

MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

Report No. 20/2018

DOI: 10.4171/OWR/2018/20

Nonlinear Data: Theory and Algorithms

Organised by
Philipp Grohs, Wien
Oliver Sander, Dresden
Jean-Luc Starck, Gif-sur-Yvette
Johannes Wallner, Graz

22 April – 28 April 2018

ABSTRACT. Techniques and concepts from differential geometry are used in many parts of applied mathematics today. However, there is no joint community for users of such techniques. The workshop on Nonlinear Data assembled researchers from fields like numerical linear algebra, partial differential equations, and data analysis to explore differential geometry techniques, share knowledge, and learn about new ideas and applications.

Mathematics Subject Classification (2010): 65-XX Numerical analysis, 53-XX Differential geometry.

Introduction by the Organisers

Differential geometry appears in a variety of fields of applied mathematics today. Various linear algebra problems can be written as minimization problems on Riemannian manifolds. Partial differential equations with values in nonlinear spaces appear in physical applications. The analysis of shapes in geometry processing and computer graphics involves analysis on infinite-dimensional nonlinear shape spaces. However, no real community exists for these aspects of applied mathematics, and there is little exchange on techniques from differential geometry between the different fields.

Setting out to change this, the workshop “Nonlinear Data” gathered people from several different fields, involving nonlinear data analysis, numerics of partial differential equations, numerical linear algebra, and geometric analysis. This was a great success, with many participants stating how much they enjoyed meeting so

many new people. There was lots of interaction between the different communities, and a very fruitful exchange of ideas.

At the workshop there were 45 participants from 11 different countries, 5 participants even from overseas. 24 talks were given, which could be loosely grouped as follows:

- **Shape analysis:** Problems involving optimal shapes or matching of shapes are typically formulated in nonlinear Shape Space.
- **Image processing:** Several talks considered processing of manifold-valued images, which is receiving more and more interest.
- **Partial differential equations for manifold-valued functions:** Equations from continuum mechanics, micromagnetics, and general relativity were investigated both from the analytical as well as from the numerical side.
- **Non-Euclidean Data analysis:** Various kinds of real-life data naturally live in nonlinear spaces, and data analysis techniques have to take this nonlinearity into account.
- **Linear algebra:** Problems involving subspaces can sometimes be reformulated in terms of Grassmann and Stiefel manifolds. A number of talks therefore explained the geometry of these and related spaces.

Plenty of room was left in between sessions for further discussions. After the traditional Wednesday afternoon excursion, a problem session was held in the evening. There, several participants gave short informal presentations of open problems that they considered interesting and relevant. This included better description of certain manifold cut loci, and a nonlinear variant of Rippla's theorem.

The workshop was very successful because it established various new connections between different communities, and we expect interesting new advances as a result from this in the future.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1641185, "US Junior Oberwolfach Fellows".

Workshop: Nonlinear Data: Theory and Algorithms**Table of Contents**

Xavier Pennec	
<i>Geometric Statistics for Computational Anatomy - Overview & recent advances</i>	1165
Gabriele Steidl (joint with Sebastian Neumayer, Johannes Persch)	
<i>Morphing of Manifold-Valued Images</i>	1166
Alex Bronstein (joint with O. Litany, E. Rodolà, and M. Bronstein)	
<i>Spectral Partial Shape Matching</i>	1168
Melvin Leok (joint with Evan Gawlik, James Hall, Joris Vankerschaver)	
<i>Variational Discretizations of Gauge Field Theories using Group-equivariant Interpolation</i>	1171
Philipp Harms	
<i>Analysis of Geometric Shapes in a Riemannian Setting</i>	1172
Rodolphe Sepulchre (joint with Giacomo Baggio, Augusto Ferrante)	
<i>Distances in Cones</i>	1175
Tamás Pusztai (joint with László Gránásy, Bálint Korbuly, James A. Warren, Mathis Plapp, Hervé Henry)	
<i>Orientation-Field Based Phase-Field Models for Solidification – Mathematical and Topological Challenges</i>	1176
Knut Hüper	
<i>Endpoint Geodesics on Graßmannians</i>	1179
Hanne Hardering (joint with Oliver Sander, Philipp Grohs)	
<i>Numerical Analysis for Maps into a Riemannian manifold</i>	1181
Thomas Blesgen (joint with Stephan Luckhaus and Günther Gottstein)	
<i>Simulations of Cosserat materials and Dynamic Recrystallisation</i>	1184
Laurent Younes	
<i>Elastic Distance Between Curves under the Metamorphosis Viewpoint</i> ..	1186
Edward Chien (joint with David Palmer, Paul Zhang, Justin Solomon, Heng Liu, David Bommes)	
<i>Generation of Frame Fields for Hexahedral Meshing</i>	1189
Timo Klock (joint with Zeljko Kereta, Mauro Maggioni, Valeriya Naumova)	
<i>Estimation of Nonlinear Single Index Models</i>	1192

Benedikt Wirth (joint with Behrend Heeren, Martin Rumpf) <i>The Existence Problem of Cubic Spline Interpolation on Riemannian Manifolds</i>	1195
Andreas Weinmann (joint with Kristian Bredies, Martin Holler, Martin Storath) <i>An Observation Concerning the Parallel Transport Variant of Total Generalized Variation for Manifold-Valued Data</i>	1198
Lek-Heng Lim <i>Friends of Grassmannians</i>	1202
Julie Digne (joint with Sébastien Valette, Raphaëlle Chaine, Yohann Béarzi) <i>Local and Non-Local Shape Analysis</i>	1202
Amit Boyarski (joint with Matthias Vestner, Zorah Löhner, Or Litany, Ron Slossberg, Tal Remez, Emanuele Rodolà, Alex Bronstein, Michael Bronstein, Ron Kimmel, Daniel Cremers) <i>Efficient Deformable Shape Correspondence via Kernel Matching</i>	1204
Ingo Muench <i>Structural Optimization with Cosserat Phase Field Modeling</i>	1209
Daniel Kressner (joint with Michael Steinlechner, Bart Vandereycken) <i>Riemannian Optimization with Low-Rank Tensors</i>	1213
Ralf Zimmermann <i>A Note on Rank-One Subspace Modifications and some Remarks on the Canonical Stiefel Logarithm</i>	1217
Olga Mula (joint with P. Binev, A. Cohen, J. Nichols) <i>Optimal Sensor Selection Using Reduced Models</i>	1221
Radu Ignat (joint with Luc Nguyen, Valeriy Slastikov, Arghir Zarnescu) <i>Some Uniqueness Results for Minimisers of Ginzburg–Landau Functionals</i>	1225
Max Wardetzky (joint with Henrik Schumacher) <i>Variational Convergence of Discrete Minimal Surfaces</i>	1228

Abstracts

Geometric Statistics for Computational Anatomy - Overview & recent advances

XAVIER PENNEC

The talk presents the general setting of Geometric Statistics (statistics on objects with a geometric structure) with an emphasis on the Riemannian and affine connection spaces. At the interface of geometry, statistics, image analysis and medicine, computational anatomy aims at analyzing and modeling the biological variability of the organs shapes and their dynamics at the population level. The goal is to model the mean anatomy, its normal variation, its motion / evolution and to discover morphological differences between normal and pathological groups. Since shapes and deformations live in non-linear spaces, this requires a consistent statistical framework on manifolds and Lie groups, which has motivated the development of Geometric Statistics during the last decade. The reformulation of the notion of mean as a minimization (Fréchet or Karcher mean) or as an implicit locus (exponential barycenter) has allowed the extension to Riemannian manifolds of a number of statistical methods (Gaussian distributions, tangent PCA, Mahalanobis distance, ...) and image processing algorithms (interpolation, filtering, restoration of missing data).

A recent advance in geometric statistics focuses on sequences of properly nested subspaces (flags) for generalizing PCA to Riemannian manifolds [2]. Barycentric subspaces and affine spans, defined as the (completion of the) locus of weighted means to a number of reference points, can be naturally nested by defining an ordering of the reference points. This allows the construction of forward or backward nested sequence of subspaces. However, optimizing for one subspace at a time cannot optimize the unexplained variance simultaneously for all the subspaces of the flag. In order to obtain a global criterion, PCA in Euclidean spaces was rephrased as the optimization on the flags of linear subspaces of the accumulated unexplained variance criterion. This generalizes nicely to flags of affine spans in Riemannian manifolds to realize a particularly appealing generalization of PCA on manifolds: Barycentric Subspaces Analysis (BSA).

The second part of the talk partially extends the statistical Riemannian framework to affine connection spaces and more particularly to Lie groups provided with the canonical Cartan-Schouten connection (a non-metric connection). This provides strong theoretical bases for the use of one-parameter subgroups and this allows to define bi-invariant means on Lie groups even in the absence of a bi-invariant metric [4]. This is an important achievement since bi-invariant metrics only exist on the direct product of compact and Abelian groups. Lie groups can also be considered as globally affine symmetric space structure thanks to the symmetry $s_p(q) = pq^{-1}p$. The symmetric Cartan-Schouten connection turns also to be the canonical connection associated with this symmetric structure. This notably simplifies some algorithms like parallel transport: we could prove that the pole

ladder algorithm, previously introduced in [1], is actually a third order scheme in general affine connection spaces and is even an exact scheme in one step only in symmetric spaces [3]. These properties make pole ladder a very attractive alternative to other methods for parallel transport along geodesics in general affine manifolds.

REFERENCES

- [1] M. Lorenzi and X. Pennec. Efficient Parallel Transport of Deformations in Time Series of Images: from Schild's to Pole Ladder. *Journal of Mathematical Imaging and Vision*, 50(1-2):5–17, Oct. 2013.
- [2] X. Pennec. Barycentric Subspace Analysis on Manifolds. *Annals of Statistics*, 2017. <http://arxiv.org/abs/1607.02833v2>
- [3] X. Pennec. Parallel Transport with Pole Ladder: a Third Order Scheme in Affine Connection Spaces which is Exact in Affine Symmetric Spaces. working paper or preprint, May 2018. <http://hal.archives-ouvertes.fr/hal-01799888/>
- [4] X. Pennec and V. Arsigny. Exponential Barycenters of the Canonical Cartan Connection and Invariant Means on Lie Groups. In F. Barbaresco, A. Mishra, and F. Nielsen, editors, *Matrix Information Geometry*, pages 123–168. Springer, May 2012.

Morphing of Manifold-Valued Images

GABRIELE STEIDL

(joint work with Sebastian Neumayer, Johannes Persch)

Smooth image transition, also known as image morphing, is a frequently addressed task in image processing and computer vision, and there are various approaches to tackle the problem. For example, in feature based morphing only specific features are mapped to each other and the whole deformation is then calculated by interpolation. This talk is related to a special kind of image morphing, the so-called metamorphosis introduced by Miller, Trouvé and Younes [5,6]. The metamorphosis model can be considered as an extension of the flow of diffeomorphism model and its large deformation diffeomorphic metric mapping framework in which *each* image pixel is transported along a trajectory determined by a diffeomorphism path. As an extension the metamorphosis model allows the variation of image intensities along trajectories of the pixels.

This talk builds up on a time discrete geodesic paths model by Berkels, Effland and Rumpf [1], but considers images in $L^2(\Omega, \mathcal{H})$, where $\Omega \subset \mathbb{R}^n$, $n \geq 2$, is an open, bounded connected domain with Lipschitz boundary and \mathcal{H} a finite dimensional Hadamard manifold. Hadamard manifolds are simply connected, complete Riemannian manifolds with non-positive sectional curvature. Typical examples are hyperbolic spaces and symmetric positive definite matrices with the affine invariant metric. As an important fact we will use that the distance in Hadamard spaces is jointly convex which will imply weak lower semicontinuity of certain functionals involving the distance function.

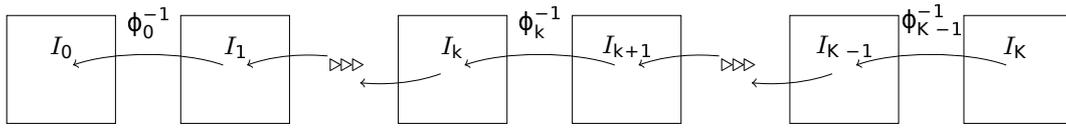


FIGURE 1. Illustration of the time discrete morphing path.

We aim in finding a minimizing sequence $\mathbf{I} = (I_1, \dots, I_{K-1}) \in (L^2(\Omega, \mathbb{R}))^{K-1}$ of the *discrete path energy*

$$\mathcal{J}(\mathbf{I}) := \sum_{k=1}^K \inf_{\varphi_k \in \mathcal{A}} \int_{\Omega} W(D\varphi_k(x)) + \gamma |D^m \varphi_k(x)|^2 dx + \frac{1}{\delta} \int_{\Omega} d_2(I_k \circ \varphi_k^{-1}, I_{k+1})^2 dx,$$

(1) subject to $I_0 = T, I_K = R,$

where $\delta, \gamma > 0$, d_2 denotes the distance in $L^2(\Omega, \mathcal{H})$,

$$\mathcal{A} := \{\varphi \in (W^{m,2}(\Omega))^n : \det(D\varphi) \geq \epsilon, \varphi(x) = x \text{ for } x \in \partial\Omega\}, \quad m > 1 + \frac{n}{2}$$

is an admissible set of deformations and the function W has to satisfy certain properties. A particular choice of W is given by the linearized elastic potential. We prove that a minimizer of (1) exists.

Dealing with digital images we have to introduce a space discrete model. We establish a finite difference model on a staggered grid together with a multiscale strategy. We have used this discretization already for gray-value images in [2]. For finding a minimizer, we also propose an alternating algorithm fixing either the deformation or the image sequence:

- i) For a fixed image sequence, we have to solve certain registration problems for *manifold-valued images* in parallel to get a sequence $(\varphi_1, \dots, \varphi_K)$ of deformations. Necessary interpolations were performed via Karcher means computation.
- ii) For a fixed deformation sequence, we need to find a minimizing image sequence (I_1, \dots, I_{K-1}) of

$$\sum_{k=1}^K d_2^2(I_k \circ \varphi_k^{-1}, I_{k+1}) \quad \text{subject to } I_0 = T, I_K = R$$

where d_2 denotes the distance in $L^2(\Omega, \mathcal{H})$.

In the second part of the talk we turn to real-valued image processing. In certain applications it makes sense to account for qualitative prior image information to improve the image reconstruction. Typical examples are tomographic imaging problems with sparsely or limited angle sampled sinogram data. One possibility to incorporate a reference images into the reconstruction process is to take its deformation towards the image of interest, which is only indirectly given by measurements, into account. We enlarge the above ideas to the solution of inverse problems with a given reference image. To this end, we combine the discrete

geodesic path model with a variational model, actually the L^2 - TV model. We minimize

$$\mathcal{J}(\mathbf{I}, \varphi) := \mathcal{F}(\mathbf{I}, \varphi) + \beta \mathcal{E}(I_0; B) \quad \text{subject to} \quad I_K = R,$$

where $\beta > 0$,

$$\mathcal{E}(I; B) := \|AI - B\|_Y^2 + \alpha TV(I), \quad \alpha \geq 0.$$

and

$$\mathcal{F}(\mathbf{I}, \varphi) := \sum_{k=0}^{K-1} \int_{\Omega} W(D\varphi_k) + \nu |D^m \varphi_k|^2 + |I_k \circ \varphi_k^{-1} - I_{k+1}|^2 dx.$$

We prove that the space continuous model has a minimizer and propose a minimization procedure which alternates over the involved sequences of diffeomorphism and images. The minimization with respect to the image sequence exploits recent algorithms from convex analysis to minimize the L^2 - TV functional.

For further information we refer to [2–4].

REFERENCES

- [1] B. Berkels, A. Effland, and M. Rumpf. Time discrete geodesic paths in the space of images. *SIAM Journal on Imaging Sciences* **8(3)** (2015), 1457–1488.
- [2] J. Persch, F. Pierre, and G. Steidl. Exemplar-based face colorization using image morphing. *Journal of Imaging* **3(4)** (2017) :ArtNum 48.
- [3] S. Neumayer, J. Persch, and G. Steidl. Morphing of manifold-valued images inspired by discrete geodesics in image spaces. ArXiv Preprint 2017, *SIAM Journal on Imaging Sciences*, submitted.
- [4] S. Neumayer, J. Persch, F. Pierre, and G. Steidl. Regularization of inverse problems via time discrete geodesics in image spaces. *ArXiv Preprint*, 2018.
- [5] A. Trounev and L. Younes. Local geometry of deformable templates. *SIAM Journal of Mathematical Analysis* **37(2)** (2005), 17–59.
- [6] A. Trounev and L. Younes. Metamorphoses through Lie group action. *Foundations in Computational Mathematics* **5(2)** (2005), 173–198.

Spectral Partial Shape Matching

ALEX BRONSTEIN

(joint work with O. Litany, E. Rodolà, and M. Bronstein)

The main drawback of partial functional maps [4] and the follow-up works [1, 2] is their explicit model of the part, requiring a somewhat cumbersome solver alternating between optimization in the spatial domain (over the part indicator function) and in the spectral domain (over the correspondence matrix). Furthermore, the complexity of the spatial domain optimization depends on the number of mesh vertices and scales poorly. One of the main contributions of our paper is a simple observation allowing to formulate the partial functional maps problem *entirely in the spectral domain*. Our method bears resemblance to joint approximate diagonalization.

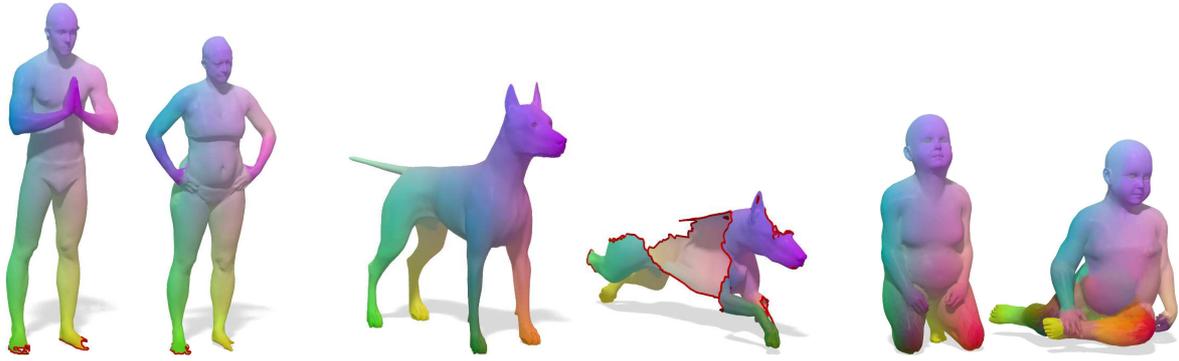


FIGURE 1. Examples of dense correspondence computed with our method on real 3D scans (left pair, the areas of contact are glued together), missing parts (middle) and strong topological artifacts (right, touching parts are glued together). Corresponding points are encoded with the same color.

A key feature of partial functional maps lies in their *spatially localized* behavior, that is, it describes a map $T : L^2(\mathcal{M}) \rightarrow L^2(\mathbb{N})$ that is supported on some region $\mathbb{N}' \subseteq \mathbb{N}$ of the full model, meaning that for all $y \in \mathbb{N} \setminus \mathbb{N}'$ the approximate equality $(Tf)(y) \approx 0$ holds for any $f \in L^2(\mathcal{M})$. This can be easily seen by noting that the image of \mathbf{A} under \mathbf{C} must be localized to the region indicated by v in order for the data term $\|\mathbf{CA} - \mathcal{B}(v)\|$ to reach a minimum; in other words, the functional map \mathbf{C} must localize the correspondence.

This localization property comes at the price of modeling the region $\mathbb{N}' \subseteq \mathbb{N}$ explicitly. Here we propose to absorb the spatial mask into a new basis $\{\hat{\psi}_j\}$ for $L^2(\mathbb{N})$; in doing so, we dispose of the explicit part v and obtain a simpler optimization problem, as we elucidate in the following.

Assume \mathbf{C}, v describe a partial functional map, such that $\mathbf{CA} = \mathcal{B}(v)$ holds approximately, and consider two functions $f \in L^2(\mathcal{M}), g \in L^2(\mathbb{N})$ whose spectral representations are columns of \mathbf{A} and \mathcal{B} respectively. In the spatial domain, the equality becomes

$$(1) \quad \sum_{ij} \langle f, \phi_i \rangle_{\mathcal{M}} c_{ji} \psi_j = \sum_{i=1}^k \langle v \cdot g, \psi_i \rangle_{\mathbb{N}} \psi_i \approx v \cdot g,$$

where the approximation is due to truncation to the first k terms. By defining a new basis $\hat{\psi}_i = \sum_{j=1}^k c_{ji} \psi_j$, we get to

$$(2) \quad \sum_{ij} \langle f, \phi_i \rangle_{\mathcal{M}} \hat{\psi}_j \approx v \cdot g,$$

in other words, the modified basis $\{\hat{\psi}_j\}$ induces the sought localization. Importantly, in order for (2) to hold for general f and g , the new basis functions themselves must be localized, *i.e.*, $\hat{\psi}_i = v \cdot \psi_i$ for all i .

Using the fact that orthogonal \mathbf{C} implies orthogonal $\{\hat{\psi}_j\}$, we can phrase (2) in the spectral domain as:

$$(3) \quad \mathbf{A} \approx \mathbf{C}\mathbf{T}\mathcal{B}(v) = \mathbf{C}\mathbf{T}\mathcal{B};$$

in the last equality, we absorbed the indicator function v into the new basis functions $\{\hat{\psi}_j\}$.

We consider the following manifold optimization problem:

$$(4) \quad \min_{\mathbf{Q} \in S(k,r)} \text{off}(\mathbf{Q}^\top \mathbf{A}_\mathbb{N} \mathbf{Q}) + \mu \|\mathbf{A}_r - \mathbf{Q}^\top \mathcal{B}\|_{2,1},$$

where $S(k, r)$ is the Stiefel manifold of orthogonal $k \times r$ matrices (ortho-projections), and $\mathbf{A}_r = \mathbf{W}_r \mathbf{A}$ with $\mathbf{W}_r = (\mathbf{I}_{r \times r} \mathbf{0}_{r \times k-r})$ denotes the $r \times k$ matrix containing the first r rows of \mathbf{A} . The value of r is directly related to the rank of the partial functional map \mathbf{C} , and can be estimated simply from the area ratio θ , or optimized for explicitly by solving (4) for a range of r 's. The rank r and the orthogonality of \mathbf{Q} act as partiality priors, since they are related to the underlying map being area-preserving [3, 4].

The optimization problem (4) models partial correspondence as the search for a new basis that is localized to a latent part of the full shape. In this view, the matrix \mathbf{Q} is not regarded as a functional map between shapes, but rather as a matrix of transformation coefficients for the basis (the off-diagonal regularity term ensures that the transformation is smooth). This interpretation will allow us to tackle part-to-part settings as a simple modification to (4).

The first r functions $\{\hat{\psi}_1, \dots, \hat{\psi}_r\}$ of the new orthogonal basis $\hat{\psi}_i = \sum_{j=1}^k q_{ji} \psi_j$ obtained as the result of such a transformation would be approximately orthogonal to $\{\phi_1, \dots, \phi_r\}$ under the functional correspondence,

$$\langle T\phi_i, \hat{\psi}_j \rangle_{\mathbb{N}} \approx \delta_{ij}; \quad i, j = 1, \dots, r.$$

It is important to remark that, while the correct partial correspondence is a solution to our problem by Eq. (1–3), this is not necessarily unique as it directly depends on the input data. Not all such optima are localized to the correct region, and some might even have global support. The choice of the input corresponding functions $\{f_i, g_i\}$ ultimately determines the quality of the localization. In practice, it is enough to employ dense descriptor fields that are sufficiently similar on the corresponding regions in order to drive the optimization to the correct solution.

REFERENCES

- [1] L. Cosmo, E. Rodolà, J. Masci, A. Torsello, and M. Bronstein. Matching deformable objects in clutter. In *Proc. 3DV*, 2016.
- [2] Or Litany, Emanuele Rodolà, Alex M. Bronstein, Michael M. Bronstein, and Daniel Cremers. Non-rigid puzzles. *Computer Graphics Forum*, 35(5):135–143, August 2016.
- [3] M. Ovsjanikov, M. Ben-Chen, J. Solomon, A. Butscher, and L. Guibas. Functional maps: a flexible representation of maps between shapes. *ACM Trans. Graph.*, 31(4):30:1–30:11, July 2012.
- [4] E. Rodolà, L. Cosmo, M. M. Bronstein, A. Torsello, and D. Cremers. Partial functional correspondence. *Computer Graphics Forum*, 2016.

Variational Discretizations of Gauge Field Theories using Group-equivariant Interpolation

MELVIN LEOK

(joint work with Evan Gawlik, James Hall, Joris Vankerschaver)

Variational integrators [4] are geometric structure-preserving numerical methods that preserve the symplectic structure, satisfy a discrete Noether's theorem, and exhibit excellent long-time energy stability properties. The exact discrete Lagrangian,

$$L_d^E(q_0, q_1; h) \equiv \underset{\substack{q \in C^2([0, h], Q) \\ q(0) = q_0, q(h) = q_1}}{\text{ext}} \int_0^h L(q(t), \dot{q}(t)) dt,$$

arises from Jacobi's solution of the Hamilton-Jacobi equation, and it generates the exact flow of a Lagrangian system. By approximating the exact discrete Lagrangian using an appropriate choice of interpolation space and quadrature rule, we obtain a systematic approach for constructing variational integrators. It was established in [2, 3] that the convergence rates of such variational integrators are related to the best approximation properties of the interpolation space.

One particularly important class of Lagrangian field theories is that of *gauge field theories*, where the Lagrangian density is equivariant under a *gauge symmetry*, which is a local symmetry action. Examples include electromagnetism, Yang-Mills, and general relativity. A consequence of gauge symmetries is that when the field theory is formulated as an initial-value problem, the evolution of the field theory is not uniquely specified by the initial conditions, that is to say that they are underdetermined. In particular, the fields can be decomposed into *dynamic fields*, whose evolution is described by well-posed equations, and *kinematic fields* which have no physical significance. In relativity, the former are the metric and extrinsic curvature on a spatial hypersurface, and the latter are the lapse and shift. Besides the indeterminacy in the evolution equations, there are initial-value constraints (typically elliptic) on the Cauchy data. *Noether's first theorem* applied to the nontrivial rigid subgroup of the gauge group implies that there exists a *Noether current* that obeys a continuity equation, and integrating this over a Cauchy surface yields a conserved quantity called a *Noether charge*. What is more interesting for gauge field theories is *Noether's second theorem*, which can recover some of the equations of motion automatically from the gauge symmetry, and this is particularly important for covariant field theories, such as general relativity.

