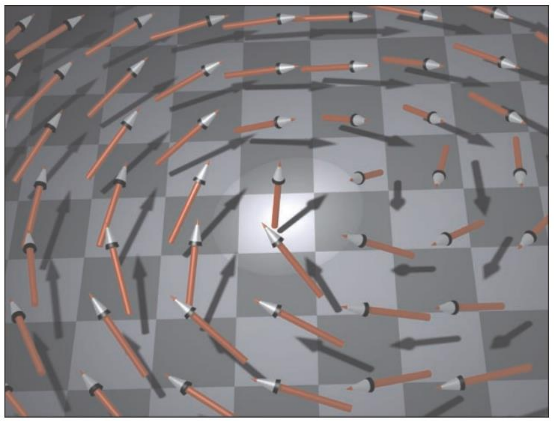
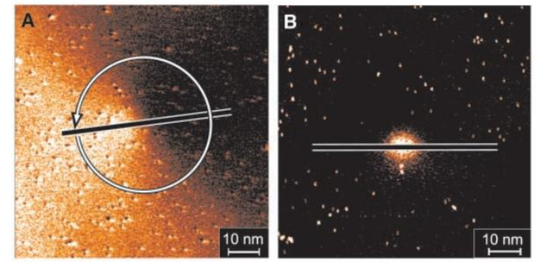


Fig. 1. Schematic of a vortex core. Far away from the vortex core the magnetization continuously curls around the center with the orientation in the surface plane. In the center of the core the magnetization is perpendicular to the plane (highlighted).

Wachowiak et al (2002): Direct Observation of Internal Spin Structure of Magnetic Vortex Cores



