

EXTENDED ABSTRACTS

Math in Maine: A Workshop centered around Geometry, Shapes and PDEs

Organised by
Martin Bauer (FSU) and Nicolas Charon (University of Houston)

Andover, Maine from August 18 to August 24.

Topic and history of the workshop

The workshop brought together researchers in geometry, shape analysis, fluid dynamics, geometric mechanics and stochastics to discuss current research topics, work on open research problems, and develop future research directions.

The participants of the workshop were:

- (1) Martin Bauer,
- (2) Nicolas Charon,
- (3) Karen Habermann,
- (4) Emmanuel Hartman,
- (5) Boris Khesin,
- (6) Elodie Maignant,
- (7) Peter Michor,
- (8) Klas Modin
- (9) Tom Needham
- (10) Xavier Pennec,
- (11) Stephen Preston,
- (12) Stefan Sommer,
- (13) Alice Barbara Tumpach,
- (14) Francois-Xavier Vialard

This year's edition was the 11th instance of a series of workshops, which were held at varying locations around the world. For further information we refer to the workshop webpage <https://sites.google.com/view/shape-workshop>.

Acknowledgement: This workshop was supported by the National Science Foundation, through grant No. 2402555.

Overview

Martin Bauer (Florida State University)	
<i>A Dynamic Model For Optimal Riemannian Metric Transport</i>	3
Nicolas Charon (University of Houston)	
<i>K-means clustering in high-dimensional Riemannian manifolds</i>	5
Karen Habermann (University of Warwick)	
<i>Geodesic and stochastic completeness for landmark space</i>	8
Emmanuel Hartman (University of Houston)	
<i>Conic Gromov Wasserstein Distance for</i>	
<i>Unbalanced Measure Networks</i>	9
Boris Khesin (University of Toronto)	
<i>Helicity and domino tilings</i>	10
Elodie Maignant (Zuse Institute Berlin)	
<i>RNA velocity fields</i>	11
Peter Michor	
<i>The orbit of the diffeomorphism group through the Euclidean metric</i>	
<i>on $\mathbb{R}^d$</i>	12
Klas Modin (Chalmers University of Technology)	
<i>A Zeitlin discretization for axisymmetric Euler on S^3</i>	14
Tom Needham (Florida State University)	
<i>The Z-Gromov-Wasserstein Distance</i>	15
Stephen C. Preston (Brooklyn College)	
<i>Curvatures of Sobolev metrics on diffeomorphism groups</i>	17
Stefan Sommer (University of Copenhagen)	
<i>Conditionally positive kernels and landmark shape spaces with induced</i>	
<i>metrics</i>	19
Alice Barbara Tumpach (Institut CNRS Pauli, Vienna)	
<i>Riemannian metrics on Quotient manifolds</i>	20
François-Xavier Vialard	
<i>Entropic regularization and specific relative entropy</i>	23

A Dynamic Model For Optimal Riemannian Metric Transport

MARTIN BAUER (FLORIDA STATE UNIVERSITY)

In this talk, we discuss ongoing work with Peter Michor and Francois-Xavier Vialard which aims to introduce a dynamic version of unbalanced optimal transport on the space of all Riemannian metrics. The starting point of our investigations is the Wasserstein-Fisher-Rao metric [1, 2] from unbalanced optimal transport on the space of densities.

We start by very briefly recapping this framework: for (M, g_0) a Riemannian manifold we consider the space of all smooth densities $\text{Dens}(M)$. Given densities $\rho_0, \rho_1 \in \text{Dens}(M)$ we then introduce the Lagrangian

$$(1) \quad \text{Lag}(g) = \int_0^1 \int_M (g_0(v(t, \cdot), v(t, \cdot)) + \lambda f^2) \rho dt$$

which we aim to minimize over all $\rho : [0, 1] \rightarrow \text{Dens}(M)$ with $\rho(0) = \rho_0$ and $\rho(1) = \rho_1$ subject to the continuity equation

$$(2) \quad \dot{\rho}(t) = -\mathcal{L}_{v(t)}\rho(t) + f\rho.$$

Here $\lambda \in \mathbb{R}$ is a balancing parameter between the cost on the transport field v (given by the Otto metric of L^2 -optimal transport) and the cost on the source term $f\rho$ (given by the Fisher-Rao metric on the space of densities).