Many gauge field theories can be formulated variationally using a multisymplectic Lagrangian formulation, and we will present a characterization of the exact generating functionals [5] that generate the multisymplectic relation. By discretizing these using group-equivariant spacetime finite element spaces, we obtain methods that exhibit a discrete multimomentum conservation law. We will then briefly describe an approach for constructing group-equivariant interpolation spaces [1] that take values in the space of Lorentzian metrics that can be efficiently computed using a generalized polar decomposition. This relies on the observation that the

space of Lorentzian metrics is a symmetric space, which is endowed with a generalized polar decomposition. This decomposition induces the Cartan decomposition at the Lie algebra level,

$$\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k},$$

and the following diagram commutes,

$$\begin{array}{ccccc}
 \mathfrak{k} & \xrightarrow{\text{exp}} & G^\sigma & & \\
 \downarrow \iota & & \searrow \iota & & \\
 \mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k} & \xrightarrow{\text{exp}} & G & & \\
 \uparrow \iota & & \nearrow \iota & & \\
 \mathfrak{p} & \xrightarrow{\text{exp}} & G_\sigma & \xrightarrow{\psi} & G/G^\sigma \xrightarrow{\bar{\varphi}} \mathcal{S} \\
 & & \searrow \psi & & \downarrow \pi \\
 & & & & \mathcal{S}
 \end{array}$$

Critically, the maps along its bottom row are local diffeomorphisms, which allow the symmetric space \mathcal{S} to be identified with the Lie triple system \mathfrak{p} , which is a linear space. This local diffeomorphism allows a symmetric space valued interpolant to be constructed from a linear space valued interpolant.

The goal is to eventually apply this to the construction of variational discretizations of general relativity, which is a second-order gauge field theory whose configuration manifold is the space of Lorentzian metrics.

REFERENCES

- [1] Gawlik, E.S., Leok, M.: Interpolation on symmetric spaces via the generalized polar decomposition. *Found. Comput. Math.* (2016). (Online First, doi:10.1007/s10208-017-9353-0)
- [2] Hall, J., Leok, M.: Spectral variational integrators. *Numer. Math.* **130**(4), 681–740 (2015)
- [3] Hall, J., Leok, M.: Lie group spectral variational integrators. *Found. Comput. Math.* **17**(1), 199–257 (2017)
- [4] Marsden, J., West, M.: Discrete mechanics and variational integrators. *Acta Numer.* **10**, 317–514 (2001)
- [5] Vankerschaver, J., Liao, C., Leok, M.: Generating functionals and Lagrangian partial differential equations. *J. Math. Phys.* **54**(8), 082901 (22 pages) (2013)

Analysis of Geometric Shapes in a Riemannian Setting

PHILIPP HARMS

This is a summary of some results and open questions in infinite-dimensional Riemannian shape analysis. Results which are cited in the survey article [4] are not cited here.

1. GEOMETRIC DATA

A key paradigm of shape analysis and, more generally, functional data analysis is that geometric data should be seen as *infinite-dimensional* and *nonlinear*. This should be contrasted with the alternative point of view, which is sometimes taken in numerical applications, where geometric data is represented in coordinates and

seen as an element of a high-dimensional linear space. Riemannian geometry has established itself as a successful tool for dealing with both the non-linearity and infinite-dimensionality of geometric data.

2. SHAPE SPACES

A first step towards a Riemannian setting for shape analysis is the definition of shape manifolds. This is achieved in great generality in the following theorem.

Theorem. Let M and N be manifolds, M compact. Then the spaces $C^\infty(M, N)$, $\text{Emb}(M, N)$ and $\text{Diff}(M)$ are Fréchet manifolds, and the projection from $\text{Emb}(M, N)$ to the quotient space $B(M, N) := \text{Emb}(M, N)/\text{Diff}(M)$ defines a principal fiber bundle of Fréchet manifolds.

One may consider larger spaces of mappings such as immersions or mappings of finite degrees of smoothness, but this leads to singularities in the quotient space. Note that the quotient space $B(M, N)$ is nonlinear even if N is linear.

3. RIEMANNIAN METRICS

Riemannian geometry formalizes an intuitive notion of similarity between shapes: two shapes are regarded as similar if they differ by a “small” deformation. A deformation between two embeddings f_0 and f_1 is a smooth path f in $\text{Emb}(M, N)$ which connects f_0 to f_1 . The distance between f_0 and f_1 is the infimal length of all such deformations, i.e.,

$$\text{dist}(f_0, f_1) = \inf_f \int_0^1 \|\dot{f}(t)\| dt.$$

Here $\dot{f}(t)$ belongs to the tangent space at $f(t)$, which consists of all vector fields along f , and the length of $\dot{f}(t)$ is measured using a Riemannian metric. The following is a rich and widely applicable class of Sobolev-type Riemannian metrics on $\text{Emb}(M, N)$.

Definition. Let M be a compact manifold and (N, \bar{g}) a Riemannian manifold. Then the H^ℓ metric at $f \in \text{Emb}(M, N)$ and $h, k \in T_f \text{Emb}(M, N)$ is defined as

$$G_f(h, k) = \int_M \bar{g}((1 + \Delta^g)^\ell h, k) \text{vol}(g),$$

where $g = f^*\bar{g}$.

It is crucial that $g = f^*\bar{g}$ is not fixed but depends on the foot point f . This implies that the metric is $\text{Diff}(M)$ -invariant, i.e., $G_{f \circ \phi}(h \circ \phi, k \circ \phi) = G_f(h, k)$. This in turn implies the existence of a unique Riemannian metric on $B(M, N)$ such that the projection from $\text{Emb}(M, N)$ to $B(M, N)$ is a Riemannian submersion.

4. GEODESIC DISTANCE

The geodesic distance functional is called non-degenerate if it separates points, degenerate otherwise, and vanishing if it is zero everywhere.

Theorem. The geodesic distance functional of the Sobolev metric of order ℓ is degenerate on $\text{Emb}(M, N)$ for each $\ell \leq \frac{1}{2}$, vanishes on $\text{Diff}(M)$ for each $\ell \leq \frac{1}{2}$, and is non-degenerate on $\text{Emb}(M, N)$ and $\text{Diff}(M)$ for each $\ell \geq 1$.

These statements follow from a couple of simple observations:

- The H^0 distance on $B([0, 1], \mathbb{R}^2)$ is degenerate: a straight line can be deformed into a shifted straight line at almost no H^0 cost if the intermediate shapes have many steep zig zags.
- The $H^{\frac{1}{2}}$ distance on $\text{Diff}(\mathbb{R})$ is degenerate: the identity can be deformed into a shift by 1 at almost no $H^{\frac{1}{2}}$ cost using a shock wave traveling at constant speed with very steep slope.
- The H^1 distance in $B([0, 1], \mathbb{R}^2)$ is non-degenerate: the H^1 energy of a deformation is bounded from below by the area swept out.

The case $\frac{1}{2} < \ell < 1$ on general spaces of embeddings and diffeomorphisms remains open.

5. GEODESICS

The geodesic equations of Sobolev metrics are in many cases well-known PDEs from mathematical physics and, in particular, hydrodynamics. For example, the SQG, Euler, Burger, KdV, mCLM, Camassa–Holm, and Hunter–Saxton equations are of this type.

Theorem. [3, 5, 6] Local well-posedness of the initial value problem for geodesics holds on $\text{Emb}(S^1, \mathbb{R}^2)$ for each $\ell \geq \frac{1}{2}$, on $\text{Diff}(\mathbb{R}^m)$ for each $\ell \geq \frac{1}{2}$, and on $\text{Emb}(M, N)$ for each $\ell \in \mathbb{N}_{\geq 1}$. Global well-posedness of the initial value problem for geodesics holds on $\text{Emb}(S^1, \mathbb{R}^2)$ for each $\ell \in \mathbb{N}_{\geq 2}$ and on $\text{Diff}(\mathbb{R}^m)$ for each $\ell > \frac{m}{2} + 1$. Existence of solutions of the boundary value problem for geodesics on the metric completion of the respective spaces holds on $\text{Emb}(S^1, \mathbb{R}^2)$ for each $\ell \in \mathbb{N}_{\geq 2}$ and on $\text{Diff}(\mathbb{R}^m)$ for each $\ell > \frac{m}{2} + 1$.

Here well-posedness is understood in the sense of Hadamard. The method for establishing local well-posedness goes back to Ebin and Marsden: one verifies that the geodesic spray extends to a smooth vector field on a suitable Sobolev completion of $T\text{Emb}(M, N)$. The cases which are not mentioned in the theorem remain open. This includes foremost non-natural orders ℓ . Some aspects of the boundary value problem are discussed in the following section.

6. NUMERICS

For some H^1 metrics the geodesic equation on $B(S^1, \mathbb{R}^2)$ can be solved explicitly using isometries to spheres or solitons [1], but not for general metrics. For H^2 metrics we implemented the initial and boundary value problems using spline

discretizations [2]. The code is available on github.com/h2metrics. Solving a boundary value problem with a spatial resolution of 100 points and a time resolution of 10 points takes under 5 seconds. Convergence is guaranteed by the following result.

Theorem. [2] The closure of $B(S^1, \mathbb{R}^2)$ with respect to the geodesic distance of the H^2 metric is a complete length space. Discrete minimizers converge weakly to minimizers, up to the selection of a subsequence.

An analysis of the convergence rate and an extension to more general shape spaces remain open.

7. APPLICATIONS AND EXTENSIONS

The above-mentioned results form a basis for statistical and machine learning frameworks. Conceptually, they extend to many other shape spaces such as spaces of Riemannian metrics or densities.

REFERENCES

- [1] Martin Bauer, Martins Bruveris, Philipp Harms, and Peter W Michor. Soliton solutions for the elastic metric on spaces of curves. *Discrete Contin. Dyn. Sys.*, 38(3), 2018.
- [2] Martin Bauer, Martins Bruveris, Philipp Harms, and Jakob Møller Andersen. A numerical framework for Sobolev metrics on the space of curves. *SIAM J. Imaging Sci.*, 10(1):47–73, 2017.
- [3] Martin Bauer, Martins Bruveris, and Boris Kolev. Fractional sobolev metrics on spaces of immersed curves, 2017. [arXiv:1703.03323](https://arxiv.org/abs/1703.03323).
- [4] Martin Bauer, Martins Bruveris, and Peter W. Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *J. Math. Imaging Vision*, 50(1-2):60–97, 2014.
- [5] Martins Bruveris, Peter W. Michor, and David Mumford. Geodesic completeness for Sobolev metrics on the space of immersed plane curves. *Forum Math. Sigma*, 2:e19, 38, 2014.
- [6] Martins Bruveris and François-Xavier Vialard. On completeness of groups of diffeomorphisms, 2017. [arXiv:1403.2089](https://arxiv.org/abs/1403.2089).

Distances in Cones

RODOLPHE SEPULCHRE

(joint work with Giacomo Baggio, Augusto Ferrante)

This talk draws from recent results presented in [1]. The main question is how to select a distance to process a specific type of nonlinear data. The design of distances with the aim of solving computational engineering problems is a rich topic because of the interplay between mathematical, modelling, and computational considerations. It has received far less attention for *intensities* nonlinear data than for *spherical* nonlinear data. The paper presents new results for spectral densities by acknowledging that those data lie in a cone. It generalises Thompson and Hilbert metric to the space of spectral densities. The resulting complete metric space has the differentiable structure of a Finsler manifold with explicit geodesics. The resulting distances are filtering invariant, can be computed efficiently, and

admit geodesic paths that preserve rationality; these are properties of fundamental importance in many engineering applications.

REFERENCES

- [1] Giacomo Baggio, Augusto Ferrante, Rodolphe Sepulchre, *Conal Distances Between Rational Spectral Densities*, arXiv:1708.02818, 2017.

Orientation-Field Based Phase-Field Models for Solidification – Mathematical and Topological Challenges

TAMÁS PUSZTAI

(joint work with László Gránásy, Bálint Korbuly, James A. Warren, Mathis Plapp, Hervé Henry)

INTRODUCTION

The phase-field method is a versatile tool for modelling pattern formation, such as microstructure evolution during solidification [1–7]. Depending on the problem addressed, the material is described with one or more continuous order parameters and other fields. The most important one is the phase-field, $\phi(\mathbf{r}, t)$, which represents the local phase state of the matter ($\phi=0$ and 1 in the bulk liquid and solid, respectively, and has intermediate values in phase boundaries), but often other fields, such as the chemical composition is also required. In the variational approach of the phase-field method, we start from a thermodynamic potential such as the total free energy of the system, which is expressed as the functional of the selected fields. The time evolution of these fields are then governed by the minimisation of the total free energy, a principle which translates to partial differential equations that can be derived from the free energy functional.

A very important class of materials are the polycrystals, which consist of small single-crystal grains with different crystallographic orientation. As opposed to assigning a separate phase field to each grain [4, 8–11], an efficient way of modelling these materials is introducing a new field, the orientation field, which represents the local crystallographic orientation [12–20]. One must, however, be careful when working with the orientation field. Due to its special properties, problems that are unknown with the other fields appear.

The first difficulty is already with the representation of orientations in 2D and 3D. In contrast to the other fields, where going from 2D to 3D is only a matter of using the higher dimensional form of the spatial differential operators, the orientation fields are fundamentally different. In the following two chapters I will give an overview of the special mathematical/physical problems that we, as material scientists, face when working with orientation fields in 2D and 3D.

ORIENTATION FIELD IN 2D

Given a reference orientation of a system, any of its possible orientations can be defined as the rotation that transforms the system to this orientation from the reference one. In 2D, rotations can be represented by their angle, thus the orientation field is a simple scalar field $\theta(\mathbf{r}, t)$. Since orientations are equivalent up to rotations by 2π , orientations can be uniquely defined by a cyclic field $\theta \in [0, 2\pi)$, the two ends of the interval being equated [12].

Special requirements and properties of the θ field:

- Rotational invariance requires that the free energy of the system does not depend directly on θ , only via its gradient $\nabla\theta$ [12].
- The usual $(\nabla\theta)^2$ term leads to grain boundaries that spread out. Localised grain boundaries require either a term proportional to $|\nabla\theta|$ [12, 13], or a singular $g(\phi)$ multiplier of the term $(\nabla\theta)^2$ [20].
- The gradient term requires the calculation of orientational differences. For any pair of orientations two differences exist. Physically, the one with a smaller absolute value should be considered [16]. This can be easily handled in custom finite difference codes, but rules out the use implicit or spectral codes or ready-made solvers.
- Between any two orientations, two different continuous paths exist. This means that between any two grains, there are two different grain boundary solutions [16, 21].
- The possible values of the orientation field, i.e., its order parameter space is the circumference of the unit circle, which is not simply connected. This means, that the two different solutions mentioned above cannot transform to one another continuously [16, 21].
- If these two different solutions meet on the same grain boundary, a topological defect must exist between them. We consider them unphysical and therefore to be eliminated. To achieve this, two new orientation fields are proposed which extend the order parameter space from the unit circle (a) to the surface of the 3D unit sphere, or (b) to the whole 2D plane, but with the addition of a new potential [21].
- Crystallographic symmetries are easily handled. In case of a k -fold symmetry, the unique values of the orientation field are limited to $\theta \in [0, 2\pi/k)$ [22]. This does not affect the topological properties of the field.

ORIENTATION FIELD IN 3D

Orientations in 3D can be represented in different ways. The Euler angle description is problematic because of its singularity at the poles. Other possibilities include the rotation matrix representation [23] and the Rodrigues vectors. Our approach, however, was the use of quaternions [24], which provides a simple and elegant way to extend the orientation field to 3D.

- When using the 4D unit sphere representation of quaternions, the orientational difference between two orientations is the distance between the respective two points on the 4D unit sphere through its 3D surface. For small angles, it can be well approximated by the Euclidean distance of these points. This approximation gives a particularly simple formulation of the free energy, similar to the 2D case [24].
- Handling symmetries is troublesome. Supposing e.g. the rotational symmetry group of a cube, all 24 equivalent orientations has to be considered when calculating the orientational difference. This slows down the simulations significantly [24].
- Two order parameter space of the 3D orientations is not simply connected, either. The two different grain boundary solutions and the associated topological defects seen in 2D also exist in 3D [21].
- Due to the higher dimensional order parameter space and the higher dimensional real space the topological properties are more complicated. Are there other types of topological defects in these systems?

The above points give a list of problems and challenges that had to be solved to use the orientation field with phase-field models successfully. Most of them are already handled at a certain level - in a way physicists would/could solve them. It might, however, be beneficial to see these topics through a mathematician's eye. It may turn out that better, more elegant or more efficient solutions to these problems exist. In addition to show a practical use of the orientation field model in materials science, my goal is to stimulate discussions on these topics.

REFERENCES

- [1] W. J. Boettinger, J. A. Warren, C. Beckermann, and A. Karma. *Phase-Field Simulation of Solidification*, Annual Review of Materials Research, **32** (2002), 163–194.
- [2] Tamás Pusztai, György Tegze, Gyula I. Tóth, László Környei, Gurvinder Bansal, Zhongyun Fan, and László Gránásy, *Phase-field approach to polycrystalline solidification including heterogeneous and homogeneous nucleation*, Journal of Physics: Condensed Matter, **20** (2008), 404205.
- [3] M. Asta, C. Beckermann, A. Karma, W. Kurz, R. Napolitano, M. Plapp, G. Purdy, M. Rappaz, and R. Trivedi, *Solidification microstructures and solid-state parallels: Recent developments, future directions*. Acta Materialia, **57** (2009), 941–971.
- [4] Ingo Steinbach, *Phase-field models in materials science*, Modelling and Simulation in Materials Science and Engineering, **17** (2009), 073001.
- [5] Nikolas Provatas and Ken Elder, *Phase-Field Methods in Materials Science and Engineering*, Wiley-VCH, Berlin, 2010.
- [6] Ingo Steinbach. *Why Solidification? Why Phase-Field?* JOM, **65** (2013), 1096–1102.
- [7] László Gránásy, László Rátkai, Attila Szállás, Bálint Korbuly, Gyula I. Tóth, László Környei, and Tamás Pusztai, *Phase-Field Modeling of Polycrystalline Solidification: From Needle Crystals to Spherulites—A Review*, Metallurgical and Materials Transactions A, **45** (2014), 1694–1719.
- [8] Long-Qing Chen and Wei Yang, *Computer simulation of the domain dynamics of a quenched system with a large number of nonconserved order parameters: The grain-growth kinetics*, Physical Review B, **50** (1994), 15752–15756.

- [9] N. Moelans, B. Blanpain, and P. Wollants, *Quantitative Phase-Field Approach for Simulating Grain Growth in Anisotropic Systems with Arbitrary Inclination and Misorientation Dependence*, Physical Review Letters, **101** (2008), 025502.
- [10] Reza Darvishi Kamachali and Ingo Steinbach, *3-D phase-field simulation of grain growth: Topological analysis versus mean-field approximations*, Acta Materialia, **60** (2012), 2719–2728.
- [11] Gyula I. Tóth, Tamás Pusztai, and László Gránásy, *Consistent multiphase-field theory for interface driven multidomain dynamics*, Physical Review B, **92** (20125), 184105.
- [12] James A. Warren, W. Craig Carter, and Ryo Kobayashi, *A phase field model of the impingement of solidifying particles*, Physica A: Statistical Mechanics and its Applications, **261** (1998), 159–166.
- [13] Ryo Kobayashi, James A. Warren, and W. Craig Carter, *A continuum model of grain boundaries*, Physica D: Nonlinear Phenomena, **140** (2000), 141–150.
- [14] James A. Warren, Ryo Kobayashi, and W. Craig Carter, *Modeling grain boundaries using a phase-field technique*, Journal of Crystal Growth, **211** (2000), 18–20.
- [15] László Gránásy, Tamás Börzsönyi, and Tamás Pusztai, *Nucleation and Bulk Crystallization in Binary Phase Field Theory*, Physical Review Letters, **88** (2002), 206105.
- [16] James A. Warren, Ryo Kobayashi, Alexander E. Lobkovsky, and W. Craig Carter, *Extending phase field models of solidification to polycrystalline materials*, Acta Materialia, **51** (2003), 6035–6058.
- [17] László Gránásy, Tamás Pusztai, James A. Warren, Jack F. Douglas, Tamás Börzsönyi, and Vincent Ferreiro, *Growth of 'dizzy dendrites' in a random field of foreign particles*, Nature Materials, **2** (2003), 92–96.
- [18] László Gránásy, Tamás Pusztai, Tamás Börzsönyi, James A. Warren, and Jack F. Douglas, *A general mechanism of polycrystalline growth*, Nature Materials, **3** (2004), 645–650.
- [19] László Gránásy, Tamás Pusztai, György Tegze, James A. Warren, and Jack F. Douglas, *Growth and form of spherulites*, Physical Review E, **72** (2005), 011605.
- [20] Hervé Henry, Jesper Mellenthin, and Mathis Plapp, *Orientation-field model for polycrystalline solidification with a singular coupling between order and orientation*, Phys. Rev. B, **86** (2012), 054117.
- [21] Bálint Korbuly, Mathis Plapp, Hervé Henry, James A. Warren, László Gránásy, and Tamás Pusztai, *Topological defects in two-dimensional orientation-field models for grain growth*, Phys. Rev. E, **96** (2017), 052802.
- [22] Ryo Kobayashi, James A. Warren, and W. Craig Carter, *Vector-valued phase field model for crystallization and grain boundary formation*, Physica D: Nonlinear Phenomena, **119** (1998), 415–423.
- [23] Ryo Kobayashi and James A. Warren, *Modeling the formation and dynamics of polycrystals in 3D*, Physica A: Statistical Mechanics and its Applications, **356** (2005), 127–132.
- [24] Tamás Pusztai, G. Bortel, and László Gránásy, *Phase field theory of polycrystalline solidification in three dimensions*, Europhysics Letters, **71** (2005), 131–137.

Endpoint Geodesics on Graßmannians

KNUT HÜPER

In recent years an extrinsic point of view, here we mean embedding Riemannian manifolds isometrically into some Euclidean space, has offered the opportunity to derive explicit formulas for certain differential geometric constructions, such as e.g. geodesics and parallel transport of tensor fields. The approach is based on considering isometrically embedded submanifolds as rigid bodies rolling without twist or slip along an affine tangent space. The rolling map considered as a smooth

curve in the isometry group of the embedding space then allows for the derivation of the above explicit formulas. This has turned out to be pretty useful for computer implementations in connection with optimization, interpolation, computer graphics and computer vision, and control theoretic applications, to mention just a few. One reason why this approach turns out to be rather attractive in engineering applications in particular, is that often evaluating closed formulas seems more attractive than a purely numerical approach. Last but not least a closed formula might serve as an ansatz for a new competitive numerical algorithm as well. It seems needless to mention that working with explicit coordinates instead of equivalence classes is often much better suited to numerical implementations.

After quickly recalling the concept of rolling maps, cf. [4], we focus on the Graßmann manifold following our earlier approach [3]. We consider the Graßmannian $G_{n,k}$ of all k -planes in \mathbb{R}^n as a Riemannian submanifold of the vector space of $n \times n$ -symmetric matrices Sym_n by identifying a k -plane with its orthogonal projection operator.

$$G_{n,k} := \{P \in Sym_n \mid P^2 = P, \text{rk}(P) = k\}.$$

Whereas closed formulas for geodesics on $G_{n,k}$ through P emanating in direction $\dot{P} \in T_P G_{n,k}$ are well known to be of the form, cf. [2],

$$\begin{aligned} \gamma: \mathbb{R} &\rightarrow G_{n,k}, \\ t &\mapsto e^{t[\dot{P}, P]} P e^{-t[\dot{P}, P]}, \end{aligned}$$

with $[\cdot, \cdot]$ denoting the matrix commutator, more recently a formula for so called endpoint geodesics has been found. Consider two elements $P, Q \in G_{n,k}$ being not conjugate points. Then the unique geodesic connecting P with Q is described (up to reparametrization) by the equation

$$(1) \quad Q = e^X P e^{-X}.$$

Clearly, solving for an $X \in \mathfrak{so}_n$ in (1) could be hard, as it is not unique. However, it becomes unique if we restrict our search space to the subspace $[T_P G_{n,k}, P] \subset \mathfrak{so}_n$. A closed formula for computing such an X was presented in [1]:

$$(2) \quad X = \frac{1}{2} \log(I - 2Q)(I - 2P).$$

As usual \log denotes the matrix logarithm, being unique under the above assumptions. It is remarkable to verify here that formula (2) is related to the fact that the Graßmann manifold $G_{n,k}$ can be embedded into the orthogonal group O_n via

$$\begin{aligned} \mu: G_{n,k} &\rightarrow O_n, \\ P &\mapsto I - 2P, \end{aligned}$$

its image being the isospectral manifold

$$\mu(G_{n,k}) = \left\{ X \in O_n \mid X = Q \begin{bmatrix} -I_k & \\ & I_{n-k} \end{bmatrix} Q^\top, Q \in SO_n \right\}.$$

A result interesting in its own right, not only because

$$\dim O_n = \binom{n}{2} \leq \dim Sym_n = \binom{n+1}{2}.$$

For k even, we get an embedding even into $SO_n = O_n^+$ rather than into the nonconnected component O_n^- of O_n .

Moreover, in this talk we give a geometric interpretation of the formula

$$(3) \quad (I - 2P)e^{-[\dot{P}, P]}(I - 2P) = e^{+[\dot{P}, P]}$$

for all $P \in G_{n,k}$ and all $\dot{P} \in T_P G_{n,k}$ in terms of reflections on the tangent space $T_P G_{n,k} \subset Sym_n$. Equation (3) is a key to find an explicit formula for the inverse of the Riemannian exponential on $G_{n,k}$.

As our approach does not exploit the fact that Grassmannians are symmetric spaces, to generalize the methodology to Stiefel manifolds and other homogeneous spaces of orthogonal group actions which are not symmetric spaces, is currently under consideration.

REFERENCES

- [1] E. Batzies, K. Hüper, L. Machado, and F. Silva Leite. Geometric mean and geodesic regression on Grassmannians. *Linear Algebra Appl.*, 466:83–101, 2015.
- [2] U. Helmke, K. Hüper, and J. Trumpf. Newton’s method on Grassmann manifolds, 2007. arXiv:0709.2205v2.
- [3] K. Hüper and F. Silva Leite. On the geometry of rolling and interpolation curves on S^n , SO_n , and Grassmann manifolds. *J. Dyn. Control Syst.*, 13(4):467–502, 2007.
- [4] R. Sharpe. *Differential geometry: Cartan’s generalization of Klein’s Erlangen program*. Springer, 1997.

Numerical Analysis for Maps into a Riemannian manifold

HANNE HARDERING

(joint work with Oliver Sander, Philipp Grohs)

We analyze finite element methods that are designed for maps that take their values in a Riemannian manifold. In particular, we develop $W^{1,2}$ - and L^2 - a priori error estimates for elliptic second order energy minimization problems. Our model problem is the Dirichlet problem for the harmonic map energy $\mathfrak{J}(v) := \frac{1}{2} \int_{\Omega} |dv(x)|_{g(v(x))}^2 dx$:

$$u = \arg \min_{w \in W_{\phi}^{1,2}(\Omega, M)} \mathfrak{J}(w),$$

where $W_{\phi}^{1,2}(\Omega, M)$ consists of Sobolev maps from a domain $\Omega \subset \mathbb{R}^d$ into a Riemannian manifold (M, g) with boundary and homotopy data ϕ . We assume that the manifold M and the data ϕ is such that there exists a locally unique local minimizer in $W^{m+1,2}(\Omega, M)$ to the problem.

We consider discretizations by *geometric finite elements*. By this we mean a finite-dimensional submanifold S_h^m of $W^{1,2} \cap C(\Omega, M)$, that is defined element-wise on a shape-regular, affine-equivalent, quasi-uniform mesh \mathcal{G}_h , and reduces to Euclidean Lagrangian finite elements of m -th order if $M = \mathbb{R}^n$.

The framework we introduce for the analysis directly translates ideas from the linear theory to manifolds without resorting to Nash embeddings or coordinate charts.