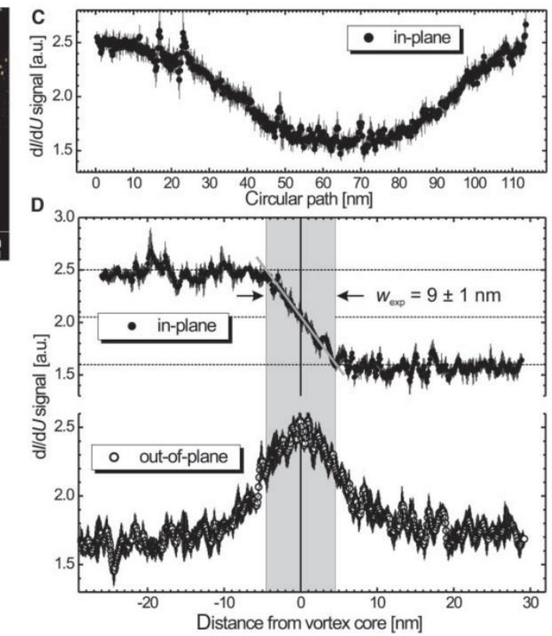


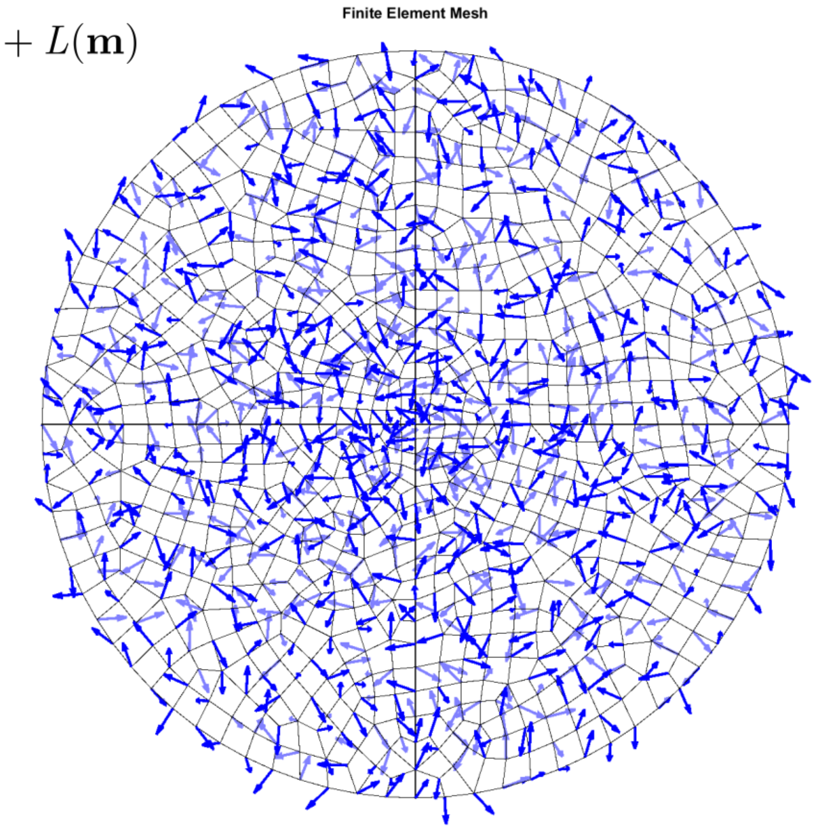
Fig. 3. Magnetic dI/dU maps as measured with an (**A**) in-plane and an (**B**) out-of-plane sensitive Cr tip. The curling in-plane magnetization around the vortex core is recognizable in (A), and the perpendicular magnetization of the vortex core is visible as a bright area in (B). (**C**) dI/dU signal around the vortex core at a distance of 19 nm [circle in (A)]. (**D**) dI/dU signal along the lines in (A) and (B). The measurement parameters were (A) I = 0.6 nA, U_0 = −300 mV and (B) I = 1.0 nA, U_0 = −350 mV.

Simulation of magnetic vortices

Simulation micromagnetostatics

$$\Pi(\tilde{\mathbf{A}}, \mathbf{m}) = \int_{\mathcal{B}} \psi_{\text{ex}}(\nabla \mathbf{m}) + \psi_{\text{demag}}(\tilde{\mathbf{A}}, \mathbf{m}) + \psi_{\text{a}}(\mathbf{m}) \, \mathrm{d}V + L(\mathbf{m})$$

$$\left\{ \tilde{\mathbf{A}}^*, \mathbf{m}^* \right\} = \arg \left\{ \min_{\mathbf{A} \in \mathbb{R}^d} \min_{\boxed{\mathbf{m} \in \mathcal{S}^{d-1}}} \Pi(\tilde{\mathbf{A}}, \mathbf{m}) \right\}.$$



Other manifolds

Optimization on Manifolds

Other manifolds

Manifold	Exponential map	Other retractions	Tangent space
Unit sphere $\mathcal{S}^{n-1}$	$\cos\|\|\Delta\mathbf{x}\|\|\mathbf{x} + \frac{\sin\|\|\Delta\mathbf{x}\|\|}{\|\|\Delta\mathbf{x}\|\|}\Delta\mathbf{x}$	$\frac{\mathbf{x}+\Delta\mathbf{x}}{\|\|\mathbf{x}+\Delta\mathbf{x}\|\|}$	$T_{\mathbf{x}}\mathcal{S}^{n-1} = \{\mathbf{y}\in\mathbb{R}^n \mid \mathbf{y}^T\mathbf{x} = 0\}$
Special linear group $\mathcal{SL}(n) = \{\mathbf{X}\in\mathbb{R}^{n\times n} \|\| \det\mathbf{X} = 1\}$	$\exp(\Delta\mathbf{X})\mathbf{X}$	$\frac{\mathbf{X}+\Delta\mathbf{X}}{\det(\mathbf{X}+\Delta\mathbf{X})^{1/n}}$	$T_{\mathbf{X}}\mathcal{SL}(n) = \{\mathbf{Y}\in\mathbb{R}^{n\times n} \|\| \operatorname{tr}\mathbf{Y} = 0\}$
Symmetric Special linear manifold $\mathcal{SSL}(n) = \{\mathbf{X}\in\mathbb{R}^{n\times n} \|\|$ $\det\mathbf{X} = 1 \wedge \mathbf{X}^T = \mathbf{X}\}$	$\exp(\Delta\mathbf{X})\mathbf{X}$	$\frac{\mathbf{X}+\Delta\mathbf{X}}{\det(\mathbf{X}+\Delta\mathbf{X})^{1/n}}$	$T_{\mathbf{X}}\mathcal{SSL}(n) = \{\mathbf{Y}\in\mathbb{R}^{n\times n} \|\| \operatorname{tr}(\mathbf{X}^{-1}\mathbf{Y}) = 0\}$
Special orthogonal group $\mathcal{SO}(n) = \{\mathbf{X}\in\mathbb{R}^{n\times n} \|\|$ $\mathbf{X}^T\mathbf{X} = \mathbf{I} \wedge \det\mathbf{X} = 1\}$	$\exp(\Delta\mathbf{X})\mathbf{X}$	QR decomposition	$T_{\mathbf{X}}\mathcal{SO}(n) = \{\mathbf{Y}\in\mathbb{R}^{n\times n} \|\| \mathbf{Y}^T = -\mathbf{Y}\}$