In particular, for $u \in W^{k,p} \cap C(\Omega, M)$ we use covariant derivatives of the differential du based on homogeneous spaces $W^{k,p} \cap W^{1,kp}$ to define the homogeneous smoothness descriptor as a generalization of Sobolev semi-norms by

$$\dot{\theta}_{k,p,\Omega}(u) = \left(\int_{\Omega} (|\nabla^k u|_g^p + |du|_g^{kp}) \, dx \right)^{\frac{1}{p}},$$

and the full inhomogeneous smoothness descriptor by

$$\theta_{k,p,\Omega}(u) = \left(\sum_{l=0}^k \dot{\theta}_{l,p,\Omega}^p(u) \right)^{\frac{1}{p}},$$

where $\dot{\theta}_{0,p,\Omega}^p(u) := \inf_{P \in M} \int_{\Omega} d^p(u(x), P) \, dx$, and we make the usual modifications for $p = \infty$.

In order to locally characterize differences of two maps $u, v \in C(\Omega, M)$ we use the point-wise inverse of the exponential map on M . This yields a vector field along one of the maps —w.l.o.g. we choose u — denoted by $\log_u v \in C(\Omega, u^{-1}TM)$. We introduce a notion of $W^{1,2}$ -error that uses the covariant derivative of this vector field along the map u , yielding a quasi-infra-distance —i.e. a distance with relaxed symmetry and triangle inequality— in a $W^{1,q}(\Omega, M)$ -ball around u , where $q > \max\{d, 2\}$, with the size of the ball depending on u and the curvature of M . When we consider a geodesic homotopy Γ within this ball starting at $\Gamma(0) = u$, we can show that given a vector field V_0 along the map u , we can use point-wise parallel transport to obtain a vector field V_t along each map $\Gamma(t)$, maintaining a bound on the $W^{1,2}$ -norm, i.e.,

$$\frac{1}{1 + Ct} \|V_0\|_{W^{1,2}(\Omega, u^{-1}TM)} \leq \|V_t\|_{W^{1,2}(\Omega, \Gamma(t)^{-1}TM)} \leq (1 + Ct) \|V_0\|_{W^{1,2}(\Omega, u^{-1}TM)}.$$

This result also holds for higher order Sobolev-norms, when we consider balls in correspondingly higher Sobolev-norms.

This crucial tool particularly allows for the transport of vector fields from the discrete minimizer to the continuous minimizer and thus for the compatibility of test vector fields in the variational formulation of both problems. This is used for the generalization of C ea’s Lemma for energy functionals that are elliptic along geodesic homotopies starting in the continuous minimizer. As long as best-approximation error estimates for S_h^m apply, one then obtains $W^{1,2}$ -discretization error estimates. However, as the lemma applies only in a ball around the continuous function u , one has to prove a priori bounds on the discrete solution. A possible way to derive these is via inverse estimates.

L^2 -error estimates rely on a generalization of the Aubin–Nitsche-Lemma. To apply this lemma, we need to assume that the problem is semi-linear, which is phrased in terms of bounds on the third variation of the energy. Using the variational formulations of the continuous and the discrete problems, and point-wise parallel transport of discrete test vector fields, one obtains an integral version of Galerkin orthogonality

$$0 = \int_0^1 \delta^2 \mathfrak{J}(\Gamma(t))(V_h(t), \dot{\Gamma}(t)) dt \quad \forall V_h(1) \in \mathcal{S}_0^h(\Omega, u_h^{-1}TM), \nabla_t V_h(t) \equiv 0.$$

A dual problem defined by linearization is given by

$$W \in W_0^{1,2}(\Omega, u^{-1}TM) : \\ \delta^2 \mathfrak{J}(u)(W, V) = -(V, \log_u u_h)_{L^2} \quad \forall V \in W_0^{1,2}(\Omega, u^{-1}TM).$$

In order to obtain the L^2 -estimates, one then needs to assume best-approximation error estimates on vector fields, as well as second-order Sobolev bounds on the discrete solution, as $W^{2,2}$ -norms of vector fields have to be transported.

Therefore, we assume semi-linearity and regularity of the dual problem (an assumption on the energy), as well as best-approximation errors of the variations of discrete maps (an assumption on the discretization). Further one needs second-order bounds on the discrete solution, which may be obtained by certain types of inverse estimates.

These results are independent of the specific approximation method used, and specify the kind of best-approximation errors a given method has to fulfill in order for the discretization errors to be derived in this manner. We give two specific examples for such interpolation methods, namely geodesic finite elements [3] and projection-based finite elements [1]. It can be shown that both methods fulfill the best-approximation error estimates as well as the inverse estimates to obtain H^1 - and L^2 -error estimates.

These ideas can be combined with techniques from the theory of surface finite elements in order to obtain discretization errors for maps from a hypersurface in \mathbb{R}^{d+1} into a Riemannian manifold. This relies on a generalization of the previously described framework to perturbed energies, which leads to a generalization of the Strang Lemmas.

REFERENCES

- [1] P. Grohs, H. Hardering, O. Sander, M. Sprecher. *Projection-Based Finite Elements for Nonlinear Function Spaces*, arXiv:1803.06576, 2018.
- [2] H. Hardering. *L^2 -Discretization Error Bounds for Maps into Riemannian Manifolds*, Numerische Mathematik (2018), 1–30.
- [3] P. Grohs, H. Hardering, O. Sander. *Optimal a priori discretization error bounds for geodesic finite elements*, Found. Comput. Math. (2015), **15**(6), 1357–1411.

Simulations of Cosserat materials and Dynamic Recrystallisation

THOMAS BLESSEN

(joint work with Stephan Luckhaus and Günther Gottstein)

A non-linear finite-strain Cosserat theory of crystal plasticity, [10], is studied which provides a detailed description of the behaviour of a material subject to mechanical forces. In contrast to the classical Prandtl-Reuss theory it is a *gradient model* giving rise to a *length scale*.

Typical of the Cosserat model is the multiplicative decomposition of the deformation tensor $F = D\phi$ in an elastic part F_e and a plastic part F_p , combined with the decomposition $F_e = R_e U_e$ in a stretching part $U_e \in \text{GL}(3)$ and a rotation part $R_e \in \text{SO}(3)$.

In the rate-independent case, the plastic dissipation can be computed analytically with the help of the Legendre-Fenchel dual Q^* analogous to the methods in [6], where Q is the plastic potential, for instance given by the von Mises condition.

Restricting the plastic slip to a-priori given material-dependent slip systems,

$$F_p = F_p(\gamma) := \text{Id} + \sum_a \gamma_a m_a \otimes n_a,$$

and parametrising R_e by Euler angles α , the complete mechanical description is obtained by the iterated minimisation of a time-discrete mechanical energy functional \mathcal{E} , see [3, 4, 7, 11],

$$\begin{aligned} \mathcal{E}(\phi, \alpha, \gamma) = & \int_{\Omega} W_{\text{st}}(R_e^t(\alpha) D\phi F_p(\gamma)^{-1}) + 2\mu_2 |\nabla \alpha|_2^2 - f_{\text{ext}} \cdot \phi - M_{\text{ext}} : R_e(\alpha) \\ & + \rho \left(\sum_a |\gamma_a - \gamma_a^0| \right)^2 + \sum_a |\gamma_a - \gamma_a^0| (\sigma_Y - 2\rho \sum_a \kappa_a^0) dx \rightarrow \min \end{aligned}$$

subject to $\phi|_{\partial\Omega} = g_d$, $\alpha|_{\partial\Omega} = \alpha_D$, where g_d and α_D are given. As shown in [4], \mathcal{E} is non-convex in α with several competing local minima. Minimising \mathcal{E} leads to the formation of sub-domains with constant rotations and interfacial layers, similar to phase transitions. The occurrence of several local minima ('lack of compactness') constitutes a major challenge to the algorithm as the simulations may get stuck without reaching the sought global minimiser. Currently, possible remedies are being investigated such as efficient preconditioning (e.g. by BPX along with extending the local basis to a suitable Haar basis) or initialising the numerical scheme with analytically-computed rotations which are optimal in special cases.

One possible important application of the Cosserat theory is the simulation of dynamic recrystallisation (DRX). DRX is a key mechanism to alter and manipulate the microstructure of a poly-crystal, especially a steel. Apart from the chemical composition, the mechanical history of the material is decisive for its properties.

In industrial processing, steels undergo hot and cold rolling to obtain the desired dislocation structure (and thus hardness). However, reliable predictions as well as a correct understanding of the physical effects within the material are lacking for

decades, despite intense research. The dislocation structure is mechanically stable at low temperatures, but thermodynamically unstable. Hence, during plastic deformations, new nuclei essentially free of dislocations are formed, leading to softening and recovery of the material. DRX is a true multiscale process and growing nuclei may range from the nanometer to the centimeter scale.

In [2], a variational model for DRX is proposed. Key components are:

- Hardening of the material by employing the afore-mentioned finite-strain Cosserat theory. The local rotations R_e are identified with the local rotations of the subgrains in the poly-crystal.
- (Sub-)Grain identification by a modified Ambrosio-Tortorelli functional from image analysis, [1], applied to the dislocation structure.
- Front-tracking and front propagation of the grain walls by a level-set algorithm, [12].
- Modeling of climbing dislocations as well as the permanent creation/ annihilation of new nuclei on small length scales by a stochastic Kolmogorov-Johnson-Mehl-Avrami equation, [9].

During the simulations, the material undergoes a cyclic uni-axial compression. Gradually, a dynamic equilibrium is generated between the hardening due to stored immobile dislocations and softening due to DRX. The simulated dislocation structure permits the computation of hysteresis and flow curves.

REFERENCES

- [1] L. Ambrosio, V.M. Tortorelli, *Approximation of Functionals Depending on Jumps by Elliptic Functionals via Γ -Convergence*, Comm. Pure Appl. Math. **43** (1990), 999–1036.
- [2] T. Blesgen, *A variational model for dynamic recrystallization based on Cosserat plasticity*, Composites Part B **115** (2017), 236–243.
- [3] T. Blesgen, *On rotation deformation zones for finite-strain Cosserat plasticity*, Acta Mechanica **226** (2015), 2421–2434.
- [4] T. Blesgen, *Deformation patterning in three-dimensional large-strain Cosserat plasticity*, Mechanics Research Communications **62** (2014), 37–43.
- [5] T. Blesgen, *Deformation patterning in Cosserat plasticity*, Modelling Simulation Mater. Sci. Eng. **21** (2013), 035001.
- [6] C. Carstensen, K. Hackl, A. Mielke, *Non-convex potentials and microstructures in finite-strain plasticity*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **458** (2002), 299–317.
- [7] I. Fonseca, G.A. Francfort, *Relaxation in BV versus quasiconvexification in $W^{1,p}$; a model for the interaction between fracture and damage*, Calculus of Variations **3** (1995), 407–446.
- [8] H. Hallberg, *A modified level set approach to 2D modeling of dynamic recrystallization*, Modelling Sim. Mat. Sci. Eng. **21** (2013), 1–33.
- [9] A.N. Kolmogorov, *Statistical theory of crystallization of metals*. (in russian), Izv. Akad. Nauk SSSR, Ser. Mat. **1** (1937), 355–359.
- [10] P. Neff, *A finite-strain elastic-plastic Cosserat theory for polycrystals with grain rotations*, Int. Journal of Eng. Science **32** (2006), 574–594.
- [11] M. Ortiz, E. Repetto, *Nonconvex energy minimization and dislocation structures in ductile single crystals*, JMPS **47** (1999), 397–462.
- [12] S.J. Ruuth, *A diffusion-generated approach to multiphase motion*, Journal Comp. Phys. **145** (1998), 166–192.

Elastic Distance Between Curves under the Metamorphosis Viewpoint

LAURENT YOUNES

There has been, over the past twenty years, a sizable amount of work exploring elastic distances between plane curves and their computation using a square-root transformation mapping the space of curves into some standard infinite-dimensional Riemannian manifold. In [1, 22, 24] a distance between parametrized plane curves was introduced, in which a transformation of the pair (ϕ, θ) (involving a square root) placed the metric in a Hilbert space context, where ϕ is the parametrization and θ is the tangent angle (as functions of the arc-length). This distance can then be optimized with respect to ϕ to yield a metric between curves modulo reparametrization (a.k.a. unparametrized curves). Existence results for minimizers were then provided in [17, 18]. Further analysis were made in the smooth case, with isometries with Stiefel and Grassman manifolds for closed curves and closed curves modulo rotations [23]. A different (but similar) approach was introduced in [6, 11] and further developed in numerous papers, among which [5, 7, 14–16, 21], to provide a metric between curves, also using a square root to reduce to a Hilbert case. More recently, the authors in [3] designed a different isometry applicable to a family of distances that includes the previous two.

In this talk, we reinterpret this line of work under the viewpoint of metamorphosis, which is described and developed in [4, 10, 12, 13, 19, 20]. This reformulation will allow us to generalize previous results on the subject by placing them in a unified context.

What we mean by elastic distances between curves are Riemannian metrics in spaces of parametrized curves that is, when evaluated at a smooth vector field along a curve, equivalent to the square norm of the derivative of this vector field with respect to the arc length. This is a small part of the range of Riemannian metrics that were considered in the literature. We refer to [8, 9] for an extensive catalog and properties and to [2, 3] for more recent developments.

Metamorphosis describes a general approach to build new Riemannian metrics on Riemannian manifolds acted upon by Lie groups [4, 10, 19, 20]. We consider a special case where curves are compared based on the orientation of their tangent. If $m : [0, L] \rightarrow \mathbb{R}^d$ is parametrized with arc length, its normalized tangent is

$$\begin{aligned} T^m : [0, 1] &\rightarrow \mathbb{R}^d \\ s &\mapsto \partial_s m(Ls). \end{aligned}$$

The function T^m characterizes m up to translation and scaling and the pair (L, T^m) characterizes m up to translation. In the following discussion, functions will depend on time $t \in [0, 1]$ and normalized arc length s , also in $[0, 1]$. For clarity, we will let $\Omega = [0, 1]$ for the arc length, i.e., write $t \in [0, 1]$, $s \in \Omega$. Given a function $h(t, s)$, $(t, s) \in [0, 1] \times \Omega$, we will denote $\dot{h} = \partial_t h$ for derivatives with respect to the time variable, and dh for derivatives with respect to the parameter.

We consider the group of diffeomorphisms G of Ω , which acts on the set M of measurable functions $a : \Omega \rightarrow S^{d-1}$ (the unit sphere in \mathbb{R}^d) by $\phi \cdot a = a \circ \phi^{-1}$ (and the group product is $\phi\psi = \phi \circ \psi$). “Tangent vectors” to G at $\phi = \text{id}$ are functions $v : \Omega \rightarrow \mathbb{R}$, with $v(0) = v(1) = 0$, and “tangent vectors” to M at α are functions $\xi : \Omega \rightarrow \mathbb{R}^d$ such that $\xi(s)^T \alpha(s) = 0$ almost everywhere on Ω . We also consider a Hilbert space V of continuous functions $v : \Omega \rightarrow \mathbb{R}$ (satisfying $v(0) = v(1) = 0$), with norm denoted $\|\cdot\|_V$. We then define the “metamorphosis energy”:

$$\int_0^1 \|\dot{\phi} \circ \phi^{-1}(t)\|_V^2 dt + \frac{1}{\sigma^2} \int_0^1 \int_{\Omega} |\dot{\alpha} \circ \phi^{-1}|^2 dt,$$

which must be minimized in (ϕ, α) with $\alpha(0) = a_0$ and $\alpha(1) = a_1 \circ \phi(1)$.

We take

$$\|v\|_V^2 = \int_{\Omega} dv(s)^2 ds,$$

which implies that functions $v \in V$ are continuous and satisfy a Hölder condition of order q for any $q < 1/2$, but are not necessarily Lipschitz continuous. An elementary computation followed by a change of variable in both integrals provides our final expression of the metamorphosis energy, namely

$$U_{\sigma}(\phi, \alpha) = \int_0^1 \int_{\Omega} \frac{(d\dot{\phi})^2}{d\phi} ds dt + \frac{1}{\sigma^2} \int_0^1 \int_{\Omega} |\dot{\alpha}|^2 d\phi ds dt$$

which needs to be minimized over all trajectories $t \mapsto \phi(t)$ and $t \mapsto \alpha(t)$, such that ϕ is at all times an increasing diffeomorphism of Ω and α a function from $\Omega \rightarrow S^{d-1}$, with boundary conditions $\alpha(0) = a_0$ and $\alpha(1) = a_1 \circ \phi(1)$. This energy coincides (up to a multiplicative constant) with the one introduced in [11]. We prove the following theorem.

Theorem 1. *Assume that $2\sigma \geq 1$, and let $\phi_1 : \Omega \rightarrow \Omega$ satisfy $\phi_1(0) = 0$, $\phi_1(1) = 1$ and $d\phi_1 > 0$. Then*

$$\begin{aligned} & \inf \{U_{\sigma}(\phi, \alpha) : \phi(1) = \phi_1, \alpha(0) = a_0, \alpha(1) = a_1 \circ \phi_1\} \\ & = 4 \arccos^2 \int_{\Omega} \sqrt{d\phi_1(s)} \cos \left(\frac{\arccos(a_0(s)^T a_1 \circ \phi_1(s))}{2\sigma} \right) ds. \end{aligned}$$

Moreover, if $2\sigma > 1$, the minimum is achieved and can be deduced from for a geodesic curve γ on the unit sphere of $L^2(\mathbb{R})$.

This induces a distance on M , given by

$$d_{\sigma}(a_0, a_1) = 2 \inf_{\phi_1} \left(\arccos \int_{\Omega} \sqrt{d\phi_1(s)} \cos (\arccos(a_0(s)^T a_1 \circ \phi_1(s))/2\sigma) ds \right)$$

minimized over all strictly increasing diffeomorphisms of Ω .

REFERENCES

- [1] Robert Azencott, François Coldefy, and Laurent Younes. A distance for elastic matching in object recognition. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, volume 1, pages 687–691. IEEE, 1996.
- [2] Martin Bauer, Martins Bruveris, Philipp Harms, and Jakob Møller-Andersen. A numerical framework for sobolev metrics on the space of curves. *SIAM Journal on Imaging Sciences*, 10(1):47–73, 2017.
- [3] Martin Bauer, Martins Bruveris, Stephen Marsland, and Peter W Michor. Constructing reparameterization invariant metrics on spaces of plane curves. *Differential Geometry and its Applications*, 34:139–165, 2014.
- [4] D. R. Holm, A. Trouvé, and L. Younes. The euler poincaré theory of metamorphosis. *Quarterly of Applied Mathematics*, 2009. (to appear).
- [5] S.H. Joshi, E. Klassen, A. Srivastava, and I. Jermyn. A novel representation for Riemannian analysis of elastic curves in r^n . In *Proceedings of CVPR'07*, 2007.
- [6] E. Klassen, A. Srivastava, W. Mio, and S. H. Joshi. Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(3):372–383, 2004.
- [7] Sebastian Kurtek, Eric Klassen, John C Gore, Zhaohua Ding, and Anuj Srivastava. Elastic geodesic paths in shape space of parameterized surfaces. *IEEE transactions on pattern analysis and machine intelligence*, 34(9):1717–1730, 2012.
- [8] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc.*, 8:1–48, 2006.
- [9] P. W. Michor and D. Mumford. An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1):74–113, 2007.
- [10] M. I. Miller and L. Younes. Group action, diffeomorphism and matching: a general framework. *Int. J. Comp. Vis.*, 41:61–84, 2001. Originally published in electronic form in: Proceedings of SCTV 99, <http://www.cis.ohio-state.edu/szhu/SCTV99.html>.
- [11] Washington Mio, Anuj Srivastava, and Shantanu Joshi. On shape of plane elastic curves. *International Journal of Computer Vision*, 73(3):307–324, 2007.
- [12] Casey L Richardson and Laurent Younes. Computing metamorphoses between discrete measures. *Journal of Geometric Mechanics*, 5(1), 2013.
- [13] Casey L. Richardson and Laurent Younes. Metamorphosis of images in reproducing kernel hilbert spaces. *Advances in Computational Mathematics*, 42(3):573–603, Jun 2016.
- [14] Chafik Samir, P-A Absil, Anuj Srivastava, and Eric Klassen. A gradient-descent method for curve fitting on riemannian manifolds. *Foundations of Computational Mathematics*, 12(1):49–73, 2012.
- [15] Anuj Srivastava and Eric P Klassen. *Functional and shape data analysis*. Springer, 2016.
- [16] Jingyong Su, Sebastian Kurtek, Eric Klassen, Anuj Srivastava, et al. Statistical analysis of trajectories on riemannian manifolds: bird migration, hurricane tracking and video surveillance. *The Annals of Applied Statistics*, 8(1):530–552, 2014.
- [17] A. Trouvé and L. Younes. Diffeomorphic matching in 1d: designing and minimizing matching functionals. In D. Vernon, editor, *Proceedings of ECCV 2000*, 2000.
- [18] A. Trouvé and L. Younes. On a class of optimal matching problems in 1 dimension. *Siam J. Control Opt.*, 39(4):1112–1135, 2001.
- [19] A. Trouvé and L. Younes. Local geometry of deformable templates. *SIAM J. Math. Anal.*, 37(1):17–59, 2005.
- [20] A. Trouvé and L. Younes. Metamorphoses through lie group action. *Found. Comp. Math.*, pages 173–198, 2005.
- [21] Qian Xie, Sebastian Kurtek, Huiling Le, and Anuj Srivastava. Parallel transport of deformations in shape space of elastic surfaces. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 865–872, 2013.

- [22] L. Younes. Computable elastic distances between shapes. *SIAM J. Appl. Math.*, 58(2):565–586, 1998.
- [23] L. Younes, P. Michor, J. Shah, and D. Mumford. A metric on shape spaces with explicit geodesics. *Rend. Lincei Mat. Appl.*, 9:25–57, 2008.
- [24] Laurent Younes. A distance for elastic matching in object recognition. *C. R. Acad. Sci. Paris Sér. I Math.*, 322(2):197–202, 1996.

Generation of Frame Fields for Hexahedral Meshing

EDWARD CHIEN

(joint work with David Palmer, Paul Zhang, Justin Solomon, Heng Liu, David Bommes)

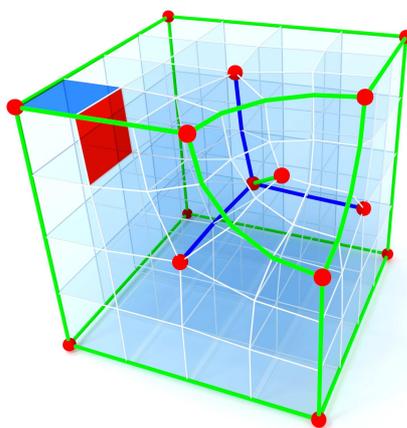


FIGURE 1. Hex mesh example. Credit: Heng Liu from [3].

A *hexahedral mesh* of a compact 3-dimensional manifold with boundary $\Omega \subset \mathbb{R}^3$ is a cell decomposition where each 3-cell is of hexahedral type. Such meshes (hex meshes, for short) are preferred in industry finite element (FEM) simulations, and a simple example is shown in Fig. 1. The colored interior/boundary edges and vertices indicate those for which the nearby cell structure is irregular, i.e., unlike the neighborhood of an interior/boundary edge or vertex in the regular \mathbb{Z}^3 -based cell decomposition of the upper half-space $\{(x, y, z) \mid z \geq 0\} \subset \mathbb{R}^3$. These edges and vertices are referred to as *singular*, and sparse singular structure for a mesh is desired, as this typically leads to lower geometric distortion of mesh elements and higher simulation fidelity.

Frame Fields. Currently, there are no automatic methods that can generate hex meshes with relatively simple singular structure. A parametrization-based approach has had success in the 2-dimensional analogue of this problem, generating a simple quad mesh of a surface by first generating a smoothly varying field of two orthogonal axes [1]. To generalize this to the 3-dimensional case, a *frame field* is needed, which is a specification of three orthogonal axes throughout Ω . These

axes guide the creation of a hex mesh by roughly specifying the orientation of mesh elements nearby.

An element $R \in \text{SO}(3)$ specifies three orthogonal axes via its column vectors: $\text{Span}(Re_1)$, $\text{Span}(Re_2)$, and $\text{Span}(Re_3)$. However, the same set of axes is specified by the coset $R(\text{Oct})$, where $\text{Oct} \in \text{SO}(3)$ is the group of orientation-preserving isometries of $[-1, 1]^3$ (or its dual polytope, the octahedron). Thus, the space of frames may be characterized as the quotient $\text{SO}(3)/\text{Oct}$, where the quotient is by the right action.

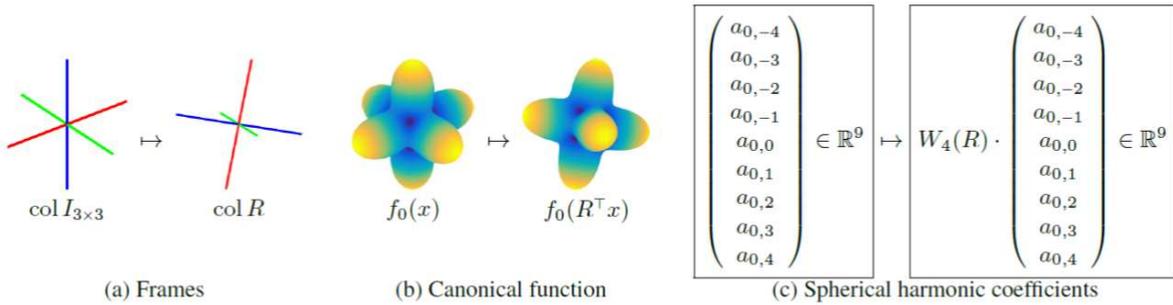


FIGURE 2. Wigner rotation of spherical harmonic coefficients. Image from [5].

Computation of smooth maps into this quotient space is a challenging task. The most successful approach to date relies on mapping into the space of $l = 4$ spherical harmonics, an idea first introduced in [4]. This is a 9-dimensional vector space of real-valued functions on \mathbb{S}^2 , given by restriction of homogeneous degree 4 polynomials $p(x, y, z)$ satisfying Laplace’s equation $\Delta p(x, y, z) = 0$. The spherical harmonics of all degrees define a Fourier basis for $L^2(\mathbb{S}^2)$, and the lower degree members are often used to illustrate atomic orbitals. We use the standard basis for the real spherical harmonics, denoted by $\{Y_4^m\}_{m=-4}^4$.

Within this space, frames are represented by functions peaked in three orthogonal directions. A canonical example is given by $f_0 = x^4 + y^4 + z^4 = \sqrt{7/12}(Y_4^0) + \sqrt{5/12}(Y_4^4)$, peaked along the standard Cartesian axes. A frame specified by $R \in \text{SO}(3)$ is represented by $f_0 \circ R^T$. Letting $a_0 = (0, 0, 0, 0, \sqrt{7/12}, 0, 0, 0, \sqrt{5/12})^T$, Fig. 2 illustrates that composition by R^T changes the spherical harmonic coefficients by application of the Wigner D-matrix, $W_4(R)$, associated with R . Let us use Γ to denote the set of spherical harmonic coefficients that represent frames:

$$\Gamma := \{a \in \mathbb{R}^9 \mid a = W_4(R)a_0 \text{ for some } R \in \text{SO}(3)\}.$$

In existing works on frame field generation, an initial map into \mathbb{R}^9 is projected onto this set in an unprincipled fashion. In an attempt to ameliorate this, we have characterized the cone over Γ , denoted $C(\Gamma)$, as a projective variety:

$$C(\Gamma) = \{a \in \mathbb{R}^9 \mid x^T P_i x = 0 \text{ for } i = 1, \dots, 5\},$$

where the matrices $P_i \in \mathbb{R}^{9 \times 9}$ are symmetric. These matrices are listed below:

$$\begin{aligned} P_1 &= L_x L_x - L_y L_y & P_3 &= L_x L_y \\ P_2 &= L_y L_y - L_z L_z & P_4 &= L_y L_z \\ & & P_5 &= L_z L_x \end{aligned}$$

where L_x, L_y, L_z denote elements of $\mathfrak{so}(9)$. These are the images of the standard Lie algebra basis of $\mathfrak{so}(3)$ under the Lie algebra homomorphism induced by the Wigner representation $R \rightarrow W_4(R)$. They represent the infinitesimal changes to the spherical harmonic coefficients under rotations of the function about the x, y, z -axes. The argument proceeds by first proving rotational invariance of the conditions, and then performs differential analysis on critical points of the functions in Γ . It is our hope that such a characterization will allow for direct mapping without projection, and suggestions and input are welcome.

Topological Investigations. We also present two additional topological facts, for discussion. Through a search of the 3-manifold literature, we have found that $SO(3)/Oct$ is a spherical 3-manifold, and admits a Seifert-fibered with base space \mathbb{S}^2 , with orbifold structure $(2,3,4)$ [2]. It would be interesting to find a method for computing frame fields which utilizes this fact.

Additionally, first steps have been made towards characterizing the possible singular structures for a hexahedral mesh. Let $\mathcal{S} = (V_{\mathcal{S}}, E_{\mathcal{S}})$ denote the *singularity graph* of a hexahedral mesh, consisting of the singular vertices and edges, respectively. Let $\text{val}_h(p)$ denote the number of hexahedral 3-cells that contain p as an edge or vertex. In an upcoming work [3], we demonstrate an analogue to the Hopf-Poincaré formula, which must hold for any singularity graph of a hexahedral mesh:

$$\begin{aligned} \sum_{v \in \partial V_{\mathcal{S}}} \frac{1}{2} \left(1 - \frac{\text{val}_h(v)}{4} \right) - \sum_{e \in \partial E_{\mathcal{S}}} \left(\frac{1}{2} - \frac{\text{val}_h(e)}{4} \right) \\ + \sum_{v \in \text{int } V_{\mathcal{S}}} \left(1 - \frac{\text{val}_h(v)}{8} \right) - \sum_{e \in \text{int } E_{\mathcal{S}}} \left(1 - \frac{\text{val}_h(e)}{4} \right) = 0. \end{aligned}$$

In the above expression, the ∂ and int prefixes denote boundary and interior vertices/edges, respectively. The result follows by a combinatorial argument, and is necessary, but not sufficient, for a singularity graph to be induced by a hex mesh of a volume Ω . Additional local characterizations are present in [3], but further results will be needed to determine the set of such singularity graphs entirely.

REFERENCES

- [1] D. Bommes, B. Lévy, N. Pietroni, E. Puppo, C. Silva, M. Tarini, D. Zorin, *Quad-Mesh Generation and Processing: A Survey*, Computer Graphics Forum **32(6)** (2013), 51–76.
- [2] W. Jaco, *Lectures on Three-Manifold Topology*, American Mathematical Society (1980), Providence, R.I.
- [3] H. Liu, P. Zhang, E. Chien, J. Solomon, D. Bommes. *Singularity-Constrained Octahedral Fields for Hexahedral Meshing*, ACM Trans. Graph. **37(4)** (2018).

- [4] J. Huang, Y. Tong, H. Wei, H. Bao, *Boundary-Aligned Smooth 3D Cross-Frame Field*, ACM Trans. Graph. **30(6)** (2011), 1–8.
- [5] J. Solomon, A. Vaxman, D. Bommes, *Boundary Element Octahedral Fields in Volumes*, ACM Trans. Graph. **36(4)** (2017).

Estimation of Nonlinear Single Index Models

TIMO KLOCK

(joint work with Zeljko Kereta, Mauro Maggioni, Valeriya Naumova)

In this work we propose an efficient method for estimating a generalisation of the single index model, which we call the *nonlinear single index model*. We begin by describing the problem.

Let $(X, Y) \sim \rho$ be a random vector in $\mathbb{R}^D \times \mathbb{R}$ that obeys the model $Y = f(X) + \varepsilon$, where f and ρ are unknown and ε represents noise. A common task in regression problems is to learn an estimator $\hat{f}(\cdot; S_N)$ of $f(\cdot)$ using a random data set $S_N = \{(X_1, Y_1), \dots, (X_N, Y_N)\} \sim \rho^N$, where the performance is measured in terms of the mean-square error $\mathbb{E}[(\hat{f}(X; S_N) - f(X))^2 | S_N]$, and should improve as the sample size N grows. It is known that, without imposing any additional assumptions, the best possible rate of estimating a \mathcal{C}^s -function f is $N^{-2s/(2s+D)}$. In other words, the rate deteriorates exponentially as the ambient dimension D increases, thus making the task of learning f intractable in high dimensions due to the demand on the number of samples. However, it has been observed that the relation between $f(X)$ and X is intrinsically low-dimensional (at least approximately) when considering real-world data sets. This lead to an increasing number of methods and models that attempt to exploit various structural assumptions to beat this so-called *curse of dimensionality*.

A popular choice is the *single index model* (SIM)

$$(1) \quad f(x) = g(\langle a, x \rangle), \quad a \in \mathbb{S}^{D-1}, \quad x \in \mathbb{R}^D,$$

that has been a subject of an enormous body of research. Assuming the direction a is known, estimating f is a univariate regression for which the learning rate $N^{-2s/(2s+1)}$ can be achieved. In practice though, a is often unknown and thus has to be estimated from data. Recently, [2–4] considered (1) for a monotonic g , and proposed estimators that are guaranteed to achieve D -independent, but also sub-optimal, learning rates. In [1] an approach is investigated that is promising, and seems to achieve the optimal rate $N^{-2s/(2s+1)}$. Although the SIM model might seem restrictive, it includes the important class of generalized linear models with an arbitrary monotone link function g .

A natural generalisation of (1) with strictly monotonous g is to assume that the feature weighting a varies as different regimes of Y are explored. This can be observed in some typical data sets where (1) is used, and naturally leads us to define the *nonlinear single index model* (NSIM)

$$(2) \quad f(x) = g(\pi_\gamma(x)).$$

Here $\gamma : \mathcal{I} \rightarrow \mathbb{R}^D$ is an arc-length parametrisation of a simple, smooth curve, and π_γ is the orthogonal projection

$$\pi_\gamma(x) := \arg \min_{z \in \text{Im}(\gamma)} \|x - z\|_2.$$

Provided γ is known and that the composition $g \circ \pi_\gamma$ is \mathcal{C}^s -smooth, we expect that, similarly to SIM, the NSIM function model can be learned at the univariate rate $N^{-2s/(2s+1)}$.

In the talk we present preliminary results on the estimation of (2). Standard manifold-estimation and -regression techniques are, to the best of our knowledge, consistent only if the marginal distribution ρ_X is supported on $\text{Im}(\gamma)$. Namely, if $\sigma := \mathbb{E}\|X - \pi_\gamma(X)\|_2$ is positive, standard estimators will fail to resolve regression tasks beyond the geometrical limit imposed by σ . Our method on the other hand, is specifically tailored to such data and makes explicit use of the assumption that ρ_X is supported in a thick tube around $\text{Im}(\gamma)$.

In order to estimate f we follow a two step method. First, we compute a discrete approximation of the tangent field $\dot{\gamma}(t)$ (which is a topic of independent interest), and second, we use the tangent approximations to find a proxy for the geodesic metric on $\text{Im}(\gamma)$ and construct an estimator \hat{f} .

Tangent estimation. Our approach resembles the one used in [1] and is based on inverse regression. We begin by dividing $[\min Y, \max Y]$ into small intervals $[c_l, c_{l+1}]$ (e.g. using a dyadic decomposition) and define level sets

$$D_l := \{X \in \{X_1, \dots, X_N\} : c_l \leq Y < c_{l+1}\}.$$

Provided that $(c_{l+1} - c_l)/\sigma$ and $\|\varepsilon\|/\sigma$ are sufficiently small, the variance of D_l will be small in the tangential direction, but large in any orthogonal direction. Moreover, the convex hull $\mathcal{I}_l = \text{Conv}(\gamma^{-1} \circ \pi_\gamma(D_l))$ is a *thin* subinterval of \mathcal{I} . We can prove that the smallest singular vector \hat{a}_l of the sample covariance matrix of D_l satisfies

$$(3) \quad \|\hat{a}_l - \dot{\gamma}(t)\|_2 \asymp \left(\frac{1}{\sqrt{N_l}} + c_\kappa \right) |\mathcal{I}_l|, \quad t \in \mathcal{I}_l, \quad \mathcal{I}_l := \text{Length}(\mathcal{I}_l),$$

with high probability on a draw of S_N , provided that the number of samples $N_l = \#D_l$ is sufficiently large. In the special case where γ is linear (i.e. in the SIM case), we can show $c_\kappa = 0$, and thus \hat{a}_l is a consistent estimator of the direction a .

Out of sample prediction and the construction of \hat{f} . To facilitate out-of-sample prediction we adhere to a modified version of the *knn* estimator, which means that for any new X we first need to identify data samples $X_i \in S_N$ that are close to X in the geodesic metric. Note now that if $\sigma > 0$ the Euclidean distance $\|X - X_i\|_2$ is locally a poor substitute for the geodesic distance (once the geodesic distance is comparable with σ). We circumvent this issue by leveraging tangential field approximations.

It is useful to consider a simpler class of curves first. Assume γ satisfies $\langle \dot{\gamma}(t_i), \dot{\gamma}(t_j) \rangle > \lambda$ for some $0 < \lambda < 1$, and all $t_i, t_j \in \mathcal{I}$, which is a property we refer to as λ *almost linearity*. For such curves, the quantity

$$\Delta(X, X_i) = |\langle \hat{a}_{l(X_i)}, X - X_i \rangle|, \quad X_i \in D_{l(X_i)}, \quad X \sim \rho_X,$$

is a proxy for the geodesic distance, and we can prove

$$\left| \Delta(X, X_i) - |\gamma^{-1}(\pi_\gamma(X)) - \gamma^{-1}(\pi_\gamma(X_i))| \right| \asymp \max_l |\mathcal{I}_l|, \quad \text{if } \kappa\sigma < \lambda.$$

Thus, in case of a λ almost linear curve, we can choose k closest neighbors through the proxy metric $\Delta(X, X_i)$, and estimate $f(X)$ by averaging the corresponding function values $f(X_i)$. If k is chosen according to $k \asymp \mathcal{O}(N^{2\alpha/(2\alpha+1)})$, such an estimator reaches optimal convergence rates for α -Hölder functions, $\alpha \leq 1$, up to an error in $\mathcal{O}(|\mathcal{I}_l|)$.

Since a typical curve will globally not be λ almost linear, we re-partition the curve into λ almost linear segments before performing the prediction step. This is done by passing through all level sets $\{D_l\}_{l=1,2,\dots}$ (on which we can establish an ordering using the monotonicity of the corresponding Y -values), and collecting them into larger families sequentially, i.e. as long as $\langle \hat{a}_l, \hat{a}_{l'} \rangle > \lambda$ holds for all elements of the family. Once the condition is violated, $\langle \dot{\gamma}(t_i), \dot{\gamma}(t_j) \rangle > \lambda$ is likely to be violated as well for some t_i, t_j along the corresponding stretch of the curve (due to (3)). Thus, the current level set family is marked as finished and we start creating a new one. We continue this process until we pass through all D_l , and by that induce a coarser re-partition of S_N where each element contains a union of contiguous level sets, and the part of the curve restricted to any element of the partition is roughly λ almost linear. Once a point has been assigned to the correct member of the partition we can again use the distance proxy $\Delta(X, \cdot)$ to identify k nearest neighbours.

We are currently investigating under which conditions the assignment to the correct almost linear element of the partition is guaranteed to yield optimal performance. Empirically, conventional classification methods such as linear SVMs show promising results to discriminate between elements of the coarser partition. We note that the number of λ almost linear subpieces of the curve depends only on λ and γ and as such, the number of elements in the coarse partition is (virtually) independent of the fineness of $\{D_l\}_{l=1,2,\dots}$. Thus, the learning complexity of classifiers used to discriminate between two segments does not increase as we strive for higher accuracy through finer level set partitioning.

We emphasize that our method has a polynomial computational complexity $\mathcal{O}(DN^2 + ND^2)$. In the talk, we have shown results with synthetic data that support the claims made. Moreover, we also identify an open problem: in the current form the estimator is not consistent if $\varepsilon \neq 0$ because $|\mathcal{I}_k| \not\rightarrow 0$, even if $c_{k+1} - c_k$ is chosen arbitrarily close to zero. Thus, we need to modify our approach to achieve consistent estimation in that case.

REFERENCES

- [1] Alessandro Lanteri, Mauro Maggioni, Stefano Vigogna, *Inverse regression for single-index models* (2018), in preparation.
- [2] L. Balzano, R. Ganti, R.D. Nowak, N. S. Rao, R. Willett, *On Learning High Dimensional Structured Single Index Models*, *Advancement of Artificial Intelligence* (2017), 1898–1904.
- [3] S. M. Kakade, V. Kanade, O. Shamir, A. Kalai, *Efficient learning of generalized linear and single index models with isotonic regression*, *Advances in Neural Information Processing Systems* (2011), 927–935.
- [4] P. Radchenko, *High dimensional single index models* *Journal of Multivariate Analysis* **139** (2015), 266–282.

The Existence Problem of Cubic Spline Interpolation on Riemannian Manifolds

BENEDIKT WIRTH

(joint work with Behrend Heeren, Martin Rumpf)

A classical data processing tool in Euclidean space \mathbb{R}^n is spline interpolation. Given (pairwise disjoint) time and data points

$$t_1, \dots, t_m \in [0, 1], \quad x_1, \dots, x_m \in \mathbb{R}^n,$$

an interpolating degree- l -spline is an $(l-1)$ times continuously differentiable curve γ with

$$\gamma(t_i) = x_i, \quad i = 1, \dots, m,$$

such that the restriction of γ to any time interval $[t_i, t_{i+1}]$ is a polynomial of degree l (these conditions have to be complemented with appropriate boundary conditions, where we shall assume natural boundary conditions throughout).

Euclidean spline interpolation enjoys several nice properties. In particular, for $m > l$ there always exists a unique interpolating degree- l -spline, the interpolation smoothness can be readily controlled by adjusting the degree l , and perturbations of data points show a rather localized influence on the interpolating curve. Furthermore, the most widely used spline interpolations obey a variational principle: Linear spline interpolation can be expressed as the interpolating curve minimizing the average squared velocity,

$$\gamma_{\text{lin}} = \arg \min \{E[\gamma] \mid \gamma(t_i) = x_i, i = 1, \dots, m\} \quad \text{with} \quad E[\gamma] = \int_0^1 |\dot{\gamma}|^2 dt,$$

while cubic spline interpolation minimizes the average squared acceleration,

$$\gamma_{\text{cub}} = \arg \min \{F[\gamma] \mid \gamma(t_i) = x_i, i = 1, \dots, m\} \quad \text{with} \quad F[\gamma] = \int_0^1 |\ddot{\gamma}|^2 dt.$$

Applications such as keyframe animation in computer vision require interpolation in non-Euclidean spaces such as Riemannian manifolds. Spline interpolation can be generalized to data in a Riemannian manifold (\mathcal{M}, g) in several ways:

- Parameterizing \mathcal{M} via a chart, classical Euclidean spline interpolation can be applied. For instance, \mathcal{M} can be parameterized by its tangent space $T_x\mathcal{M}$ in $x \in \mathcal{M}$ using the Riemannian exponential so that the interpolation problem reduces to calculating the Euclidean spline interpolation of the Riemannian logarithms $\log_x x_i$ in $T_x\mathcal{M}$.
- Degree- l -splines can be characterized as C^{l-1} -curves with vanishing $(l+1)^{\text{th}}$ derivative on each time interval $[t_i, t_{i+1}]$. Replacing derivatives with covariant derivatives, this translates to \mathcal{M} .
- Geometric spline constructions such as de Casteljau's algorithm can be generalized by replacing line segments with geodesics.
- From t_1, \dots, t_m one can compute coefficients $\alpha_i(t)$ such that Euclidean spline interpolation reads

$$\gamma(t) = \sum_{i=1}^m \alpha_i(t) x_i = \text{critical point of } \lambda \mapsto \sum_{i=1}^m \alpha_i(t) |\lambda - x_i|^2$$

in which $|\lambda - x_i|$ can be replaced with the Riemannian distance $\text{dist}(\lambda, x_i)$.

- Exploiting the variational principle, linear and cubic splines on \mathcal{M} can be defined as minimizers of generalized energies E and F .

We follow the latter approach: The velocity $\dot{\gamma}(t)$ of a curve $\gamma : [0, 1] \rightarrow \mathcal{M}$ lies in the tangent space $T_{\gamma(t)}\mathcal{M}$. Furthermore, the (intrinsic) acceleration of γ is the covariant derivative of $\dot{\gamma}$ along the curve, which for a parameterized manifold can be defined as

$$\frac{D}{dt}\dot{\gamma}(t) = \ddot{\gamma}(t) + \Gamma_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) \in T_{\gamma(t)}\mathcal{M}$$

with the Christoffel operator $\Gamma_x : T_x\mathcal{M} \times T_x\mathcal{M} \rightarrow T_x\mathcal{M}$ defined via

$$2g_x(\Gamma_x(v, w), z) = (D_x g_x)(w)(v, z) - (D_x g_x)(z)(v, w) + (D_x g_x)(v)(w, z).$$

Finally, the squared norm of tangent vectors $v \in T_x\mathcal{M}$ can be expressed as $g_x(v, v)$, hence we arrive at the generalized energies

$$E[\gamma] = \int_0^1 g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) dt, \quad F[\gamma] = \int_0^1 g_{\gamma(t)}\left(\frac{D}{dt}\dot{\gamma}(t), \frac{D}{dt}\dot{\gamma}(t)\right) dt$$

and define the linear and cubic spline interpolation of data $x_1, \dots, x_m \in \mathcal{M}$ as the interpolating curves $\gamma : [0, 1] \rightarrow \mathcal{M}$ minimizing E and F , respectively.

Linear spline interpolation on a Riemannian manifold \mathcal{M} in the above sense turns out to equal piecewise geodesic interpolation and is well-understood. Existence holds under mild, standard conditions on \mathcal{M} , while uniqueness is lost in general due to nonuniqueness of shortest geodesics. We shall instead discuss the more interesting question of existence of cubic spline interpolations on \mathcal{M} .

Consider a minimizing sequence $\gamma_1, \gamma_2, \dots : [0, 1] \rightarrow \mathcal{M}$ with $\gamma_j(t_i) = x_i$ for all $i = 1, \dots, m$ and $j \geq 1$ such that

$$\lim_{j \rightarrow \infty} F[\gamma_j] = \inf \{F[\gamma] \mid \gamma : [0, 1] \rightarrow \mathcal{M}, \gamma(t_i) = x_i, i = 1, \dots, m\}.$$

Following the direct method of the calculus of variations, the existence analysis typically requires two properties: Compactness of the minimizing sequence (or

equivalently coercivity of F) and lower semi-continuity of F . As it turns out, both properties seem to get lost in general when passing from Euclidean space to a Riemannian manifold.

Lack of compactness. The functional F controls the acceleration of a curve γ in an L^2 -sense. Likewise, the curve location is controlled by the interpolation data. In \mathbb{R}^n , this implies control of the curve velocities via the Gagliardo–Nirenberg inequality; as a result, the sequence γ_j is bounded in $W^{2,2}((0, 1); \mathbb{R}^n)$ and thus precompact with respect to weak convergence. However, on a general manifold \mathcal{M} , the curve velocity is not controlled at all: As an example, consider \mathcal{M} to be a cylinder with interpolation times $t_1 = 0, t_2 \in (0, 1), t_3 = 1$ and data $x_1 = x_3$ and x_2 lying exactly opposite of x_1 . For given integers $a, b \geq 0$ it is straightforward (since the cylinder is developable so that calculations can be reduced to Euclidean space) to calculate

$$\min\{F[\gamma] \mid \gamma(t_1) = \gamma(t_3) = x_1 = x_3, \gamma(t_2) = x_2, \gamma|_{[t_1, t_2]} \text{ and } \gamma|_{[t_2, t_3]} \text{ wind } a + \frac{1}{2} \text{ and } b + \frac{1}{2} \text{ times round } \mathcal{M}\} \sim (a + \frac{1}{2} - (a + b + 1)t_2)^2.$$

Thus, for any $t_2 = \frac{a+1/2}{a+b+1}$ with integers $a, b \geq 0$ there is an interpolating curve γ with $F[\gamma] = 0$, a minimizer. However, for $t_2 \notin \mathbb{Q}$, Dirichlet’s approximation theorem implies that large enough a and b (resulting in arbitrarily high curve velocities) can make $F[\gamma]$ arbitrarily small, but never zero. Hence there is no minimizer.

Noting that E controls curve velocities, compactness can be regained by replacing F with

$$F_\varepsilon = F + \varepsilon E$$

for a small regularization parameter $\varepsilon > 0$. Under the simplifying assumptions of $\mathcal{M} \equiv X$ for a Hilbert space X and the (spatially varying) Riemannian metric g being equivalent to the inner product on X , the argument is as follows. Since $E[\gamma]$ behaves like the Sobolev semi-norm $|\gamma|_{W^{1,2}}$, we already know uniform boundedness of $|\gamma_j|_{W^{1,2}}$ (and by Poincaré’s inequality of the full Sobolev norm). Now, can it happen that $\|\gamma_j\|_{W^{2,2}}$ diverges? This would require $\|\ddot{\gamma}_j\|_{L^2}^2 \rightarrow \infty$, while at the same time $\|\dot{\gamma}_j\|_{L^2}^2$ and $\|\ddot{\gamma}_j + \Gamma_{\gamma_j}(\dot{\gamma}_j, \dot{\gamma}_j)\|_{L^2}^2 \sim F[\gamma_j]$ are uniformly bounded. However, this implies that $\|\ddot{\gamma}_j\|_{L^2}^2$ and $\|\Gamma_{\gamma_j}(\dot{\gamma}_j, \dot{\gamma}_j)\|_{L^2}^2 \sim \|\dot{\gamma}_j\|_{L^4}^4$ must diverge at the same rate which one can show contradicts boundedness of $\|\dot{\gamma}_j\|_{L^2}^2$. Thus we have boundedness of $\|\gamma_j\|_{W^{2,2}}$, and we can extract a weakly converging subsequence $\gamma_j \rightharpoonup \gamma$ in $W^{2,2}$.

Lack of lower semi-continuity. $F[\gamma]$ is quadratic in the intrinsic acceleration $\frac{D}{dt}\gamma$. Hence, for lower semi-continuity of F along $\gamma_j \rightharpoonup \gamma$ it is natural to require weak convergence of $\frac{D}{dt}\gamma_j = \ddot{\gamma}_j + \Gamma_{\gamma_j}(\dot{\gamma}_j, \dot{\gamma}_j)$ to $\frac{D}{dt}\gamma$ in L^2 . In both a finite-dimensional space and in (potentially infinite-dimensional) Euclidean space this is implied by $\gamma_j \rightharpoonup \gamma$ in $W^{2,2}$: Indeed, in Euclidean space the Christoffel operator is uniformly zero,

$$\Gamma \equiv 0,$$

implying automatically the desired convergence. Similarly, in a finite-dimensional manifold, the weak convergence $\gamma_j \rightharpoonup \gamma$ in $W^{2,2}$ implies by Sobolev embedding

strong convergence $\gamma_j \rightarrow \gamma$ in a Hölder space $C^{1,\alpha}$, from which one can deduce

$$\Gamma_{\gamma_j}(\dot{\gamma}_j, \dot{\gamma}_j) \rightarrow \Gamma_{\gamma}(\dot{\gamma}, \dot{\gamma}) \quad \text{in } C^0.$$

However, in an infinite-dimensional manifold, the above Sobolev embedding no longer holds (one only has $W^{2,2}((0,1); X) \hookrightarrow C^1([0,1]; Y)$ for some space Y into which X compactly embeds), and $\Gamma_{\gamma_j}(\dot{\gamma}_j, \dot{\gamma}_j)$ as a quadratic form of a weakly converging sequence cannot be expected to converge to $\Gamma_{\gamma}(\dot{\gamma}, \dot{\gamma})$ even in a weak sense, which prevents lower semi-continuity.

A remedy is to restrict to appropriate manifolds \mathcal{M} . Considering the above, lower semi-continuity is valid on flat manifolds and locally compact manifolds (which is required for the strong Sobolev embedding). It is possible to combine both settings to arrive at a more general class of manifolds: F is lower semi-continuous on manifolds $\mathcal{M} = X$ whose Riemannian metric g decomposes,

$$g_x(v, w) = Q(v, w) + g_x^c(v, w) \quad \text{for } v, w \in T_x\mathcal{M},$$

into a flat, spatially constant metric Q and a spatially varying metric g_x^c with respect to which the manifold \mathcal{M} is locally compact.

REFERENCES

- [1] B. Heeren, M. Rumpf, B. Wirth, *Variational time discretization of Riemannian splines*, accepted at IMA Journal of Numerical Analysis (2018).

An Observation Concerning the Parallel Transport Variant of Total Generalized Variation for Manifold-Valued Data

ANDREAS WEINMANN

(joint work with Kristian Bredies, Martin Holler, Martin Storath)

Total variation type approaches for regularizing manifold valued data and images have gained a lot of interest in recent years [1, 2, 5, 6, 8–10]. To overcome drawbacks from staircasing and oversmoothing, we have introduced the total generalized variation (TGV) functional for regularizing manifold-valued data in [3]. More precisely, we have considered the variational problem

$$(1) \quad \arg \min_{u \in \mathcal{M}^K} \text{dist}(u, f)^2 + \text{TGV}_{\lambda}(u),$$

where f denotes the given data, dist denotes the distance on the product manifold \mathcal{M}^K , induced by a Riemannian metric on \mathcal{M} , and u is the argument to optimize for. The idea is to find an estimate u which both approximates the data f well and is more regular than f . Here, the regularity of u is measured by the total generalized variation $\text{TGV}_{\lambda}(u)$. In the univariate setting,

$$\text{TGV}_{\lambda}(u) = \arg \min_v \lambda_0 \sum_i \text{dist}(u_{i+1}, v_i) + \lambda_1 \sum_i D([u_i, v_i], [u_{i-1}, v_{i-1}]),$$

where λ_1, λ_2 are positive parameters, and where D generalizes the distance between difference vectors to the manifold setting. We here represent difference vectors as

tuples $[u_i, v_i]$ of points u_i, v_i in the manifold \mathcal{M} . One instantiation of D employed in [3] uses the parallel transport in \mathcal{M} ; it reads

$$(2) \quad D([u_i, v_i], [u_{i-1}, v_{i-1}]) = \|\log_{u_i} v_i - \text{pt}_{u_{i-1}, u_i} \log_{u_{i-1}} v_{i-1}\|_{u_i}.$$

Here \log denotes the inverse of the Riemannian exponential mapping, the symbol pt_{u_{i-1}, u_i} denotes the parallel transport along a shortest geodesic connecting u_{i-1} and u_i such that $\log_{u_i} v_i - \text{pt}_{u_{i-1}, u_i} \log_{u_{i-1}} v_{i-1}$ is a tangent vector sitting in u_i , and $\|\cdot\|_{u_i}$ denotes the norm induced by the Riemannian metric in u_i . In [3] we have taken an axiomatic point of view and formulated requirements for a reasonable generalization of D to the manifold setting, and we have considered a further instantiation of D based on the Schild’s ladder; in the multivariate setup, we have introduced an additional term based on a symmetrized gradient for manifold valued data in the regularizing term; for details, we refer to [3]. We here concentrate on the parallel transport variant. Most algorithms dealing with problems of the form (1), require the computation of the (sub)gradients of (2) w.r.t. the variables $u_i, u_{i-1}, v_i, v_{i-1}$; in [3], we have derived rather explicit representations of these (sub)gradients. More precisely, we have observed that the problem is symmetric w.r.t. interchanging $[u_i, y_i]$ and $[u_{i-1}, y_{i-1}]$ and so reduces to finding the gradients w.r.t u_i and y_i . W.r.t. y_i we have obtained expressions in terms of Jacobi fields which, in symmetric spaces, results in solving eigenvalue problems and systems of equations of typically small size. Computing the gradient w.r.t. u_i again involves computing Jacobi fields, and essentially the derivatives of sections of the form

$$(3) \quad q \mapsto V_q = \text{pt}_{p,q} v_p$$

from \mathcal{M} to the tangent bundle \mathcal{TM} . Here v_p denotes a tangent vector at $p \in \mathcal{M}$ and the vector field V (which is almost everywhere defined on \mathcal{M}) is given by the parallel transport $\text{pt}_{p,q}$ along the shortest geodesic joining p and the argument q .