Optimization on Manifolds

Retractions, Symmetric positive definite matrices

$$\mathcal{SPD}(n) = \{X \in \mathbb{R}^{n\times n} : X = X^T \wedge X \succ 0\}$$

$$T_{\mathbf{X}}\mathcal{SPD}(n) = \{\mathbf{Y} \in \mathbb{R}^{n\times n} : \mathbf{Y} = \mathbf{Y}^T\}$$

Exponential map: $\mathbf{X}_{k+1} = \mathbf{X}_k \exp(\mathbf{X}_k^{-1}\Delta\mathbf{X})$

Second order: $\mathbf{X}_{k+1} = \mathbf{X}_k + \Delta\mathbf{X}_k + \frac{1}{2}\Delta\mathbf{X}_k\mathbf{X}_k^{-1}\Delta\mathbf{X}_k$ Stays always on SPD!

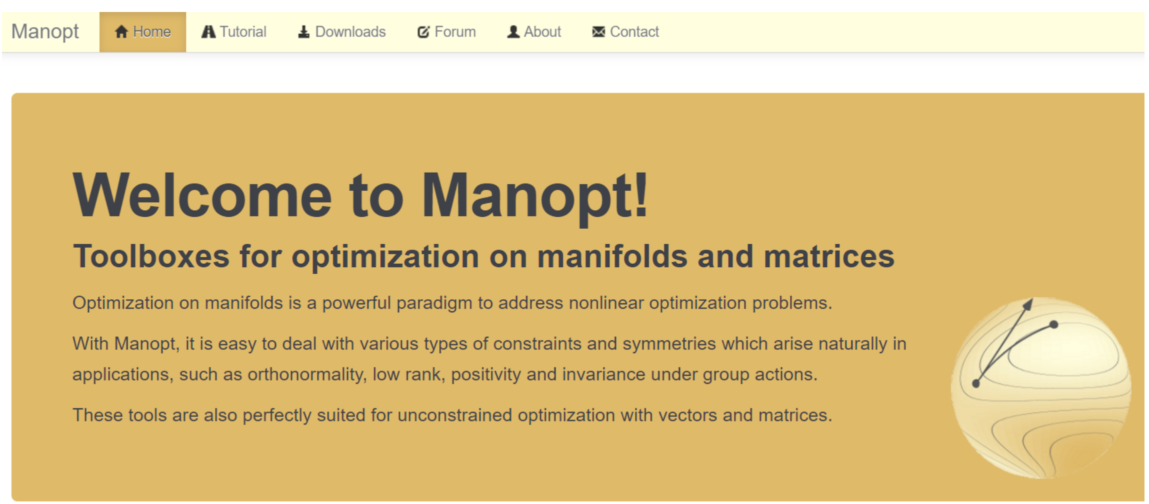
First order: $\mathbf{X}_{k+1} = \mathbf{X}_k + \Delta\mathbf{X}_k$ Stays not always on SPD!