The purpose of this note is to derive expressions which are more general than the ones in [3] for the covariant derivative of the section V_q in a nondegenerate point q , (i.e., q is not in the cut locus of p , for details see [4, 7].) In particular, the derived expressions result in eigenvalue problems and linear systems of equations in the setup of general symmetric spaces.

Observations for the parallel transport variant of TGV. In the following we let $\gamma : [0, 1] \rightarrow \mathcal{M}$ be a geodesic connecting the points $p = \gamma(0)$ and $q = \gamma(1)$. We assume that γ is a shortest geodesic and notice that, as such, γ is unique (up to parametrization) for almost all input $p, q \in \mathcal{M}$. We consider a tangent vector $y_q \in \mathcal{TM}$ with base point q . We use y_q to define the geodesic variation $c : [0, 1] \times (-\varepsilon, \varepsilon) \rightarrow \mathcal{M}$ along γ by

$$c(t, s) = [p, \exp_q(sy_q)]_t.$$

Here the bracket $[x, y]_t := \exp_x(t \log_x y)$ denotes the point on the geodesic joining x and y at time t . We note that $c(t, 0) = \gamma(t)$ for all $t \in [0, 1]$. Then we define the Jacobi field $Y(t, s)$ corresponding to this geodesic variation by

$$Y(t, s) = \frac{d}{ds} [p, \exp_q(sy_q)]_t.$$

We recall that the Riemannian curvature tensor R is defined by $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$ for vector fields X, Y, Z . Here $[X, Y]f = XYf - YXf$ denotes the Lie bracket of vector fields. Further, the sectional curvature $K(X, Y)$ is given by $K(X, Y) = \frac{\langle R(X, Y)X, Y \rangle}{\|X\|^2 \|Y\|^2}$.

Lemma 1. *Let \mathcal{M} be a Riemannian manifold. Then the covariant derivative of the vector field $\nabla_Y V$, V given by (3), along γ is given by $\nabla_{\dot{\gamma}(t)} \nabla_Y V = R(\dot{\gamma}(t), Y)V$. Further, the covariant derivative $\nabla_{y_q} V$ of the vector field V in the point q in direction y_q is given by*

$$(4) \quad \langle \nabla_{y_q} V, Z \rangle = \int_0^1 \langle R(\dot{\gamma}(t), Y)V, Z_{\gamma(t)} \rangle dt,$$

where $Z_r = \text{pt}_{q,r} z_q$ for a test vector z_q at the point q .

Proof. We consider the vector field V along the geodesic variation c , i.e., the mapping $V'(t, s) := V_{c(t,s)}$. Using the symbols $\frac{D}{dt}, \frac{D}{ds}$, for the covariant derivative w.r.t. the arguments t, s , we have $\nabla_{\dot{\gamma}(t)} \nabla_Y V = \frac{D}{dt} \frac{D}{ds} |_{s=0} V'(t, s)$. Then,

$$(5) \quad \frac{D}{dt} \frac{D}{ds} V'(t, s) = \frac{D}{ds} \frac{D}{dt} V'(t, s) + R\left(\frac{d}{dt} c(t, s), \frac{d}{ds} c(t, s)\right) V'(t, s).$$

By the construction of $c(t, s)$ we have that $\frac{D}{dt} V'(t, s) = 0$ since the covariant derivative of a parallel field equals zero. This yields the equality $\frac{D}{dt} \frac{D}{ds} V'(t, s) = R\left(\frac{d}{dt} c(t, s), \frac{d}{ds} c(t, s)\right) V'(t, s)$. Hence, $\nabla_{\dot{\gamma}(t)} \nabla_Y V = R(\dot{\gamma}(t), Y)V$. Concerning (4) we have

$$\begin{aligned} \langle \nabla_{y_q} V, Z \rangle &= \langle \frac{D}{ds} |_{s=0} V_{c(1,s)}, Z \rangle = \int_0^1 \frac{d}{dt} \langle \frac{D}{ds} |_{s=0} V_{c(t,s)}, Z_{\gamma(t)} \rangle dt \\ &= \int_0^1 \left(\langle \frac{D}{dt} \frac{D}{ds} |_{s=0} V_{c(t,s)}, \text{pt}_{q,\gamma(t)} z_q \rangle + \langle \frac{D}{ds} |_{s=0} V_{c(t,s)}, \frac{D}{dt} \text{pt}_{q,\gamma(t)} z_q \rangle \right) dt. \end{aligned}$$

We have $\frac{D}{dt} \text{pt}_{q,\gamma(t)} z_q = 0$ which makes the second summand in the bracket vanish. By (5), $\frac{D}{dt} \frac{D}{ds} |_{s=0} V_{c(t,s)} = R(\dot{\gamma}(t), Y(t, 0))V(t, 0)$ from which we conclude the validity of (4). □

In a symmetric space this expression can be made more explicit.

Theorem 2. *Let \mathcal{M} be a Riemannian symmetric space. Let y_q be an eigenvector of the Riemannian curvature operator $X \mapsto R(\dot{\gamma}(1), X)\dot{\gamma}(1)$. Then, the covariant derivative $\nabla_{y_q} V$ of the parallel transport along γ is given by*

$$(6) \quad \nabla_{y_q} V = F(K \|y_q\|^2) R(\dot{\gamma}(1), y_q) V_q$$

where $K = K(\dot{\gamma}(1), y_q)$ denotes the sectional curvature of the submanifold spanned by $\dot{\gamma}(1), y_q$ and the function F is given by $F(\lambda) = \frac{1 - \cos \sqrt{\lambda}}{\sqrt{\lambda} \sin \sqrt{\lambda}}$, if $\lambda > 0$, $F(\lambda) = \frac{\cosh(\sqrt{-\lambda}) - 1}{\sqrt{-\lambda} \sinh \sqrt{-\lambda}}$ if $\lambda < 0$, and by $F(\lambda) = 1/2$ if $\lambda = 0$.

Proof. We start from (4) and first invoke the fact, that, in a symmetric space, the Jacobi field $Y(t, 0)$ is given by

$$(7) \quad Y(t, 0) = f(\lambda_q, t) \text{pt}_{q,\gamma(t)} y_q$$

where $\lambda_q = K(\dot{\gamma}(1), y_q) \|y_q\|^2$ denotes the eigenvalue of the curvature operator $X \mapsto R(\dot{\gamma}(1), X)\dot{\gamma}(1)$ for the eigenvector y_q ; further, the function f is, depending on the sign of λ , given by $f(\lambda, t) = \sin(\sqrt{\lambda}t)/\sin\sqrt{\lambda}$, if $\lambda > 0$, $f(\lambda, t) = \sinh(\sqrt{-\lambda}t)/\sinh\sqrt{-\lambda}$ if $\lambda < 0$, $f(\lambda, t) = t$ if $\lambda = 0$. Now, using (7) in (4), we get, for an arbitrary tangent vector z_q at the point q ,

$$\begin{aligned} \langle \nabla_{y_q} V, z_q \rangle &= \int_0^1 \langle R(\dot{\gamma}(t), f(\lambda_q, t) \text{pt}_{q, \gamma(t)} y_q) V, \text{pt}_{q, \gamma(t)} z_q \rangle dt \\ &= \int_0^1 f(\lambda_q, t) \langle R(\text{pt}_{\gamma(t), q} \dot{\gamma}(t), y_q) V_q, z_q \rangle dt = \langle R(\dot{\gamma}(1), y_q) V_q, z_q \rangle F(\lambda_q). \end{aligned}$$

The second equality is a consequence of the invariance of the curvature tensor R under parallel transport in a symmetric space. Hence, the scalar product in the second integral is constant which implies the last equality together with noting that F is the integral w.r.t. f on $[0, 1]$. \square

Remark 3. Concerning the implementation of Formula (6), we note that the time consuming part for computing $F(K \|y_q\|^2)$ is computing $K \|y_q\|^2$. Since these arguments are also needed in other parts of the scheme of [3], they may be reused. For the computation of $R(\dot{\gamma}(1), y_q)v_q$ there are explicit formulas available in many symmetric spaces; further, it does not have to be carried for each y_q of a basis of \mathcal{TM}_q ; a related computation is needed only once by linearity.

Acknowledgement. We acknowledge a valuable discussion on the topic of this report with Xavier Pennec at the Oberwolfach Conference on “Nonlinear Data: Theory and Algorithms.”

REFERENCES

- [1] M. Bačák, R. Bergmann, G. Steidl, and A. Weinmann. A second order non-smooth variational model for restoring manifold-valued images. *SIAM Journal on Scientific Computing*, 38(1):A567–A597, 2016.
- [2] R. Bergmann, F. Laus, G. Steidl, and A. Weinmann. Second order differences of cyclic data and applications in variational denoising. *SIAM Journal on Imaging Sciences*, 7(4):2916–2953, 2014.
- [3] K. Bredies, M. Holler, M. Storath, and A. Weinmann. Total generalized variation for manifold-valued data. *arXiv preprint arXiv:1709.01616*, 2017.
- [4] J. Cheeger and D. Ebin. *Comparison theorems in Riemannian geometry*, volume 9. North-Holland, 1975.
- [5] D. Cremers and E. Strekalovskiy. Total cyclic variation and generalizations. *Journal of Mathematical Imaging and Vision*, 47(3):258–277, 2013.
- [6] P. Grohs and M. Sprecher. Total variation regularization on Riemannian manifolds by iteratively reweighted minimization. *Information and Inference*, 5(4):353–378, 2016.
- [7] J. Itoh and M. Tanaka. The dimension of a cut locus on a smooth riemannian manifold. *Tohoku Mathematical Journal, Second Series*, 50(4):571–575, 1998.
- [8] J. Lellmann, E. Strekalovskiy, S. Koetter, and D. Cremers. Total variation regularization for functions with values in a manifold. In *International Conference on Computer Vision (ICCV)*, pages 2944–2951, 2013.

- [9] E. Strelakovski and D. Cremers. Total variation for cyclic structures: Convex relaxation and efficient minimization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1905–1911, 2011.
- [10] A. Weinmann, L. Demaret, and M. Storath. Total variation regularization for manifold-valued data. *SIAM Journal on Imaging Sciences*, 7(4):2226–2257, 2014.

Friends of Grassmannians

LEK-HENG LIM

We discuss various objects connected to the Grassmannian – Schubert varieties, Sato Grassmannians, doubly-infinite Grassmannian, affine Grassmannians, and flag manifolds. We show how the first three objects may be used to define natural distances on subspaces of different dimensions and how the last two objects could serve as immensely useful platforms for problems in multivariate statistical analysis and signal processing.

Local and Non-Local Shape Analysis

JULIE DIGNE

(joint work with Sébastien Valette, Raphaëlle Chaine, Yohann Béarzi)

In this talk we explore two contributions for shape analysis. In a first case, we consider surfaces and how local analysis of the angular oscillations and polynomial radial behavior around surface points leads to accurate normal estimation and new integral invariants. A direct application of these integral invariants is geometric detail exaggeration. This however assumes that the surface can be represented as a height function over some parameterization plane in a neighbourhood of *fixed* radius that is the same for all the surface.

In many cases, however, shapes, as they are acquired by laser scanners, might not fulfill this hypothesis: they can have isotropically sampled areas or curve parts. For example, street cables can be considered as curves, depending on the acquisition accuracy. We call this case the *mixed dimension case*. Armed with a well-defined probing operator associating a point of the ambient space to a point on the shape, we define Local Probing Fields, and analyze them in a non-local manner. Hence, we are able to extract and describe data self-similarities. Exploiting these allows us to revisit various surface processing tasks such as denoising, compression and shape resampling.

Let us first consider the first case, and assume we have a smooth surface \mathcal{S} . Many shape processing methods need accurate surface derivatives estimates. Surface derivatives are indeed useful to estimate important shape features such as normals or curvatures. The signal processing viewpoint is slightly different: instead of analyzing signal derivatives, signals are often processed using a frequency analysis and by devising filters operating on the Fourier coefficients. To bring together these two trends, we describe a function basis taking into account both the local surface derivatives and the angular oscillations around each point of the

surface [2]. This formulation, which we term Wavejets, in reference to their fully polynomial counterpart, the Osculating Jets [1], gives valuable information on the shape.

Given a local polar coordinate system (r, θ) in the tangent plane of a sample p of a smooth surface, a Wavejet is the set of complex coefficients $\phi_{k,n}$ such that the height field describing the surface around p is written as follows:

$$f(r, \theta) = \sum_{k=0}^{\infty} \sum_{n=-k}^k r^k \phi_{k,n} e^{in\theta} = \sum_{n=-\infty}^{\infty} \sum_{k=|n|}^{\infty} r^k \phi_{k,n} e^{in\theta}$$

Wavejets require a local parameterization. We demonstrate theoretically that if this parameterization is not done over the tangent plane but over a parameterization plane with a small angle to the true tangent plane, it is possible to correct the plane. It is also possible to correct the coefficients to those computed over the tangent plane. Furthermore, Wavejets coefficients can be used to compute interesting indicators of differential volumes that can be an alternative to using curvatures and further surface derivatives (see Figure 1), those indicators are new integral invariants that can be efficiently computed.

From a practical point of view, given a point set acquired from an object surface, it is possible to estimate Wavejets coefficients by least squares regression. The new integral invariants are then used to filter the shapes. We show two applications of these filters working directly on point sets, shape detail exaggeration, either by working on point positions or by working on points normals.

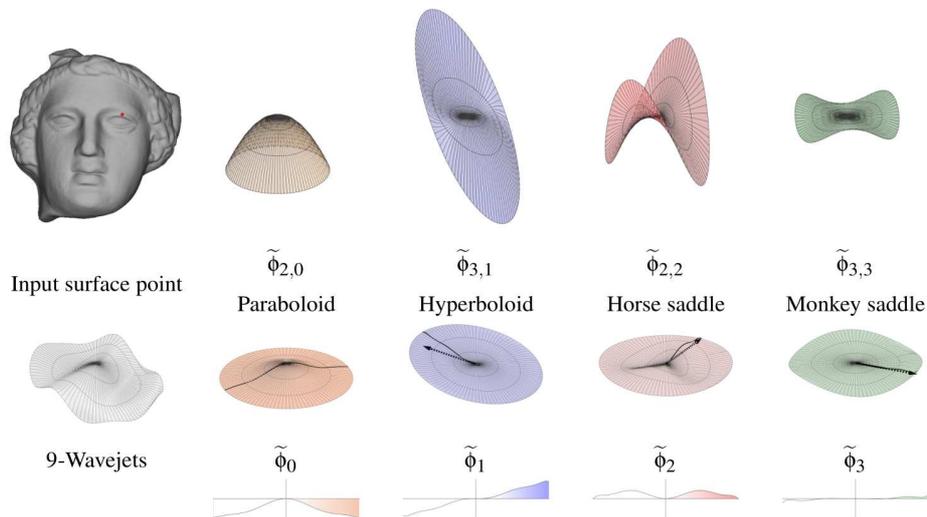


FIGURE 1. Wavejets decomposition around a point of a surface. Left: approximated 9-Wavejets surface. Let $\tilde{\phi}_{k,n}(r, \theta) = r^k (\phi_{k,n} e^{in\theta} + \phi_{k,-n} e^{-in\theta})$ and $\tilde{\phi}_n = \sum_{k=0}^{\infty} \tilde{\phi}_{k,n}$.

The second part of this talk deals with the mixed dimension case [3]. Mixed dimension shapes cannot be described everywhere via local height maps over some

parameterization plane [4]. We introduce a different type of descriptor that can reliably represent curves or isotropic shape parts while mimicking a height map in areas where the surface assumption holds. By analyzing jointly these vector fields through a nonlocal similarity optimization process, the descriptors of similar areas tend to align and reveal the similarity, if any.

Let us consider a probing operator \mathcal{P} mapping a point p in the ambient space to a point q on the shape (\mathcal{P} can be the nearest point projector for example). Given a generic pattern of points $(u_i)_{i=1\dots M}$ parameterized by a local coordinate frame and located at an anchor point s in the ambient space but *close to* the shape, we encode the set of deformation vectors $v_i = \mathcal{P}(s + u_i) - (s + u_i)$ as a vector $V \in \mathbb{R}^{3M}$. The following process then aims at enhancing the similarities between the descriptors by optimizing the positions of the anchors and orientations of the sampling pattern:

$$\begin{aligned} \min_{V, D, \alpha} \quad & \sum_{j=0}^{N-1} \|V_j - D\alpha_j\|_2^2 + \lambda \|\alpha\|_1 \\ \text{s.t.} \quad & \{V_j\}_{j=0\dots N-1} \text{ cover the whole shape} \\ & D \in \mathbb{R}^{3M \times d}, V_j \in \mathbb{R}^{3M} \end{aligned}$$

In a nutshell, we optimize the positions and orientations of the sampling pattern so that the local probing fields V_j are efficiently encoded on a dictionary D with d atoms. λ is a parameter giving a loose control over the sparsity of the decomposition coefficients α . From this optimized dictionary decomposition, the shape can then be resampled, or denoised.

REFERENCES

- [1] F. Cazals. Estimating differential quantities using polynomial fitting of osculating jets. *Eurographics*, pages 177–187, 2003.
- [2] Y. Béarzi, J. Digne, R. Chaine. Wavejets: A Local Frequency Framework for Shape Details Amplification *Computer Graphics Forum, Proc. Eurographics*, 2018
- [3] J. Digne, S. Valette, R. Chaine, Sparse Geometric Representation Through Local Shape Probing *IEEE TVCG*, 2018.
- [4] J. Digne, R. Chaine, S. Valette Self-similarity for the accurate compression of point set surfaces *Computer Graphics Forum, Proc. Eurographics*, 2014

Efficient Deformable Shape Correspondence via Kernel Matching

AMIT BOYARSKI

(joint work with Matthias Vestner, Zorah Lähler, Or Litany, Ron Slossberg, Tal Remez, Emanuele Rodolà, Alex Bronstein, Michael Bronstein, Ron Kimmel, Daniel Cremers)

The correspondence problem is a fundamental problem lying at the heart of computer vision, graphics and pattern recognition, with applications including shape comparison, texture transfer, and shape interpolation just to name a few. Given

two three-dimensional objects \mathcal{X} and \mathcal{Y} , modeled as compact two-dimensional Riemannian manifolds, a correspondence is a map $\varphi : \mathcal{X} \rightarrow \mathcal{Y}$ assigning points from one domain to points from the other. φ should ideally satisfy a few properties: It should be bijective, continuous in both directions in the sense that nearby points on \mathcal{X} should be mapped to nearby points on \mathcal{Y} (and vice versa), and map between similar points. For the simplicity of the introduction, we assume the two shapes \mathcal{X} and \mathcal{Y} to be sampled at n points each. The proposed framework can be easily extended to handle the case of a different number of samples. Assuming a consistent sampling (e.g. via farthest point sampling with a sufficiently large number n of points), the discrete counterpart to the correspondence φ is a mapping $\pi : \{x_1, \dots, x_n\} \rightarrow \{y_1, \dots, y_n\}$, which admits a representation as a permutation matrix $\mathbf{\Pi} \in \{0, 1\}^{n \times n}$ satisfying $\mathbf{\Pi}^\top \mathbf{1} = \mathbf{\Pi} \mathbf{1} = \mathbf{1}$ with $\mathbf{1}$ being a column vector of ones. We henceforth denote the space of $n \times n$ permutation matrices by \mathcal{P}_n .

A common way to approach the correspondence problem is by phrasing it as a score maximization problem (e.g. [8, 9])

$$\mathbf{\Pi}^* = \arg \max_{\mathbf{\Pi} \in \mathcal{P}_n} E(\mathbf{\Pi}),$$

where $E(\mathbf{\Pi})$ is usually a weighed aggregate of two terms

$$E(\mathbf{\Pi}) = \alpha g(\mathbf{\Pi}) + h(\mathbf{\Pi}).$$

The first term $g(\mathbf{\Pi})$ is a fidelity term trying to align a set of *pointwise descriptors* encoding the similarity between points, while the second term $h(\mathbf{\Pi})$ is a regularization term promoting the continuity of the correspondence by aligning a set of *pairwise descriptors* encoding global/local relations between pairs of points. The parameter α governs the tradeoff between the two terms. The constraint $\mathbf{\Pi} \in \mathcal{P}_n$ guarantees bijectivity of the correspondence, and together with the two terms h and g , provide a trade-off between complexity, fidelity and regularity. To that end, we aim at maximizing the following functional over \mathcal{P}_n ,

$$(1) \quad \arg \max_{\mathbf{\Pi} \in \mathcal{P}_n} E(\mathbf{\Pi}) = \arg \max_{\mathbf{\Pi} \in \mathcal{P}_n} \alpha \langle \mathbf{\Pi}, \mathbf{F}_\mathcal{Y} \mathbf{F}_\mathcal{X}^\top \rangle + \langle \mathbf{\Pi}, \mathbf{K}_\mathcal{Y} \mathbf{\Pi} \mathbf{K}_\mathcal{X} \rangle,$$

where $\mathbf{F}_\mathcal{X}, \mathbf{F}_\mathcal{Y}$ are matrices of point-wise descriptors and $\mathbf{K}_\mathcal{X}, \mathbf{K}_\mathcal{Y}$ are kernel matrices defined on \mathcal{X} and \mathcal{Y} , respectively. We advocate the use of heat kernels, motivated by the fact that they are positive-definite, easy to approximate, and admit an interesting physical interpretation (see below). Similar results can be obtained using geodesic Gaussian kernels [1].

A popular technique to handle problems like (1) is via relaxation. The discrete set \mathcal{P}_n is replaced by a larger continuous set, i.e., the set of doubly-stochastic matrices \mathcal{B}_n , and the solution of the relaxed problem is later projected on \mathcal{P}_n , possibly using Euclidean projection. The choice of the relaxation and projection is usually related to the properties of the optimization problem (i.e., a quadratic assignment problem) rather than the particular choice of descriptors. This leads to poorly understandable algorithms whose justification remains elusive. We stress in our work that taking into account the choice of descriptors can lead to a simpler relaxation scheme, as simple as linear filtering.

Due to the heat kernels being (strictly) positive-definite, (1) is equivalent to the relaxed problem

$$\arg \max_{\mathbf{P} \in \mathcal{B}_n} E(\mathbf{P}).$$

Our algorithm consists of the following projected gradient iteration

$$\mathbf{P}^{k+1} = \arg \max_{\mathbf{P} \in \mathcal{B}_n} \langle \mathbf{P}, \nabla E(\mathbf{P}^k) \rangle,$$

where for our choice of $E(\mathbf{P})$, the gradient is given by

$$\nabla E(\mathbf{P}) = \alpha \mathbf{F}_Y \mathbf{F}_X^\top + \mathbf{K}_Y \mathbf{P} \mathbf{K}_X.$$

Since \mathbf{P}^k is guaranteed to be a permutation matrix, we henceforth use $\mathbf{\Pi}^k$ to denote the iterates themselves, yielding the following step

$$(2) \quad \mathbf{\Pi}^{k+1} = \arg \max_{\mathbf{\Pi} \in \mathcal{P}_n} \langle \mathbf{\Pi}, \alpha \mathbf{F}_Y \mathbf{F}_X^\top + \mathbf{K}_Y \mathbf{\Pi}^k \mathbf{K}_X \rangle$$

that is guaranteed to produce a series on increasing values of $E(\mathbf{\Pi})$. In the experiments presented in [2], we use the data fidelity term $\langle \mathbf{\Pi}, \mathbf{F}_Y \mathbf{F}_X^\top \rangle$, mainly to initialize the process:

$$\mathbf{\Pi}^0 = \arg \max_{\mathbf{\Pi} \in \mathcal{P}_n} \langle \mathbf{\Pi}, \mathbf{F}_Y \mathbf{F}_X^\top \rangle.$$

Algorithm (2) neatly extends to the partial case by replacing the row-sum constraint $\mathbf{\Pi} \mathbf{1} = \mathbf{1}$ with $\mathbf{\Pi} \mathbf{1} \leq \mathbf{1}$ and using slack variables. For more information and proofs of the claims presented here, see [2]. Some results obtained with our algorithm are presented in Figures 1 and 2.

To intuitively understand the efficacy of kernel alignment for the purpose of finding correspondences, consider the k -th iteration (for simplicity, without the data term):

$$\max_{\mathbf{\Pi} \in \mathcal{P}_n} \text{tr} (\mathbf{\Pi}^\top \mathbf{K}_Y \mathbf{\Pi}^k \mathbf{K}_X).$$

Let us denote by δ^j the discrete indicator function of vertex j on shape \mathcal{X} , representing initial heat distribution concentrated at vertex j . This heat is propagated via the application of the heat kernel \mathbf{K}_X to the rest of the vertices, resulting in the new heat distribution on \mathcal{X} given by $\mathbf{k}_X^j = \mathbf{K}_X \delta^j$. This heat distribution, whose spread depends on the time parameter t , is mapped via $\mathbf{\Pi}^k$ onto the shape \mathcal{Y} , where it is propagated via the heat kernel \mathbf{K}_Y . The ij -th element of the matrix $\mathbf{K}_Y \mathbf{\Pi}^k \mathbf{K}_X$,

$$(\mathbf{K}_Y \mathbf{\Pi}^k \mathbf{K}_X)_{ij} = (\mathbf{k}_Y^i)^\top \mathbf{\Pi}^k \mathbf{k}_X^j = \sum_m (\mathbf{K}_Y)_{i, \pi^k(m)} (\mathbf{K}_X)_{jm},$$

represents the probability of a point i on \mathcal{Y} being in correspondence with the point j on \mathcal{X} . This is affected by both the distance between i and $\pi^k(m)$ on \mathcal{Y} for every m on \mathcal{X} , encoded in the entries of $(\mathbf{K}_Y)_{i, \pi^k(m)}$, and by the distance between m and j on \mathcal{X} , encoded in the entries of $(\mathbf{K}_X)_{jm}$. This process, as illustrated in Figure 3, resembles the alternating diffusion process described in [4]. Its success in uncovering the latent correspondence is based on the following statistical assertions on the

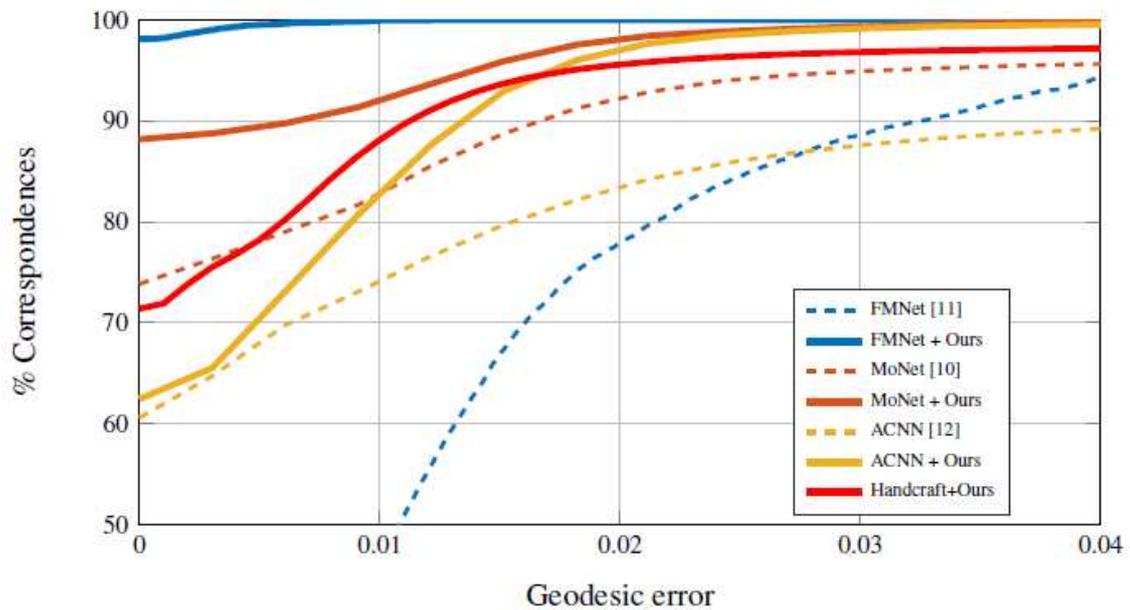


FIGURE 1. Correspondence accuracy on FAUST (template subset). The dashed curves indicate the raw performance of recent deep learning methods, while solid curves are the results obtained after using our method as post-processing. Our method based on handcrafted descriptors (SHOT [3]) is denoted as ‘Handcrafted+Ours’.

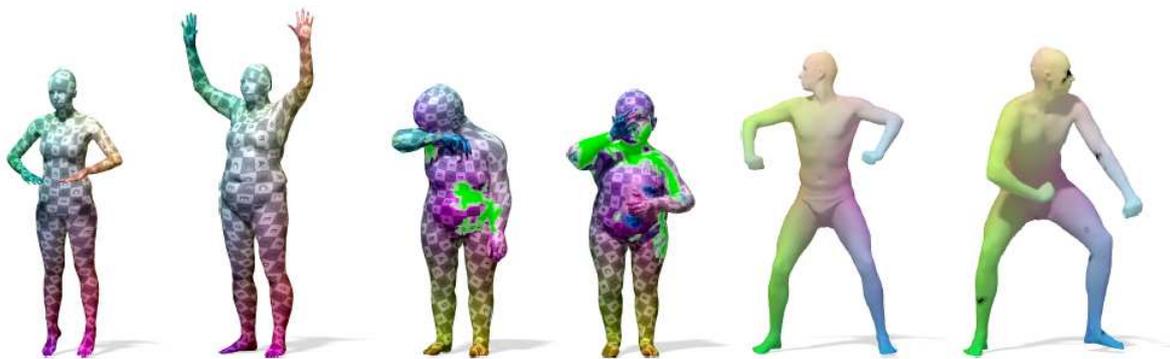


FIGURE 2. Qualitative examples on FAUST [5] models (left), SHREC’16 [6] (middle) and SCAPE [7] (right). In the SHREC experiment, the green parts mark where no correspondence was found. Notice how those areas are close to the parts that are hidden in the other model. The missing matches (marked in black) in the SCAPE experiment are an artifact due to the multiscale approach.

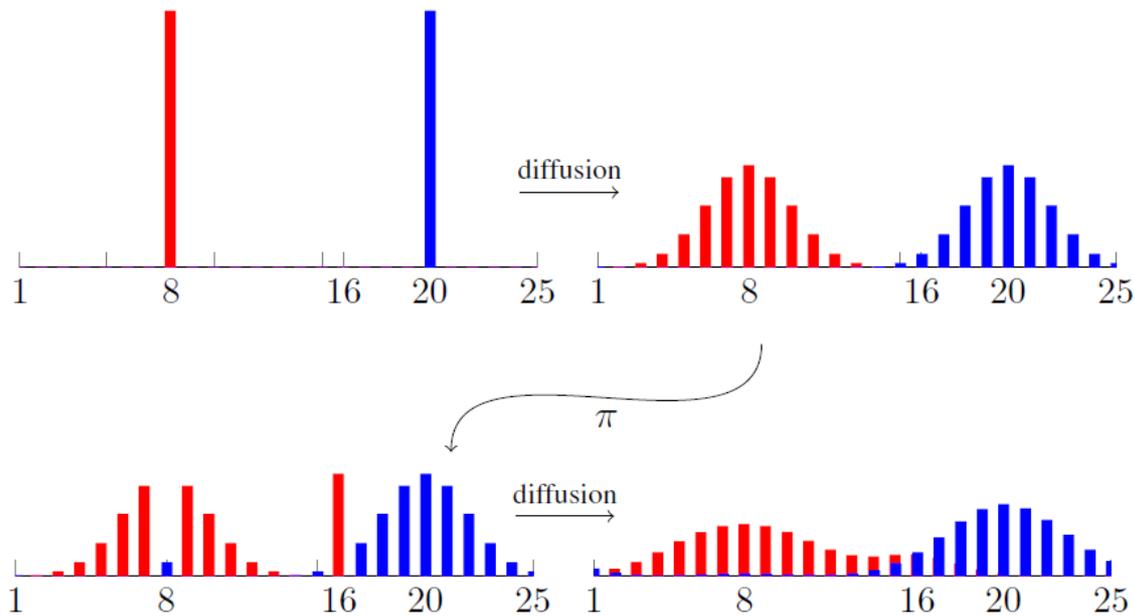


FIGURE 3. **Illustration of the alternating diffusion process** initialized with a noisy correspondence that wrongly maps $\pi(8) = 16$ and $\pi(16) = 8$ but correctly maps $\pi(x) = x$ elsewhere. Top left: Indicator functions on the source shape, one on a point with a wrong correspondence (red) and one with a correct correspondence (blue). Top right: Both indicator functions are diffused. Bottom left: The diffused functions are transported to the target shape via π . Notice how the red maximum is in the wrong position. Bottom right: Diffusion on the target shape. Now the maximum of the previously wrong matched point (red) is at the correct position.

distribution of correspondences in the initial assignment: we tacitly assume that a sufficiently large number of (uniformly distributed) points are initially mapped correctly while the rest are mapped randomly, such that when averaging over their “votes” they do not bias towards any particular candidate. These concepts will be presented more rigorously in future research.

REFERENCES

- [1] Vestner, Matthias, et al. *Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space*, Proceedings of CVPR. 2017.
- [2] Vestner, Matthias, et al. “Efficient Deformable Shape Correspondence via Kernel Matching.” arXiv preprint arXiv:1707.08991 (2017).
- [3] Tombari, Federico, Samuele Salti, and Luigi Di Stefano. “Unique signatures of histograms for local surface description.” European conference on computer vision. Springer, Berlin, Heidelberg, 2010.
- [4] Lederman, Roy R., and Ronen Talmon. “Learning the geometry of common latent variables using alternating-diffusion.” Applied and Computational Harmonic Analysis (2015).

- [5] Bogo, Federica, et al. *FAUST: Dataset and evaluation for 3D mesh registration*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.
- [6] Cosmo, L., et al. *SHREC'16: Partial matching of deformable shapes*. Proc. 3DOR 2.9 (2016): 12.
- [7] Anguelov, Dragomir, et al. *SCAPE: shape completion and animation of people*. ACM transactions on graphics (TOG). Vol. 24. No. 3. ACM, 2005.
- [8] Bronstein, Alexander M., Michael M. Bronstein, and Ron Kimmel. *Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching*. Proceedings of the National Academy of Sciences 103.5 (2006): 1168-1172.
- [9] Chen, Qifeng, and Vladlen Koltun. *Robust nonrigid registration by convex optimization*. Proceedings of the IEEE International Conference on Computer Vision. 2015.
- [10] Monti, Federico, et al. *Geometric deep learning on graphs and manifolds using mixture model CNNs*. Proc. CVPR. Vol. 1. No. 2. 2017.
- [11] Litany, Or, et al. *Deep functional maps: Structured prediction for dense shape correspondence*. Proc. ICCV. Vol. 2. 2017.
- [12] Boscaini, Davide, et al. *Learning shape correspondence with anisotropic convolutional neural networks*. Advances in Neural Information Processing Systems. 2016.

Structural Optimization with Cosserat Phase Field Modeling

INGO MUENCH

We present a continuum-type optimality algorithm for the evolution of load-bearing solid structures with linear Cosserat theory. The model combines two sensitivity functions to homogenize equivalent stress and equivalent couple stress, too. The phase field variable φ defines the density and stiffness of the substance

$$(1) \quad f(\varphi) = \frac{e^{\alpha \varphi}}{e^{\alpha \varphi} + 1}, \quad \rho_\varphi = f(\varphi) \rho_0, \quad \mathbb{C}_\varphi = f(\varphi) \mathbb{C}_0, \quad \mathbb{C}_\varphi^{\text{Coss}} = f(\varphi) \mathbb{C}_0^{\text{Coss}},$$

where ρ_0 , \mathbb{C}_0 , and $\mathbb{C}_0^{\text{Coss}}$ represent material and $f(\varphi) \rightarrow 0$ yield voids. The continuous function $f(\varphi)$ increases monotonously with lower and upper limits

$$\lim_{\varphi \rightarrow -\infty} \frac{e^{\alpha \varphi}}{e^{\alpha \varphi} + 1} = 0, \quad \lim_{\varphi \rightarrow +\infty} \frac{e^{\alpha \varphi}}{e^{\alpha \varphi} + 1} = 1, \quad \text{for } \alpha > 0.$$

We allocate the interval $[-1, 1]$ for the phase field variable φ to describe the transition from void to material and choose the exponent $\alpha = 14$ to obtain $f(-1) \approx 1 \cdot 10^{-6}$, and $f(1) \approx 1$. Double-well potential and gradient energy

$$(2) \quad \psi_w = \varphi^6 - \varphi^4 - \varphi^2 + 1, \quad \psi_g = \frac{1}{2} L_c \|\text{Grad}[\varphi]\|^2 dV,$$

specify the phase field part of the model. The standard kinematical equations for the linear Cosserat model are given via displacements u and rotations α reading

$$\varepsilon = \text{Grad}[u] + \epsilon \cdot \alpha, \quad \kappa = \text{Grad}[\alpha].$$

The inner strain and curvature energy of material are considered by

$$\psi_s = \mu \|\text{sym } \varepsilon\|^2 + \mu_c \|\text{skew } \varepsilon\|^2 + \frac{\lambda}{2} (\text{tr}[\varepsilon])^2, \quad \psi_c = \mu X_c \|\text{dev sym } \kappa\|^2.$$

Therefore, the inherent stiffness of the substance in eq.(1) is given by

$$\mathbb{C}_0 = \frac{\partial^2 \psi_s}{\partial \varepsilon^2}, \quad \mathbb{C}_0^{\text{Coss}} = \frac{\partial^2 \psi_c}{\partial \kappa^2}.$$

The overall inner energy of the system is considered

$$(3) \quad \Pi_i = \int_{\mathcal{B}} \underbrace{\psi_w + \psi_g - \frac{1}{2} f(\varphi) (\psi_s + \psi_c)}_{\psi} dV.$$

In accordance to elasticity we define stress and couple stress by

$$(4) \quad \sigma = -\frac{\partial \psi}{\partial \varepsilon} = f(\varphi) \mathbb{C}_0 : \varepsilon, \quad \mathbf{m} = -\frac{\partial \psi}{\partial \kappa} = f(\varphi) \mathbb{C}_0^{\text{Coss}} : \kappa.$$

Phase field sensitivities read

$$(5) \quad \eta = \frac{\partial \psi}{\partial \varphi} = -\frac{1}{2} \underbrace{f'(\varphi)}_{>0} \underbrace{[\varepsilon : \mathbb{C}_0 : \varepsilon + \kappa : \mathbb{C}_0^{\text{Coss}} : \kappa]}_{>0} + \underbrace{6\varphi^5 - 4\varphi^3 - 2\varphi}_{\eta_\varphi},$$

$$\xi = \frac{\partial \psi}{\partial \text{Grad}[\varphi]} = L_c \text{Grad}[\varphi].$$

The derivative of the double well potential in eq.(2)₁ concerning φ is defined as η_φ in eq.(5). For $-1 < \varphi < 0$ one obtains $\eta_\varphi > 0$. It indicates that $\eta > 0$ yields the evolution of voids. Therefore, we have designed the first term in eq.(5) such that substance generate in regions with eminent stress or couple stress. Finally, it motivates our definition of ψ in eq.(3), and the unusual sign in eq.(4).

The outer work of our model includes volume forces $\rho_\varphi b$, volume couples $\rho_\varphi c$, and a source of material γ within the body. Such a source of material implies that we do not conserve mass. However, this is inspired by bio-mechanical processes, e.g., the formation of bones. Within a phenomenological formulation, the variation of mineral material in the periosteum may be captured by such a source term γ . Further, we consider the injection or rejection of material y on the surface $\partial \mathcal{B}$ of the design space \mathcal{B} . Similarly, mechanical loading on $\partial \mathcal{B}$ is given by tractions t and surface couples d . In summary, we define

$$(6) \quad \Pi_a = \int_{\mathcal{B}} \rho_\varphi b \cdot u + \rho_\varphi c \cdot \alpha + \gamma \varphi dV + \int_{\partial \mathcal{B}} t \cdot u + d \cdot \alpha + y \varphi dV.$$

From the principle of virtual work one can derive the Euler equations of the model. It is the balance of linear momentum

$$\text{Div } \sigma + f(\varphi) \rho_0 b = 0,$$

the balance of angular momentum

$$\text{Div } \mathbf{m} - \sigma : \varepsilon + f(\varphi) \rho_0 c = 0,$$

and the balance equation for the phase field parameter

$$(7) \quad \eta - \text{Div } \xi + f'(\varphi) \rho_0 b \cdot u + f'(\varphi) \rho_0 c \cdot \alpha + \gamma = 0.$$

Generally, the phase field parameter does not satisfy eq.(7) in the initial configuration of a system. Allen and Cahn [1] suggested to extend such balance equations by the rate of φ yielding

$$\eta - \text{Div } \xi + f'(\varphi) \rho_0 b \cdot u + f'(\varphi) \rho_0 c \cdot \alpha + \gamma + \beta \dot{\varphi} = 0.$$

We use Galerkin’s method to set up

$$G = \int_{\mathcal{B}} - (\text{Div } \sigma + f(\varphi) \rho_0 b) \cdot \delta u - (\text{Div } \mathbf{m} - \sigma : \epsilon + f(\varphi) \rho_0 c) \cdot \delta \alpha + (\eta - \text{Div } \xi + f'(\varphi) \rho_0 b \cdot u + f'(\varphi) \rho_0 c \cdot \alpha + \gamma + \beta \dot{\varphi}) \delta \varphi \, dV.$$

Linearization of $G = 0$, backward Euler time integration, and the standard finite element method yield a linear system of equation for increments Δu , $\Delta \alpha$, and $\Delta \varphi$ for the Newton-Raphson scheme. More information about numerical aspects are given in [2]. In the context of a linear Cosserat model the update of rotations is additive.

Generally, one has to consider an initial distribution of substance within the design space \mathcal{B} to set up the boundary value problem. In our case it is the specification of the phase field variable $\varphi = \varphi_0$. Obviously, $\varphi_0 \equiv 0$ within \mathcal{B} is a neutral choice. Then, we obtain $f(0) = 1/2$ and begin the evolution of a structure with homogeneously distributed substance having 50% density and stiffness compared to the material phase. Let us define the filling level

$$(8) \quad K = \frac{\int_{\mathcal{B}} f(\varphi) \, dV}{\int_{\mathcal{B}} dV} \in [0, 1].$$

In case of $\varphi_0 \equiv 0$ we obtain the initial filling level $K_0 = 0.5$. Since we do not conserve mass the first in term in eq.(5) will lead to $K > K_0$ if there is mechanical deformation within \mathcal{B} . To control the filling level within \mathcal{B} we need to introduce a side condition into our formulation. Therefore, we couple γ in eq.(6) to the desired filling level \hat{K} of the system. Since the filling level is given as global quantity by eq.(8) we resign consistent variation and linearization of γ in the above equations. Thus, the material source γ becomes a deterministic variable. Further, we account for scalar norms of stress and couple stress, which we denote σ_V and \mathbf{m}_V . Then, we define thresholds $\bar{\sigma}_V$ and $\bar{\mathbf{m}}_V$ with help of the following algorithm: σ_V and \mathbf{m}_V in all integration points n_{GP} at position x_{GP} are stored into arrays

$$g_i = \sigma_V(x_{GP}), \quad h_i = \sigma_V(x_{GP}), \quad i = 1 \dots n_{GP}.$$

Then, the components of g_i and h_i are sorted in ascending order, e.g.,

$$g_i \mapsto g_j, \quad g_{j+1} \geq g_j, \quad i, j = 1 \dots n_{GP}.$$

With help of the rounded index

$$k = \text{int} \left[\left(1 - \hat{K} \right) \cdot n_{GP} \right],$$

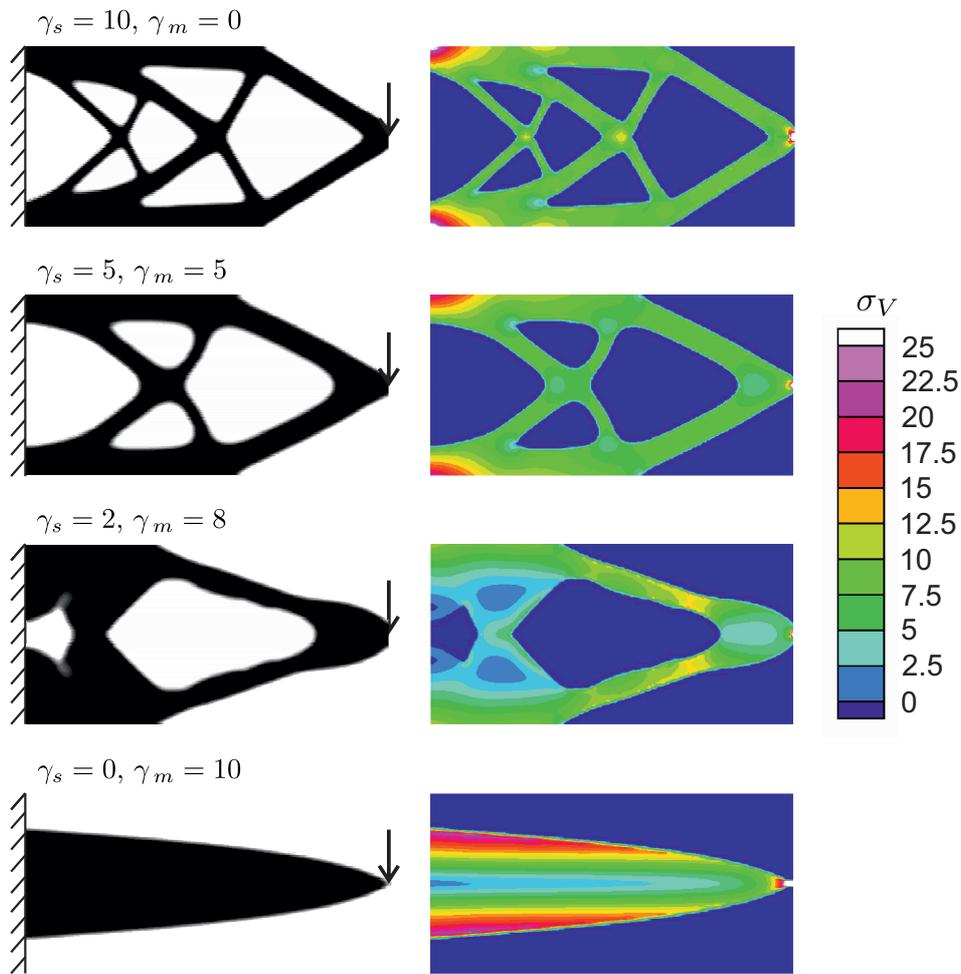
$2n + 1$ components of g_j and h_j are used to define

$$\bar{\sigma}_V := \sum_{j=k-n}^{j=k+n} \frac{g_j}{2n+1}, \quad \bar{\mathbf{m}}_V := \sum_{j=k-n}^{j=k+n} \frac{h_j}{2n+1}.$$

We found that $n = \text{int}[0.001 n_{GP}]$ reduces random noise in $\bar{\sigma}_V$ and $\bar{\mathbf{m}}_V$ during the evolution process to an acceptable level. We treat $\bar{\sigma}_V$ and $\bar{\mathbf{m}}_V$ as constants within each time step. It is computationally expensive and even unfavorable to update these thresholds after each Newton-Raphson iteration step. We use both thresholds to define the objective function

$$F = \int_{\mathcal{B}} \underbrace{\gamma_s \frac{\bar{\sigma}_V - \sigma_V}{\bar{\sigma}_V} + \gamma_m \frac{\bar{\mathbf{m}}_V - \mathbf{m}_V}{\bar{\mathbf{m}}_V}}_{=\gamma} \varphi \, dV \rightarrow \min \text{ w.r.t. } u, \alpha, \varphi.$$

In the following example we show the effect of couple stresses onto the evolution of topologies. Let us consider a rectangular design space of length 40, height 20, numerically treated as plate of thickness 1. It is clamped at the left border and loaded by a single force $P = 20$ at in the mid of the left border. We fix the constitutive parameters $L_c = 1$, $\mu = \mu_c = 5000$, $\lambda = 0$, $X_c = 1$ and vary the weight of σ_V and \mathbf{m}_V onto the source of material γ by factors γ_s and γ_m . For the variant $\gamma_s = 10$, $\gamma_m = 0$ we obtain a framework with 12 straight-lined trusses. The equivalent stress field is almost homogeneous. However, one can find minor stress concentrations at intersections. Further, major stress concentrations are at the point of loading and at the boundary. For $\gamma_s = 5$, $\gamma_m = 5$ the number of trusses is reduced to 8 and we observe curved bars. The stress concentrations at intersections have almost vanished. Actually, the stress concentrations at the loading and at the boundary become weaker. Considering $\gamma_s = 2$, $\gamma_m = 8$ the homogenization of the equivalent stress becomes less important than the homogenization of the equivalent couple stress. Therefore, the structure is not a classical framework but a combination of bars and plates. Finally, for $\gamma_s = 0$, $\gamma_m = 10$ the algorithm evolves a cantilever beam, where the equivalent stress is inhomogeneous. However, it is obvious that the proposed algorithm yields a wide spectrum of structural solutions, which cannot be found without the Cosserat continuum modeling.



REFERENCES

- [1] Cahn, J. and Allen, S., *A microscopic theory for domain wall motion and its experimental verification in Fe-Al Alloy domain growth kinetics*, Journal de Physique Colloques **38**(C7) (1977), 51–54.
- [2] Muench, I., *Evolution of load-bearing structures with phase field modeling*, to appear in Numerical Mathematics and Advanced Applications.

Riemannian Optimization with Low-Rank Tensors

DANIEL KRESSNER

(joint work with Michael Steinlechner, Bart Vandereycken)

A very active research area during the last decade, optimization on matrix manifolds [1] has been applied to a wide variety of applications. Recently, optimization on tensor manifolds has been considered and is increasingly applied in settings where existing algorithms based, e.g., on alternating optimization exhibit limitations.

A tensor is a multivariate array $\mathcal{X} \in \mathbb{R}^{n_1 \times \dots \times n_d}$. For various low-rank tensor formats, including the Tucker, tensor train, and hierarchical Tucker format, it has been shown that the set of tensors of fixed ranks (with respect to the corresponding

format) forms a smooth embedded submanifold of $\mathbb{R}^{n_1 \times \dots \times n_d}$; see [5, 7, 13, 14]. Combined with the highly compressed representation in such a format, this yields efficient Riemannian optimization algorithms.

1. INGREDIENTS OF RIEMANNIAN OPTIMIZATION

To illustrate the ingredients needed for Riemannian optimization, let us consider a tensor in Tucker decomposition:

$$(1) \quad \mathcal{X} = \mathcal{C} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d,$$

where $\mathcal{C} \in \mathbb{R}^{k_1 \times \dots \times k_d}$ is the core tensor, $U_1 \in \mathbb{R}^{n_1 \times k_1}, \dots, U_d \in \mathbb{R}^{n_d \times k_d}$, and \times_μ denotes μ -mode matrix multiplication [8]. The multilinear rank is the smallest possible tuple $\mathbf{k} = (k_1, \dots, k_d)$ for which \mathcal{X} admits a representation of the form (1). The fact that k_μ equals the rank of the matricization

$$X_{(\mu)} \in \mathbb{R}^{n_\mu \times \prod_{\nu \neq \mu} n_\nu}$$

is essential in showing that $\mathcal{M}_{\mathbf{k}}$, the set of tensors of fixed multilinear rank, forms an embedded submanifold.

Because of the multilinearity of the format (1), the tangent space of $\mathcal{M}_{\mathbf{k}}$ at \mathcal{X} takes the form

$$T_{\mathcal{X}}\mathcal{M}_{\mathbf{k}} = \left\{ \mathcal{G} \times_{\mu=1}^d U_\mu + \sum_{\mu=1}^d \mathcal{C} \times_\mu V_\mu \times_{\nu \neq \mu} U_\nu \mid V_\mu^T U_\mu = 0 \right\},$$

where $\mathcal{G} \in \mathbb{R}^{k_1 \times \dots \times k_d}$ and $V_\mu \in \mathbb{R}^{n_\mu \times k_i}$ are the free parameters [7]. Among others, this allows to conveniently express $P_{T_{\mathcal{X}}\mathcal{M}_{\mathbf{k}}}$, the orthogonal projection onto the tangent space.

Consider an optimization problem of the form

$$(2) \quad \min_{\mathcal{X} \in \mathcal{M}_{\mathbf{k}}} f(\mathcal{X})$$

with a smooth function $f : \mathbb{R}^{n_1 \times \dots \times n_d} \rightarrow \mathbb{R}$. A basic first-order method for solving (2), Riemannian steepest descent aims at improving an iterate \mathcal{X}_i by taking a step in the direction of the Riemannian gradient, which equals the orthogonal projection of the Euclidean gradient onto $T_{\mathcal{X}}\mathcal{M}_{\mathbf{k}}$:

$$\mathcal{X}_i + \alpha_i P_{T_{\mathcal{X}}\mathcal{M}_{\mathbf{k}}}(\nabla f(\mathcal{X}_i)).$$

Such a step, in general, leaves the manifold. While the exponential map is the canonical way of mapping from a tangent space back to the manifold, it is often computationally too expensive. The more general class of retractions [1] is equally suitable in the context of optimization and often includes significantly cheaper alternatives. In [9], we have proven that the higher-order SVD (HOSVD), an SVD-based quasi-optimal compression [3], yields a retraction. Letting the HOSVD-based compression to multilinear rank \mathbf{k} be denoted by $\mathcal{T}_{\mathbf{k}}$, one step of Riemannian steepest descent becomes

$$(3) \quad \mathcal{X}_{i+1} \leftarrow \mathcal{T}_{\mathbf{k}}(\mathcal{X}_i + \alpha_i P_{T_{\mathcal{X}}\mathcal{M}_{\mathbf{k}}}(\nabla f(\mathcal{X}_i))).$$

As demonstrated in [9], the standard Armijo backtracking scheme for choosing $\alpha_i > 0$ greatly benefits from being initialized with the step size obtained from a linearized line search procedure.