HUANG, W. (2017). INTRODUCTION TO RIEMANNIAN BFGS METHODS. LINK

Optimization on Manifolds

Other manifolds

Manifold	Definition
Symmetric positive semidefinite fixed-rank	$\{\mathbf{X} \in \mathbb{R}^{n \times n} : \mathbf{X} = \mathbf{X}^T \succeq 0, \mathrm{rank}(\mathbf{X}) = k\}$
Generalized sphere	$\{\mathbf{X} \in \mathbb{R}^{n \times m} : \lVert\mathbf{X}\rVert_{\mathrm{F}} = 1\}$
Fixed rank	$\{\mathbf{X} \in \mathbb{R}^{m \times n} : \mathrm{rank}(\mathbf{X}) = k\}$
Manifold defined by some function $\mathbf{h}(\mathbf{X}) = \mathbf{0}$	…



Boumal, N. and Mishra, B. and Absil, P.-A. and Sepulchre, R., Manopt, a Matlab Toolbox for Optimization on Manifolds, 2014

https://www.manopt.org/

Optimization on Manifolds

Summary

- Optimization of cost functions with constraints can be interpreted as optimization on manifolds
- Lots of customization points exist, e.g., retractions
- May be superior to classical optimization techniques
- Many manifolds can be found in literature

Not mentioned: Finite Elements

Lots of pitfalls!

Examples:

The important Sobolev space $W^{1,2}(\Omega, \mathcal{M})$ does not even always possess the structure of a Banach manifold

Interpolation has to stay on the manifold

Geodesics for interpolation are not always unique

References

Baustatik und Baudynamik

Algebraic consideration of optimization of manifolds:

Rosen JB (1961) *The gradient projection method for nonlinear programming. Part II. nonlinear constraints*. SIAM 9(4):514–532, doi:10.1137/0109044

Luenberger, David G. (1973) *Introduction to linear and nonlinear programming*. Vol. 28. Reading, MA: Addison-wesley,.

Adler RL, Dedieu JP, Margulies JY, Martens M, Shub M (2002) *Newton's method on Riemannian manifolds and a geometric model for the human spine*. IMA J Numer Anal 22(3):359–390, doi:10.1093/imanum/22.3.359

Absil PA, Mahony R, Sepulchre R (2008) *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, doi:10.1515/9781400830244

Absil PA, Malick J (2012) *Projection-like retractions on matrix manifolds*. SIAM J Optim 22(1):135–158, doi:10.1137/100802529

Absil PA, Mahony R, Trumpf J (2013) *An extrinsic look at the riemannian hessian*. In: Nielsen F, Barbaresco F (eds) Geometric Science of Information, pp 361–368, doi:10.1007/978-3-642-40020-9_39

Huang W (2017) *). Introduction to riemannian bfgs methods*. Available online, URL Link to online resource

Boumal N (2020) *An introduction to optimization on smooth manifolds*. Available online, URL Link to online resource

Finite elements for manifolds:

Grohs P (2011) *Finite elements of arbitrary order and quasiinterpolation for data in Riemannian manifolds*. Tech. Rep. 2011-56, Seminar for Applied Mathematics, ETH Zürich, URL https://www.sam.math.ethz.ch/sam_reports/reports_final/reports2011/2011-56.pdf

Sander O (2012) *Geodesic finite elements on simplicial grids*. Int J Numer Methods Eng 92(12):999–1025, doi:10.1002/nme.4366

Grohs P, Hardering H, Sander O (2015) *Optimal A Priori Discretization Error Bounds for Geodesic Finite Elements*. Found Comut Math 15(6):1357–1411, doi:10.1007/s10208-014-9230-z

Hardering H (2018) *L2-discretization error bounds for maps into Riemannian manifolds*. Numer Math (Heidelb) 139(2):381–410, doi:10.1007/s00211-017-0941-3

Physical simulations (Small sample):

Sander O, Neff P, Bırsan M (2016) *Numerical treatment of a geometrically nonlinear planar Cosserat shell model*. Comput Mech 57(5):817–841, doi:10.1007/s00466-016-1263-5

AM, Manfred Bischoff (2022) *A Consistent Finite Element Formulation of the Geometrically Non-linear Reissner-Mindlin Shell Model*. DOI :10.1007/s11831-021-09702-7.

Universität Stuttgart

Tack så mycket! Vielen Dank!

Alex Müller

Email mueller@ibb.uni-stuttgart.de

University of Stuttgart
Institute for Structural Mechanics
Pfaffenwaldring 7, 70569 Stuttgart, Germany