More complex first-order Riemannian optimization methods may require additional ingredients. In particular, Riemannian conjugate gradient (CG) requires the combination of elements from different tangent spaces, which in turn requires to transport elements from one tangent space to another. For the Tucker format, such a vector transport can be implemented efficiently [9].

The convergence analysis (3) is complicated by the fact that the method may converge to a tensor of lower multilinear rank. This can be avoided, e.g., by a suitable regularization of f . In practice, such a regularization is not needed and also in theory this can be circumvented more elegantly by considering instead of $\mathcal{M}_{\mathbf{k}}$ the algebraic variety of tensors of multilinear rank *at most* \mathbf{k} ; see [11].

The discussion above focuses on the Tucker format but it extends in a seamless manner to the hierarchical Tucker [2] and the tensor train [12] formats.

2. APPLICATION TO TENSOR COMPLETION

The goal of tensor completion is to fill in missing entries of a partially known tensor under a low-rank constraint. When considering tensors of multilinear rank \mathbf{k} , one can formulate tensor completion as the Riemannian optimization problem

$$(4) \quad \min_{\mathcal{X} \in \mathcal{M}_{\mathbf{k}}} \frac{1}{2} \|\Pi_{\Omega} \mathcal{X} - \Pi_{\Omega} \mathcal{A}\|^2,$$

where \mathcal{A} contains the known entries and

$$\Pi_{\Omega} \mathcal{X} := \begin{cases} \mathcal{X}_{i_1 i_2 \dots i_d} & \text{if } (i_1, i_2, \dots, i_d) \in \Omega, \\ 0 & \text{otherwise,} \end{cases}$$

with the sampling set $\Omega \subset [1, n_1] \times \dots \times [1, n_d]$. Riemannian optimization applies in a straightforward manner to (4). In particular, the Riemann gradient at \mathcal{X}_i is obtained by projecting $\Pi_{\Omega} \mathcal{X} - \Pi_{\Omega} \mathcal{A}$ onto the tangent space. In [9], we have demonstrated for a variety of applications, including the recovery of multidimensional images and the approximation of multivariate functions, that Riemannian CG performs very well in this context, in terms of recovery, robustness, and efficiency. More recent developments include an effective preconditioned quotient manifold approach [6] and a Riemannian trust-region method [4] for low-rank tensor completion.

3. APPLICATION TO LARGE LINEAR SYSTEMS

The numerical solution of partial differential equations on high-dimensional domains gives rise to computationally challenging linear systems. When using standard discretization techniques, the size of the linear system grows exponentially with the number of dimensions, making the use of classic iterative solvers unfeasible. Under certain conditions, e.g., when using tensorized FEM on a hypercube,

this linear system can be rephrased as

$$(5) \quad \mathcal{A}(\mathcal{X}) = \mathcal{B},$$

for a given linear operator $\mathcal{A} : \mathbb{R}^{n_1 \times \dots \times n_d} \rightarrow \mathbb{R}^{n_1 \times \dots \times n_d}$ and right-hand side \mathcal{B} . Reformulating (5) as an optimization problem and restricting the set of admissible \mathcal{X} to low-rank tensors potentially mitigates the curse of dimensionality. Compared to tensor completion, additional complications arise because \mathcal{A} is usually ill-conditioned and plain first-order methods can therefore be expected to exhibit unsatisfactory converge rates. In [10], we propose two preconditioned gradient methods on low-rank tensor manifolds: A Riemannian version of the preconditioned Richardson method as well as an approximate Newton scheme based on the Riemannian Hessian. For the latter, considerable attention is given to the efficient solution of the resulting Newton equation. In numerical experiments, we compare the efficiency of our Riemannian algorithms with other established tensor-based approaches such as a truncated preconditioned Richardson method and the alternating linear scheme. The results show that our approximate Riemannian Newton scheme is significantly faster in cases when the application of the linear operator is expensive. For the special case of approximating the discretized d -dimensional Laplace equation in the Tucker format, we obtain a direct extension of an existing low-rank solver for Lyapunov matrix equations [15].

REFERENCES

- [1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, Princeton, NJ, 2008.
- [2] C. Da Silva and F. J. Herrmann. Optimization on the hierarchical Tucker manifold—applications to tensor completion. *Linear Algebra Appl.* 481, 131–173, 2015.
- [3] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, 2000.
- [4] G. Heidel and V. Schulz. A Riemannian trust-region method for low-rank tensor completion arXiv preprint 1703.10019, 2017.
- [5] S. Holtz, T. Rohwedder, and R. Schneider. On manifolds of tensors of fixed TT-rank. *Numer. Math.*, 120(4):701–731, 2012.
- [6] H. Kasai and B. Mishra. Low-rank tensor completion: a Riemannian manifold preconditioning approach. International Conference on Machine Learning, 2016
- [7] O. Koch and Ch. Lubich. Dynamical tensor approximation. *SIAM J. Matrix Anal. Appl.*, 31(5):2360–2375, 2010.
- [8] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.
- [9] D. Kressner, M. Steinlechner, and B. Vandereycken. Low-rank tensor completion by Riemannian optimization. *BIT*, 54(2):447–468, 2014.
- [10] D. Kressner, M. Steinlechner, and B. Vandereycken. Preconditioned low-rank Riemannian optimization for linear systems with tensor product structure. *SIAM J. Sci. Comput.* 38, no. 4, A2018–A2044, 2016.
- [11] R. Schneider, Reinhold and A. Uschmajew. Convergence results for projected line-search methods on varieties of low-rank matrices via Łojasiewicz inequality. *SIAM J. Optim.* 25, no. 1, 622–646, 2015.
- [12] M. Steinlechner. Riemannian optimization for high-dimensional tensor completion. *SIAM J. Sci. Comput.* 38, no. 5, S461–S484, 2016.

- [13] A. Uschmajew. *Zur Theorie der Niedrigrangapproximation in Tensorprodukten von Hilberträumen*. PhD thesis, Technische Universität Berlin, 2013.
- [14] A. Uschmajew and B. Vandereycken. The geometry of algorithms using hierarchical tensors. *Linear Algebra Appl.*, 439(1):133–166, 2013.
- [15] B. Vandereycken and S. Vandewalle. A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations. *SIAM J. Matrix Anal. Appl.*, 31(5):2553–2579, 2010.

A Note on Rank-One Subspace Modifications and some Remarks on the Canonical Stiefel Logarithm

RALF ZIMMERMANN

As with the corresponding talk, this extended abstract is divided into two virtually independent parts. The first part is on a geometric approach to rank-one subspace modifications while the second part reports on a recent development in computing the Riemannian logarithm map on the Stiefel manifold with respect to the canonical metric.

1. PART: RANK-ONE SUBSPACE UPDATES

Introduction. Investigations on the behaviour of matrix decompositions under perturbations of restricted rank have a long tradition [2–5, 7, 9]. Of special importance in many applications are rank-one modifications $X_{new} = X + ab^T$ of a given matrix $X \in \mathbb{R}^{n \times p}$ with either known (thin) singular value decomposition (SVD) or known (compact) QR-decomposition. In the former case, the matrix X is factorised as $X = U\Sigma V^T$, where $U \in \mathbb{R}^{n \times p}$ is column-orthogonal, $\Sigma \in \mathbb{R}^{p \times p}$ is diagonal and $V \in \mathbb{R}^{p \times p}$ is orthogonal. In the latter, the matrix X is factorised as $X = QR$, where $Q \in \mathbb{R}^{n \times p}$ is column-orthogonal and $R \in \mathbb{R}^{p \times p}$ is upper triangular.

Observe that in both cases, the matrix decomposition is of the form $X = UW$, with a column-orthogonal matrix U such that $U^T U$ is the p -by- p identity matrix I_p . Hence, the SVD and the QR-decomposition provide an *orthonormal basis* (ONB) for the range of X , i.e., the subspace $\mathcal{X} = \text{colspan}(X)$. From an abstract perspective, U and Q are to be considered as points on the *Stiefel manifold*

$$St(n, p) = \{U \in \mathbb{R}^{n \times p} \mid U^T U = I_p\}$$

that represent subspaces $[U] = \text{colspan}(X) = [Q]$ on the *Grassmann manifold*

$$Gr(n, p) = \{\mathcal{X} \subset \mathbb{R}^n \mid \mathcal{X} \text{ subspace, } \dim(\mathcal{X}) = p\},$$

where we consider the Grassmannian as a quotient of the Stiefel manifold under actions of the orthogonal group O_p in the following sense:

$$Gr(n, p) = St(n, p)/O_p = \{\{UR \mid R \in O_p\} \mid U \in St(n, p)\} = \{[U] \mid U \in St(n, p)\}.$$

For more details, see [1, 6].

Rank-one subspace updates. Let $X = UW$, where $U \in St(n, p)$ is an ONB for $\mathcal{X} = [U]$. The classical approach [3–5] to the rank-one update problem

$$U_{new}W_{new} = X_{new} = X + ab^T = UW + ab^T$$

is via writing the update in factorised form as

$$UW + ab^T = (U | q)K, \quad (U | q) \in St(n, p + 1), \quad K \in \mathbb{R}^{(p+1) \times p}.$$

Here, the component q of the vector a that is orthogonal w.r.t. U is appended as an extra column and the update problem reduces to computing the SVD or the QR-decomposition of the matrix K . If $K = \tilde{U}\tilde{W}$ is established, then $U_{new} = (U | q)\tilde{U}$ is the updated subspace representative. Neglecting the costs of the decomposition of K , this requires a matrix product of $\mathcal{O}(n(p+1)p) = \mathcal{O}(np^2)$ FLOPS.¹

Our contribution is a method that provides the subspace $[U_{new}]$ together with an ONB at a reduced cost of $\mathcal{O}(np)$ FLOPS and in closed form. This is achieved via a geometric approach [11]. One can show that the modified subspace $\mathcal{X}_{new} = \text{colspan}(X + ab^T)$ can be reached via a geodesic path starting in $[U] = \text{colspan}(X)$ with a suitable tangent velocity vector Δ

$$t \mapsto \text{Exp}_{[U]}^{Gr}(t\Delta),$$

where $\text{Exp}_{[U]}^{Gr}$ denotes the Riemannian exponential on the Grassmann manifold. In this way, the rank-one update problem is reduced to the problem of finding a suitable starting velocity $\Delta \in T_{[U]}Gr(n, p)$. The following theorem provides both Δ and $[U_{new}]$. For details and proofs, see [11].

Theorem 1 ([11]). *Consider the rank-one update problem $X_{new} = X + ab^T = UW + ab^T$. From this data, compute the following quantities:*

$$\begin{aligned} \tilde{q} &= (I - UU^T)a, \quad q = \frac{\tilde{q}}{\|\tilde{q}\|}, \quad \tilde{w} = -W^{-T}b, \quad w = \frac{\tilde{w}}{\|\tilde{w}\|}, \quad \omega = \frac{1}{\|\tilde{q}\|}(1 - a^T U \tilde{w}), \\ g &= (\tilde{w}, \omega)^T, \quad \alpha = \frac{|\omega|}{\|g\|} - 1, \quad \beta = -\text{sign}(\omega) \frac{\|\tilde{w}\|}{\|g\|}. \end{aligned}$$

Then, the unit tangent vector $\Delta := qw^T$ is such that the geodesic that starts at $[U]$ with velocity Δ meets the point $[U_{new}] = \text{colspan}(X_{new})$. In particular,

$$[U_{new}] = \text{Exp}_{[U]}^{Gr}(t^* \Delta) = [U + (\alpha U w + \beta q) w^T] \in Gr(n, p),$$

where $t^ = \arccos\left(\frac{|\omega|}{\|g\|}\right)$ is the Riemannian distance between $[U_{new}]$ and $[U]$.*

2. PART: SOME REMARKS ON THE CANONICAL STIEFEL LOGARITHM

Introduction. Many data processing methods on Riemannian manifolds make use of local *normal coordinates* for mapping data back and forth between the curved manifold and its flat tangent space. This is via the Riemannian logarithm

¹floating point operations

and the Riemannian exponential, which depend on the chosen metric. On the Stiefel manifold, the tangent space at $U \in St(n, p)$ is

$$T_U St(n, p) = \{\Delta \in \mathbb{R}^{n \times p} | \Delta = UA + U^\perp B, \quad a \in \mathbb{R}^{p \times p} \text{skew}, B \in \mathbb{R}^{n-p \times p}\}.$$

The Euclidean metric on $T_U St(n, p)$ is the one inherited from the ambient $\mathbb{R}^{n \times p}$:

$$g^E(\Delta, \Delta) = \text{trace}(\Delta^T \Delta) = \text{trace}(A^T A) + \text{trace}(B^T B).$$

The canonical metric arises from intrinsic considerations, see [6], and reads

$$g^C(\Delta, \Delta) = \text{trace}(\Delta^T (I - 1/2UU^T)\Delta) = \frac{1}{2}\text{trace}(A^T A) + \text{trace}(B^T B).$$

An algorithm for computing the exponential w.r.t. the canonical metric was derived in [6] and proceeds as follows.

- Input: $U \in St(n, p)$, $\Delta \in T_U St(n, p)$.
 - $QR := (I - UU^T)\Delta$, (QR of normal component)
 - $A := U^T \Delta$, (horizontal component)
 - $\begin{pmatrix} M \\ N \end{pmatrix} := \exp_m \left(\begin{pmatrix} A & -R^T \\ R & \mathbf{0} \end{pmatrix} \right) \begin{pmatrix} I_p \\ 0 \end{pmatrix} \in \mathbb{R}^{2p \times p}$
- Output: $\tilde{U} := Exp_U^{St}(\Delta) = UM + QN \in St(n, p)$

The Riemannian logarithm. It turns out that the essential building block for computing the associated Riemannian logarithm is solving the following nonlinear matrix equation [8, 10]:

$$(1) \quad (0 | I_p) \log_m \left(\begin{pmatrix} M & X_0 \\ N & Y_0 \end{pmatrix} \begin{pmatrix} I_p & 0 \\ 0 & \Phi \end{pmatrix} \right) \begin{pmatrix} 0 \\ I_p \end{pmatrix} =: (0, | I_p) \log_m(VW) \begin{pmatrix} 0 \\ I_p \end{pmatrix} \stackrel{!}{=} 0.$$

The matrix blocks M, N are obtained from the input data and X_0, Y_0 constitute a suitable orthogonal completion such that $V = \begin{pmatrix} M & X_0 \\ N & Y_0 \end{pmatrix} \in O_{2p}$. The unknown variable is $\Phi \in O_p$. Note that (1) asks us to find Φ such that the lower p -by- p diagonal block of the matrix $\log_m(VW)$ cancels. In [10], I introduced an algorithm based on the Baker–Campbell–Hausdorff(BCH) series for the matrix logarithm

$$\begin{aligned} \log_m(VW) &= \log_m(V) + \log_m(W) + \frac{1}{2}[\log_m(V), \log_m(W)] \\ &+ \frac{1}{12}([\log_m(V), [\log_m(V), \log_m(W)]] + [\log_m(W), [\log_m(W), \log_m(V)]]) + \dots \end{aligned}$$

Write $\log_m(V) := \begin{pmatrix} A & -B^T \\ B & C \end{pmatrix}$, $\log_m(W) := \begin{pmatrix} 0 & 0 \\ 0 & X \end{pmatrix}$. The algorithm [10, Alg. 1] and the associated convergence analysis rely on the observation that with the choice of $X = -C$ and the corresponding $\Phi = \exp_m(X)$, the lower p -by- p block of $\log_m(VW)$ cancels up to terms of third order in the BCH series.

Towards an improved algorithm. A closer look at the BCH series reveals that when X is chosen to satisfy the Sylvester equation

$$C = \left(\frac{1}{12}BB^T - I_p \right) X + X \left(\frac{1}{12}BB^T \right),$$

then the third-order terms in the BCH series of the lower p -by- p block cancel up to terms that are quadratic in C and X . Preliminary experiments show that this choice of X at every iteration step of [10, Alg. 1] improves the iteration count by a factor of 2 and accelerates the algorithm by a factor of 1.5. For

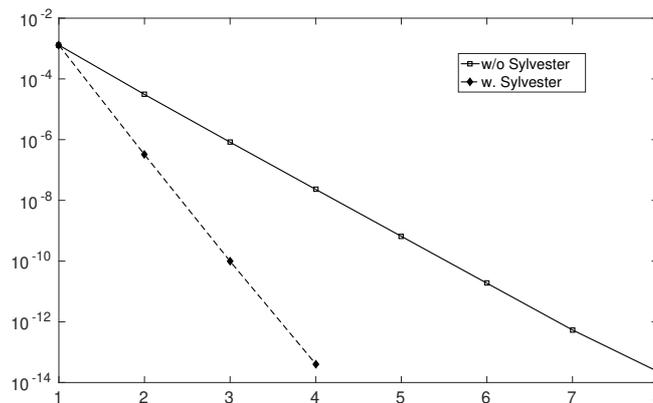


FIGURE 1. Performance of the canonical Stiefel logarithm with and without solving a Sylvester equation to obtain the next iterate for input data on $St(n, p)$, $n = 100,000$, $p = 500$.

random data $U_0, U_1 \in St(n, p)$, $n = 100,000$, $p = 500$, with Euclidean distance $\|U_0 - U_1\|_2 = 0.4734$, the original algorithm [10, Alg. 1] takes 8 iterations and ca. 16.3s while the Sylvester-enhanced version converges to the same accuracy in 4 iterations and 10.5s, see Figure 1.

This observation will be the starting point for further investigations on the convergence domain and the convergence rate of the Sylvester-enhanced Stiefel log algorithm.

Remark: A brute-force solution of the non-linear matrix equation (1), say, via MATLAB's pre-installed 'fsolve'-function is non-competitive. For the case at hand with $n = 100,000$, $p = 500$ it is even unfeasible. For a case with $n = 10,000$, $p = 100$, it is 2000 times slower. (less than 0.3s vs. 576s \approx 10min)

REFERENCES

- [1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, New Jersey, 2008.
- [2] L. Balzano and S.J. Wright, *On GROUSE and Incremental SVD*, 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 1–4, 2013.
- [3] M. Brand. Fast low-rank modifications of the thin singular value decomposition. *Linear Algebra and its Applications*, 415:20–30, 2006.

- [4] J. R. Bunch and C. P. Nielsen. Updating the singular value decomposition. *Numerische Mathematik*, 31:111–129, 1978.
- [5] J. W. Daniel, W. B. Gragg, L. Kaufman, and G. W. Stewart. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. *Mathematics of Computation*, 30:772–795, 1976.
- [6] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM Journal on Matrix Analysis and Application*, 20(2):303–353, 1999.
- [7] M. Gu and S. C. Eisenstat. A stable and efficient algorithm for the rank-one modification of the symmetric eigenproblem. *SIAM Journal on Matrix Analysis and Applications*, 15:1266–1276, 1994.
- [8] Q. Rentmeesters, *Algorithms for data fitting on some common homogeneous spaces*. PhD thesis, Université Catholique de Louvain, Louvain, Belgium, 2013.
- [9] R. C. Thompson. The behavior of eigenvalues and singular values under perturbations of restricted rank. *Linear Algebra and its Applications*, 13:69–78, 1976.
- [10] R. Zimmermann, *A matrix-algebraic algorithm for the Riemannian logarithm on the Stiefel manifold under the canonical metric*. *SIAM Journal on Matrix Analysis and Applications*, 38(2):322–342, 2017.
- [11] R. Zimmermann, *A closed-form update for orthogonal matrix decompositions under arbitrary rank-one modifications*. arXiv:1711.08235v2, 2018

Optimal Sensor Selection Using Reduced Models

OLGA MULA

(joint work with P. Binev, A. Cohen, J. Nichols)

1. SETTING

We consider the recovery problem introduced in [1,2] whose setting is the following (see also [3,4] for very closely related formulations). Let V be a Hilbert space over a domain $\Omega \subset \mathbb{R}^d$ with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. We want to approximate an unknown function $u \in V$ for which we observe m measurements

$$z_i := \ell_i(u), \quad i = 1, \dots, m,$$

where the ℓ_i are independent continuous linear functionals over V . The knowledge of $z = (z_i)_{i=1}^m$ is equivalent to that of the orthogonal projection

$$w = P_{W_m} u,$$

where $W_m := \text{span}\{\omega_1, \dots, \omega_m\}$ and ω_i are the Riesz representers of the linear functionals ℓ_i . Since there are infinitely many $v \in V$ such that $P_{W_m} v = w$, the only way to recover u up to a guaranteed accuracy is to combine the measurements with some a-priori information on u . Our additional assumption is that u belongs to a manifold \mathcal{M} of functions from V . This hypothesis is relatively natural in the sense that, if we interpret our setting from a physical point of view, our goal is to recover from m measures a function $u \in V$ which is the state of a physical system. In general, the system may present different working conditions so the manifold \mathcal{M} is the set of all possible states. Also, note that the ℓ_i (or their Riesz representers ω_i) can be seen as the mathematical model for the sensor devices that are placed in the experiment.

Since in general \mathcal{M} does not have a simple geometry (lack of convexity, for instance), it is not simple to recover u by working on this set. For certain relevant types of manifolds like the ones given by the set of solutions of elliptic PDEs, it is possible to build linear spaces $V_n \subset V$ of dimension n such that

$$\sup_{u \in \mathcal{M}} \|u - P_{V_n} u\| \leq \varepsilon(n)$$

where $\varepsilon(n)$ decays at a comparable rate to the optimal one given by the Kolmogorov n -width [5]. Therefore, for a given target accuracy ε , one can work with subspaces V_n of moderate dimension when the width presents a fast decay and this element is an important ingredient to build recovery schemes that are both accurate, easy and quickly computable. These subspaces are usually called reduced models.

Given a reduced model V_n and the observation w from the measurements, the optimal recovery solution $u^*(w)$ is

$$u^*(w) := \arg \min \{ \|v - P_{V_n} v\| : P_{W_m} v = w \}.$$

It can be computed from the data w by solving a finite set of linear equations. The worst case performance for this reconstruction is given by

$$(1) \quad \sup_{u \in \mathcal{M}} \|u - u^*(P_W u)\| = \frac{1}{\beta(V_n, W_m)} \varepsilon_n,$$

where

$$\beta(V_n, W_m) := \inf_{v \in V_n} \frac{\|P_{W_m} v\|}{\|v\|} \in [0, 1]$$

plays the role of a stability constant. Note that $\beta(V_n, W_m) > 0$ requires that $m \geq n$.

In the rest of this short note, we recall some recent results on optimal sensor selection that use the above methodology (see [6] for the details) and mention some connections of the present setting with different fields.

2. GREEDY ALGORITHMS FOR OPTIMAL SENSOR PLACEMENT (SEE [6])

For a given reduced model space V_n with accuracy ε_n , it follows from the performance formula (1) that one natural objective is to ensure $\beta(V_n, W_m) \geq \underline{\beta} > 0$, with a number of measurements $m \geq n$ as small possible. Note that taking $W_m = V_n$ would automatically give the maximal value $\beta(V_n, W_m) = 1$ with $m = n$. However, in a typical data acquisition scenario, the sensors ω_i that span the basis of W_m are chosen from within a limited class. This is the case for example when placing m pointwise sensors at various locations within the physical domain Ω .

We model this restriction by asking that the ℓ_i are picked within a *dictionary* \mathcal{D} of V' , that is a set of linear functionals normalized according to

$$\|\ell\|_{V'} = 1, \quad \ell \in \mathcal{D},$$

which is *complete* in the sense that $\ell(v) = 0$ for all $\ell \in \mathcal{D}$ implies that $v = 0$. With an abuse of notation, we identify \mathcal{D} with the subset of V that consists of all Riesz representers ω of the above linear functionals ℓ .

Our task is therefore to pick $\{\omega_1, \dots, \omega_m\} \in \mathcal{D}$ in such a way that

$$(2) \quad \beta(V_n, W_m) \geq \underline{\beta} > 0,$$

for some prescribed $0 < \underline{\beta} < 1$, with m larger than n but as small as possible. In particular, we may introduce

$$m^* = m^*(\underline{\beta}, \mathcal{D}, V_n),$$

the minimal value of m such that there exists $\{\omega_1, \dots, \omega_m\} \in \mathcal{D}$ satisfying (2).

Since finding the optimal measurements from \mathcal{D} is in general an np-hard combinatorial problem, we explore in [6] two greedy algorithms that are more amenable for computation. We recall briefly the main results for one of them in the following. One important point is to compare the quality of the selected measurements with respect to the optimal choice. In [6], this is done for simple cases where the optimal selection is a priori known.

A collective Orthogonal Matching Pursuit algorithm:

Setting $W_0 := \{0\}$, we iteratively select for $k \geq 1$

$$(3) \quad \omega_k = \arg \max_{\omega \in \mathcal{D}} \max_{v \in V_n, \|v\|=1} |\langle \omega, v - P_{W_{k-1}} v \rangle| = \arg \max_{\omega \in \mathcal{D}} \|P_{V_n}(\omega - P_{W_{k-1}} \omega)\|,$$

where $W_{k-1} := \text{span}\{\omega_1, \dots, \omega_{k-1}\}$. Note that in the case $n = 1$, we obtain the original OMP algorithm applied to a single function $v \in V$.

As to the implementation of this algorithm, we take (ϕ_1, \dots, ϕ_n) to be any orthonormal basis of V_n . Then

$$\|P_{V_n}(\omega - P_{W_{k-1}} \omega)\|^2 = \sum_{i=1}^n |\langle \omega - P_{W_{k-1}} \omega, \phi_i \rangle|^2 = \sum_{i=1}^n |\langle \phi_i - P_{W_{k-1}} \phi_i, \omega \rangle|^2$$

Therefore, at every step k , we have

$$\omega_k = \arg \max_{\omega \in \mathcal{D}} \sum_{i=1}^n |\langle \phi_i - P_{W_{k-1}} \phi_i, \omega \rangle|^2,$$

which amounts to a stepwise optimization of a similar nature as in the standard OMP. Note that, while the basis (ϕ_1, \dots, ϕ_n) is used for the implementation, the actual definition of the greedy selection algorithm is independent of the choice of this basis in view of (3). It only involves V_n and the dictionary \mathcal{D} . Similar to OMP, we may weaken the algorithm by taking ω_k such that

$$\sum_{i=1}^n |\langle \phi_i - P_{W_{k-1}} \phi_i, \omega_k \rangle|^2 \geq \kappa^2 \max_{\omega \in \mathcal{D}} \sum_{i=1}^n |\langle \phi_i - P_{W_{k-1}} \phi_i, \omega \rangle|^2,$$

for some fixed $0 < \kappa < 1$. For such a basis, we introduce the residual quantity

$$r_m := \sum_{i=1}^n \|\phi_i - P_{W_m} \phi_i\|^2,$$

which is, like before, independent of the choice of the basis. This quantity allows us to control the validity of (2) since we have

$$1 - \beta^2(V_n, W_m) = \sup_{v \in V_n, \|v\|=1} \|v - P_{W_m} v\|^2 = \sup_{\sum_{i=1}^n c_i^2 = 1} \left\| \sum_{i=1}^n c_i (\phi_i - P_{W_m} \phi_i) \right\|^2 \leq r_m,$$

and therefore (2) holds provided that $r_m \leq 1 - \underline{\beta}^2$. In [6], the following convergence result is proven for $(r_n)_n$.

Theorem: Let $\Phi = (\phi_1, \dots, \phi_n)$ be an orthonormal basis of V_n and $\Psi = (\psi_1, \dots, \psi_n) \in V^n$ be arbitrary. Then the application of the collective OMP algorithm on the space V_n gives

$$r_m \leq 4 \frac{\|\Psi\|_{\ell^1(\mathcal{D})}^2}{\kappa^2} (m+1)^{-1} + \|\Phi - \Psi\|^2, \quad m \geq 1.$$

where $\|\Phi - \Psi\|^2 := \|\Phi - \Psi\|_{V^n}^2 = \sum_{i=1}^n \|\phi_i - \psi_i\|^2$.

3. CONNECTIONS OF THE PRESENT SETTING WITH DIFFERENT FIELDS

The recovery of a function $u \in V$ from a measurement w belonging to a sampling space W_m and using a well-chosen approximation space V_n arises in many different applications. A non exhaustive list comprises the sampling of bandlimited functions, Fourier sampling or data assimilation combined with reduced models. The table below lists relevant choices of spaces V_n and W_m .

Some choices for V_n	Some choices for W_m
<ul style="list-style-type: none"> • Polynomial subspaces • Trigonometric polynomials • Wavelets • Reduced models built from a parametrized PDE 	<ul style="list-style-type: none"> • Fourier samples • Convolutions • Local averages • Pointwise evaluations

We next give a connection of our setting with subspace distances on Grassmanian manifolds that was discussed during the workshop. Denoting $(\varphi_j)_{j=1}^n$ and $(\omega_i)_{i=1}^m$ two orthonormal bases of V_n and W_m respectively, we define the $m \times n$ cross-Gramian matrix

$$\mathbf{G} = (\langle \omega_i, \varphi_j \rangle)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}}$$

Denoting $0 \leq \sigma_n \leq \sigma_{n-1} \leq \dots \leq \sigma_1 \leq 1$ its singular values, we have that $\beta(V_n, W_m) = \sigma_n$ and $\theta_n = \cos^{-1} \sigma_n$ is the so-called Asimov principal angle in the literature about subspace distances on Grassmanian manifolds.

Finally, the topic of optimal sensor placement in the particular setting where the linear functionals are point evaluations or local averages has been extensively studied since the 1970's in control and systems theory. In this context, the state function to be estimated is the realization of a Gaussian stochastic process and the error is measured in the mean square sense, rather than in the worst case performance sense (1) which is the point of view adopted in our work. The function to be minimized by the sensors locations is then the trace of the error covariance, while we target at maximizing $\beta(V_n, W_m)$.

REFERENCES

- [1] Y. Maday, A.T. Patera, J.D. Penn, M. Yano, *A Parameterized-Background Data-Weak approach to variational data assimilation: formulation, analysis, and application to acoustics*, International Journal for Numerical Methods in Engineering **102** (2015), pp. 933–965.
- [2] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, P. Wojtaszczyk, *Data assimilation in reduced modeling*, SIAM/ASA Journal on Uncertainty Quantification, **5** (2017), 1–29.
- [3] Y. Maday, O. Mula *A Generalized Empirical Interpolation Method: application of reduced basis techniques to data assimilation*, Analysis and Numerics of Partial Differential Equations, **4** (2013), 221–235.
- [4] B. Adcock, A. Hansen, C. Poon, *Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem*, SIAM Journal on Mathematical Analysis, **45** (2013), 3132–3167.
- [5] A. Cohen and R. DeVore, *Kolmogorov widths under holomorphic mappings*, IMA Journal of Numerical Analysis **36** (2015), 1–12.
- [6] P. Binev, A. Cohen, O. Mula, J. Nichols *Greedy algorithms for optimal measurements selection in state estimation using reduced models*, submitted (2017), <https://hal.archives-ouvertes.fr/hal-01638177>

Some Uniqueness Results for Minimisers of Ginzburg–Landau Functionals

RADU IGNAT

(joint work with Luc Nguyen, Valeriy Slastikov, Arghir Zarnescu)

Model. This report is based on the article [3] where we consider the following Ginzburg-Landau type energy functional

$$E_\varepsilon(u) = \int_\Omega \left[\frac{1}{2} |\nabla u|^2 + \frac{1}{2\varepsilon^2} W(1 - |u|^2) \right] dx,$$

with $\varepsilon > 0$ being a fixed parameter, $\Omega \subset \mathbb{R}^m$ ($m \geq 1$) is a bounded domain (i.e., open connected set) with smooth boundary $\partial\Omega$ and the potential $W \in C^1((-\infty, 1]; \mathbb{R}_+)$ satisfies

$$W(0) = 0, W(t) > 0 \text{ for all } t \in (-\infty, 1] \setminus \{0\}, W \text{ is strictly convex.}$$

(The prototype of the nonlinear potential is $W(t) = t^2/2$.) We focus on minimisers of the energy E_ε over the following set

$$\mathcal{A} := \{u \in H^1(\Omega; \mathbb{R}^n) : u = u_{bd} \text{ on } \partial\Omega\}, \quad n \geq 1,$$

consisting of H^1 maps with a given boundary data (in the sense of $H^{1/2}$ -trace on $\partial\Omega$):

$$u_{bd} \in H^{1/2} \cap L^\infty(\partial\Omega; \mathbb{R}^n).$$

The direct method in the calculus of variations yields existence of minimizers u_ε of E_ε over \mathcal{A} for all range of $\varepsilon > 0$; moreover, any minimizer u_ε belongs to $C^1 \cap L^\infty(\Omega; \mathbb{R}^n)$ and satisfies the system of PDEs

$$(1) \quad -\Delta u_\varepsilon = \frac{1}{\varepsilon^2} u_\varepsilon W'(1 - |u_\varepsilon|^2) \quad \text{distributionally in } \Omega.$$

Aim. We are interested in the question of uniqueness (or its failure) for the minimisers of E_ε in \mathcal{A} for all range of $\varepsilon > 0$. If ε is large (i.e., $\varepsilon \geq \varepsilon_0 := (|W'(1)|/\lambda_1(\Omega))^{1/2}$ where $\lambda_1(\Omega)$ is the first eigenvalue of $(-\Delta)$ on Ω with zero Dirichlet data), then E_ε is strictly convex and thus, there exists a unique solution $u_\varepsilon \in \mathcal{A}$ of (1) which is the minimizer of E_ε over \mathcal{A} . If $\varepsilon < \varepsilon_0$, the problem is more delicate. We provide results for this problem in the special case where the boundary data is non-negative in a (fixed) direction $e \in \mathbb{S}^{n-1}$, i.e.,

$$(2) \quad u_{bd} \cdot e \geq 0 \quad \mathcal{H}^{m-1}\text{-a.e. in } \partial\Omega.$$

Example 1. In the scalar case $n = 1$ with zero boundary data $u_{bd} = 0$ on $\partial\Omega$, if $\varepsilon \geq \varepsilon_0$, then $\tilde{u}_\varepsilon = 0$ is the unique solution of (1) in \mathcal{A} (so, the unique minimizer of E_ε over \mathcal{A}). If $\varepsilon < \varepsilon_0$, then there exists a unique positive solution $u_\varepsilon \in \mathcal{A}$ (i.e., $u_\varepsilon > 0$ in Ω) of (1) with zero boundary data, see e.g. [1]; as a consequence of Theorems 1 and 3 (see below), we have that u_ε and $-u_\varepsilon$ are the only two minimizers of E_ε over \mathcal{A} and moreover, the trivial solution $\tilde{u}_\varepsilon = 0$ is unstable (i.e., the second variation of E_ε at \tilde{u}_ε is negative in a certain direction).

Example 2. For $m = 2$ and $n = 3$, we consider the unit disk $\Omega \subset \mathbb{R}^2$ and the boundary data carrying a given winding number $k \in \mathbb{Z} \setminus \{0\}$ on $\partial\Omega$:

$$u_{bd}(\cos \varphi, \sin \varphi) = (\cos(k\varphi), \sin(k\varphi), 0) \in \mathbb{S}^1 \times \{0\} \subset \mathbb{R}^3, \quad \forall \varphi \in [0, 2\pi).$$

(Note that u_{bd} satisfies (2) in the vertical direction e_3 .) As a consequence of Theorem 1 (see below), there exists $\varepsilon_k > 0$ such that

a) if $\varepsilon \geq \varepsilon_k$, the unique minimizer of E_ε over \mathcal{A} is given by

$$\tilde{u}_\varepsilon := \tilde{f}_\varepsilon(r)(\cos(k\varphi), \sin(k\varphi), 0), \quad r \in (0, 1), \varphi \in [0, 2\pi),$$

where the radial profile \tilde{f}_ε is the unique solution of the ODE (see e.g. [2])

$$\begin{cases} -\tilde{f}_\varepsilon'' - \frac{1}{r}\tilde{f}_\varepsilon' + \frac{k^2}{r^2}\tilde{f}_\varepsilon = \frac{1}{\varepsilon^2}\tilde{f}_\varepsilon W'(1 - \tilde{f}_\varepsilon^2) & \text{in } (0, 1), \\ \tilde{f}_\varepsilon(0) = 0, \tilde{f}_\varepsilon(1) = 1; \end{cases}$$

b) if $\varepsilon < \varepsilon_k$, then E_ε admits exactly two minimizers u_ε^\pm over \mathcal{A} that have the form

$$u_\varepsilon^\pm := f_\varepsilon(r)(\cos(k\varphi), \sin(k\varphi), 0) \pm g_\varepsilon(r)(0, 0, 1), \quad g_\varepsilon(r) > 0, \quad r \in (0, 1), \varphi \in [0, 2\pi),$$

where the couple $(f_\varepsilon, g_\varepsilon)$ of radial profiles is the unique solution of the system

$$\begin{cases} -f_\varepsilon'' - \frac{1}{r}f_\varepsilon' + \frac{k^2}{r^2}f_\varepsilon = \frac{1}{\varepsilon^2}f_\varepsilon W'(1 - f_\varepsilon^2 - g_\varepsilon^2) & \text{in } (0, 1), \\ -g_\varepsilon'' - \frac{1}{r}g_\varepsilon' = \frac{1}{\varepsilon^2}g_\varepsilon W'(1 - f_\varepsilon^2 - g_\varepsilon^2) & \text{in } (0, 1), \\ f_\varepsilon \geq 0, g_\varepsilon > 0 & \text{in } (0, 1), \\ f_\varepsilon(0) = 0, f_\varepsilon(1) = 1, g_\varepsilon'(0) = 0, g_\varepsilon(1) = 0. \end{cases}$$

Moreover, the solution \tilde{u}_ε of (1) (given at point a) above) is unstable if $\varepsilon < \varepsilon_k$.

These examples suggest the following phenomenology: if $V = \text{Span } u_{bd}(\partial\Omega)$ has co-dimension ≥ 1 in \mathbb{R}^n , then non-uniqueness of minimizers of E_ε over \mathcal{A} is equivalent with the existence of “escaping” solutions $u_\varepsilon \in \mathcal{A}$ of (1) (i.e., $u_\varepsilon(\Omega) \not\subset V$). This is highlighted by the following result:

Theorem 1 ([3]). *Let $u_\varepsilon \in H^1 \cap L^\infty(\Omega; \mathbb{R}^n)$ be an “escaping” critical point of the energy E_ε over \mathcal{A} such that $u_\varepsilon \cdot e > 0$ a.e. in Ω in some direction $e \in \mathbb{S}^{n-1}$ for some $\varepsilon > 0$. Then u_ε is a minimiser of E_ε over \mathcal{A} and we have the following dichotomy:*

- a) *If $u_{bd}(x_0) \cdot e > 0$ for some Lebesgue point $x_0 \in \partial\Omega$, then u_ε is the unique minimiser of E_ε over \mathcal{A} .*
- b) *If $u_{bd}(x) \cdot e = 0$ for \mathcal{H}^{m-1} -a.e. $x \in \partial\Omega$, then all minimisers of E_ε in \mathcal{A} are given by Ru_ε where $R \in O(n)$ is an orthogonal transformation of \mathbb{R}^n satisfying $Rx = x$ for all $x \in \text{Span } u_{bd}(\partial\Omega)$.*

Using the above theorem, we prove the following result which completely characterises uniqueness and its failure for minimisers of the energy E_ε over \mathcal{A} under the assumption (2) for the boundary data u_{bd} .

Theorem 2 ([3]). *Let $\varepsilon > 0$. If (2) holds in direction $e \in \mathbb{S}^{n-1}$ and $V = \text{Span } u_{bd}(\partial\Omega)$, then there exists a unique minimiser u_ε of the energy E_ε over \mathcal{A} unless both following conditions hold:*

- i) $u_{bd}(x) \cdot e = 0$ \mathcal{H}^{m-1} -a.e. $x \in \partial\Omega$,
- ii) the functional E_ε restricted to the set

$$\mathcal{A}_{res} := \{u \in \mathcal{A} : u(x) \in \text{Span}(V \cup \{e\}) \text{ a.e. in } \Omega\}$$

has an “escaping” minimiser \tilde{u}_ε with $\tilde{u}_\varepsilon(\Omega) \not\subset V$.

Moreover, if uniqueness of minimisers of E_ε in \mathcal{A} does not hold, then all minimisers of E_ε in \mathcal{A} are given by $R\tilde{u}_\varepsilon$ where $R \in O(n)$ is an orthogonal transformation of \mathbb{R}^n satisfying $Rx = x$ for all $x \in V$.

The “escaping” phenomenon is closely related to stability properties of critical points if $\text{codim}_{\mathbb{R}^n}(V) \geq 1$ with $V = \text{Span } u_{bd}(\partial\Omega)$. Indeed, by Theorem 1, every “escaping” critical point u_ε of E_ε over \mathcal{A} is in fact a minimiser and there are multiple minimisers as one can reflect u_ε about the orthogonal space to the escaping direction (so, non-uniqueness holds in this case). On the contrary, we show in the following that for a “non-escaping” critical point u_ε of E_ε over \mathcal{A} (i.e., $u_\varepsilon(\Omega) \subset V$), its stability is equivalent with its minimality and therefore, by Theorem 2, u_ε is the unique minimiser.

Theorem 3 ([3]). *Assume that $V = \text{Span } u_{bd}(\partial\Omega) \subset e^\perp = \{v \in \mathbb{R}^n : v \cdot e = 0\}$ for a direction $e \in \mathbb{S}^{n-1}$. For any fixed $\varepsilon > 0$, if u_ε is a bounded critical point of E_ε in \mathcal{A} confined in e^\perp , i.e., $u_\varepsilon \in L^\infty(\Omega; e^\perp)$ and u_ε is stable in direction e , i.e.,*

$$\left. \frac{d^2}{dt^2} \right|_{t=0} E_\varepsilon(u_\varepsilon + t\varphi e) = \int_\Omega \left[|\nabla\varphi|^2 - \frac{1}{\varepsilon^2} W'(1 - |u_\varepsilon|^2) \varphi^2 \right] dx \geq 0 \text{ for all } \varphi \in H_0^1(\Omega),$$

then u_ε is a minimiser of E_ε in \mathcal{A} . Moreover, if u_ε is “non-escaping”, i.e., $u_\varepsilon(\Omega) \subset V$, then u_ε is the unique minimiser of E_ε in \mathcal{A} .

Our results hold true also for the harmonic map problem, thus covering the well-known result of Sandier and Shafrir [4] on the uniqueness of minimising harmonic maps into a closed hemisphere. In fact, our argument does not assume the

smoothness of boundary data and does not use the regularity theory of minimising harmonic maps, which appears to play a role in the argument of [4].

REFERENCES

- [1] H. Brezis, L. Oswald, *Remarks on sublinear elliptic equations*, *Nonlinear Anal.* **10** (1986), 55-64.
- [2] R. Ignat, L. Nguyen, V. Slustikov, A. Zarnescu, *Uniqueness results for an ODE related to a generalized Ginzburg–Landau model for liquid crystals*, *SIAM J. Math. Anal.* **46** (2014), 3390-3425.
- [3] R. Ignat, L. Nguyen, V. Slustikov, A. Zarnescu, *On the uniqueness of minimisers of Ginzburg–Landau functionals*, preprint arXiv:1708.05040.
- [4] E. Sandier, I. Shafrir, *On the uniqueness of minimizing harmonic maps to a closed hemisphere*, *Calc. Var. Partial Differential Equations* **2** (1994), 113-122.

Variational Convergence of Discrete Minimal Surfaces

MAX WARDETZKY

(joint work with Henrik Schumacher)

Consider a finite set $\Gamma = \{\Gamma_1, \Gamma_2, \Gamma_3, \dots\}$ of closed embedded curves in \mathbb{R}^m . Among all surfaces of *prescribed* topology spanning Γ find those with least (or more precisely critical) area. Solutions of this problem are known as *minimal surfaces*—an extensively studied problem. In the 1930s, Radó and Douglas independently solved the least area problem for the special case of disk-like, immersed surfaces. To date, though, less is known about the existence of minimisers for the general case.

A natural question is how to compute minimal surfaces using finite dimensional approximations. Already Douglas followed this approach using finite differences. A more flexible option is to consider a given finite set Γ of embedded boundary curves in \mathbb{R}^3 , followed by spanning a triangle mesh into Γ and moving the positions of interior vertices such that the overall area of the triangle mesh is minimised. Following this approach, several authors have applied Newton-like methods for finding critical points of the area functional and gradient descent (the discrete mean curvature flow) in order to produce discrete minimisers. With these tools at hand, the question remains whether the so obtained discrete minimisers converge to smooth minimal surfaces and if so in which sense?

One approach for which such convergence of discrete minimisers could be established is based on adapting Douglas' existence proof for disk-like minimal surfaces: Instead of the area of (unparameterised) surfaces, the Dirichlet energy of conformal surface parameterizations is minimised under the constraint of the so-called *three point condition*. Several authors have utilised this idea in order to compute numerical approximations of minimal surfaces via finite element analysis. However, these energy methods (which are based on minimising the Dirichlet energy instead of the area functional) face certain difficulties. E.g., in dimension greater than two, Dirichlet energy is no longer conformally invariant and minimisers of the Dirichlet energy need not minimise area.

In contrast, the Ritz method, i.e., the approach of minimising area (or volume) among simplicial manifolds, is in principle capable of treating any dimension, co-dimension, and topological class with a single algorithm. This is the approach we follow here. Even non-manifold examples can be treated with this method. These advantages come at a cost, though, since showing convergence of the Ritz method is hampered by several difficulties, including that (i) simplicial manifolds capture smooth boundary conditions only in an approximate sense; hence, they cannot be utilised to minimise area (or volume) in the space of surfaces with smooth prescribed boundaries, (ii) smooth minimal surfaces are known to satisfy strong regularity properties (e.g., they are analytic for sufficiently nice boundary data), leading to the question in which space and topology smooth and simplicial area minimisers ought to be compared, (iii) the general least area problem is far from being convex, and (iv) area minimisers need neither be unique nor isolated; rather, they are *sets* in general. These obstacles render the use of convex optimisation approaches and monotone operators inappropriate (if not impossible) for showing convergence of discrete (i.e., simplicial) area minimisers.

For these reasons, we suggest a different route for exploring convergence of discrete minimisers, which in particular is capable of dealing with convergence of sets. Building on variational analysis, our main result is to prove *Kuratowski convergence* of discrete area (and volume) minimisers to their smooth counterparts. While Kuratowski convergence is weaker than Hausdorff convergence in general, both notions coincide in compact metric spaces. Kuratowski convergence is related to the perhaps more familiar notion of Γ -convergence: A sequence of functionals Γ -converges if and only if their epigraphs converge in the sense of Kuratowski.

We establish Kuratowski convergence by adopting the notions of *consistency* and *stability*, following the often repeated mantra from numerical analysis that consistency and stability imply convergence. In our setting, consistency refers to the existence of sampling and reconstruction operators that take smooth manifolds to simplicial ones and vice-versa, respectively, such that the discrete and smooth area functionals stay close to one another. Stability refers to a notion of growth of sublevel sets of the smooth area functional near its (set of) minimisers. Additionally, we require the notion of *proximity*, which is motivated by finding a space in which discrete and smooth minimisers can be compared. Showing consistency, stability, and proximity is somewhat technically involved and constitutes our main technical contribution.

As a consequence of Kuratowski convergence we obtain that every cluster point of discrete area minimisers is a smooth minimal surface and every smooth minimal surface that globally minimises area is the limit of a sequence of discrete (almost) minimisers of area. For details, please refer to [1].

REFERENCES

- [1] H. Schumacher and M. Wardetzky, *Variational Convergence of Discrete Minimal Surfaces*, [arXiv:1605.05285](https://arxiv.org/abs/1605.05285).

Participants

Robert Bauer

Institut für Numerische Mathematik
Technische Universität Dresden
Zellescher Weg 12-15
01069 Dresden
GERMANY

Prof. Dr. Thomas Blesgen

Fachbereich 2 (Rm 1-232)
Technische Hochschule Bingen
Berlinstraße 109
55411 Bingen am Rhein
GERMANY

Dr. Christian Böhmer

Department of Mathematics
University College London
Gower Street
London WC1E 6BT
UNITED KINGDOM

Amit Boyarski

Computer Science Department
TECHNION
Israel Institute of Technology
Haifa 32000
ISRAEL

Prof. Dr. Alexander Bronstein

Computer Science Department
TECHNION
Israel Institute of Technology
Haifa 32000
ISRAEL

Edward Chien

MIT CSAIL
The Stata Center
32 Vassar Street
Cambridge MA 02139
UNITED STATES

Dr. Julie Digne

LIRIS - Geomod
Bâtiment Nautibus
23-25, Avenue Pierre de Coubertin
69622 Villeurbanne Cedex
FRANCE

Prof. Dr. Nira Dyn

School of Mathematical Sciences
Tel Aviv University
Ramat Aviv
Tel Aviv 69978
ISRAEL

Dr. Martin Ehler

Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Dennis Elbrächter

Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Evan Gawlik

Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112
UNITED STATES

Prof. Dr. Philipp Grohs

Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Dr. Hanne Hardering

Institut für Numerische Mathematik
Fachrichtung Mathematik
Technische Universität Dresden
Willersbau C311
01062 Dresden
GERMANY

Prof. Dr. Philipp Harms

Institut für Mathematische Stochastik
Universität Freiburg
Ernst-Zermelo-Strasse 1
79104 Freiburg i. Br.
GERMANY

Svenja Hüning

Institut für Geometrie
Technische Universität Graz
Kopernikusgasse 24
8010 Graz
AUSTRIA

Prof. Dr. Knut Hüper

Institut für Mathematik
Universität Würzburg
Am Hubland Nord
97074 Würzburg
GERMANY

Prof. Dr. Radu Ignat

Institut de Mathématiques de Toulouse
Université Paul Sabatier
118, route de Narbonne
31062 Toulouse Cedex 9
FRANCE

Prof. Dr. Arieh Iserles

Department of Applied Mathematics and
Theoretical Physics
University of Cambridge
Wilberforce Road
Cambridge CB3 0WA
UNITED KINGDOM

Timo Klock

Simula Research Laboratory
P.O. Box 134
1325 Lysaker
NORWAY

Prof. Dr. Stavros Komineas

Department of Mathematics and
Applied Mathematics
University of Crete
Voutes Campus
70013 Heraklion, Crete
GREECE

Sarah Koppensteiner

Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Prof. Dr. Daniel Kressner

Institute of Mathematics
Station 8
École Polytechnique Fédérale de
Lausanne
1015 Lausanne
SWITZERLAND

Prof. Dr. Melvin Leok

Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112
UNITED STATES

Prof. Dr. Lek-Heng Lim

Department of Statistics
The University of Chicago
5747 South Ellis Ave.
Chicago, IL 60637-1514
UNITED STATES

Dr. Caroline Moosmüller
John Hopkins University
Whiting School of Engineering
3400 North Charles Street
Baltimore, MD 21218-2608
UNITED STATES

Prof. Dr. Olga Mula
CEREMADE
Université Paris Dauphine
Place du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16
FRANCE

Dr. Ingo Münch
Institut für Baustatik
Karlsruher Institut für Technologie
(KIT)
Kaiserstrasse 12
76131 Karlsruhe
GERMANY

Dr. Valeriya Naumova
Simula Research Laboratory
P.O. Box 134
1325 Lysaker
NORWAY

Lisa Julia Nebel
Institut für Numerische Mathematik
Technische Universität Dresden
Zellescher Weg 12-14
01069 Dresden
GERMANY

Dr. Xavier Pennec
INRIA Sophia Antipolis
2004 Route des Lucioles
06902 Sophia-Antipolis Cedex
FRANCE

Prof. Dr. Helmut Pottmann
Institut für Diskrete Mathematik
und Geometrie
Technische Universität Wien
Wiedner Hauptstraße 8-10
1040 Wien
AUSTRIA

Dr. Tamás Pusztai
Wigner Research Centre for Physics
P.O. Box 49
1525 Budapest
HUNGARY

Prof. Dr. Ulrich Reif
Fachbereich Mathematik
Technische Universität Darmstadt
Schloßgartenstrasse 7
64289 Darmstadt
GERMANY

Dr. David Rottensteiner
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Prof. Dr. Oliver Sander
Institut für Numerische Mathematik
Technische Universität Dresden
Zellescher Weg 12-14
01069 Dresden
GERMANY

Prof. Dr. Rodolphe Sepulchre
Engineering Department
University of Cambridge
Trumpington Street
Cambridge CB2 1PZ
UNITED KINGDOM

Prof. Dr. Gabriele Steidl

Fachbereich Mathematik
Technische Universität Kaiserslautern
67653 Kaiserslautern
GERMANY

Dr. Martin Storath

Image Analysis and Learning Group
Universität Heidelberg
Berliner Strasse 43
69120 Heidelberg
GERMANY

Dr. Irène Waldspurger

CEREMADE
Université Paris Dauphine
Place du Marechal de Lattre deTassigny
75775 Paris Cedex 16
FRANCE

Prof. Dr. Johannes Wallner

Institut für Geometrie
Technische Universität Graz
Kopernikusgasse 24
8010 Graz
AUSTRIA

Prof. Dr. Max Wardetzky

Institut für Numerische und
Angewandte Mathematik
Universität Göttingen
Lotzestrasse 16-18
37083 Göttingen
GERMANY

Dr. Andreas Weinmann

FB Mathematik und
Naturwissenschaften
Fachhochschule Darmstadt
Schöfferstrasse 3
64295 Darmstadt
GERMANY

Prof. Dr. Benedikt Wirth

Fachbereich Mathematik und Informatik
Universität Münster
Einsteinstrasse 62
48149 Münster
GERMANY

Prof. Dr. Laurent Younes

John Hopkins University
Whiting School of Engineering
3400 North Charles Street
Baltimore, MD 21218-2608
UNITED STATES

Prof. Dr. Ralf Zimmermann

Institut for Matematik og Datalogi
Syddansk Universitet
Campusvej 55
5230 Odense M
DENMARK