The goal of our present investigations is to extend the above Lagrangian to obtain a simialar framework on the space of Riemannian metrics $\text{Met}(M)$. The first step towards this goal is to generalize the corresponding continuity equation. In analogy with equation (2) it is natural to consider the continuity equation

$$(3) \quad \dot{g}(t) = -\mathcal{L}_{v(t)}g + h.$$

where $g : [0, 1] \rightarrow \text{Met}(M)$ is a path of Riemannian metrics and where $h : [0, 1] \rightarrow T_{g(t)}\text{Met}(M)$ is again a time dependent source term. Note, that h takes values in a space of matrix valued functions, which is a significant complication as compared to the situation of unbalanced optimal mass transport, where the source term is a $\mathbb{R}$ -valued function. We will use the same cost function on the transport vector field v and thus it only remains to specify a choice for the cost-function on source term. Starting from the Wasserstein-Fisher-Rao metric there is also a natural choice for this cost function on the space of Riemannian metrics, namely the Ebin Riemannian metric, which is related to the Fisher-Rao metric by the fact that the mapping $g \mapsto \text{vol}(g)$ is a Riemannian submersion.

Starting from this definition we discuss several desirable properties, that this optimal transport model on the space of Riemannian metrics possesses. In particular, we show that the mapping $g \mapsto \text{vol}(g)$ is a Riemannian submersion from the space of Riemannian metric to the space of densities, when the latter is equipped with the Wasserstein-Fisher-Rao metric. Furthermore, we obtain an analogue of

Otto's Riemannian submersion, where one has to substitute the group of diffeomorphisms with the larger space of all automorphisms of TM . Finally we discuss several open questions and future research directions.

REFERENCES

- [1] Lenaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. An interpolating distance between optimal transport and Fisher–Rao metrics. *Foundations of Computational Mathematics*, 18 (2018): 1-44.
- [2] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures. *Inventiones mathematicae*, 211.3 (2018): 969-1117.

K-means clustering in high-dimensional Riemannian manifolds

NICOLAS CHARON (UNIVERSITY OF HOUSTON)

In this talk, we discuss some ongoing work with Robert Azencott, Demetrio Labate, Andreas Mang and Ji Shi which aims at understanding the properties of the k-means clustering algorithm in certain classes of Riemannian manifolds (e.g. Cartan-Hadamard) and explore some fast numerical approaches to perform k-means clustering in manifolds of large dimension. K-means is among the most basic and common clustering algorithm in machine learning. For a dataset of points in an Euclidean space $\{x_i\}_{i=1,\dots,N}$, with $x_i \in \mathbb{R}^m$, and a given number k of clusters, it consists in looking for a partition of the dataset into k distinct groups $S = \{S_1, \dots, S_k\}$ that has minimal dispersion in the following sense:

$$(1) \quad \operatorname{argmin}_S \sum_{j=1}^k \operatorname{disp}(S_j) = \sum_{j=1}^k \sum_{x \in S_j} \|x - c_j\|^2$$

where $c_j = \frac{1}{|S_j|} \sum_{x \in S_j} x$ is the center (or mean) of the j -th cluster. Optimizing the above is known to be a NP-hard problem. However, some greedy iterative methods have been proposed to find potential suboptimal solutions, among which Lloyd's algorithm that consists in iterating, until convergence, the following cluster assignment and center update steps:

- (i) *Cluster assignment update*: assign each point x_i to the cluster S_j which center c_j is closest to x_i .
- (ii) *Cluster center update*: compute the means of the new clusters.

One can note that the assignment rule in step (i) can be interpreted as separating $\mathbb{R}^n$ into k regions corresponding to the Voronoi diagram of the centers c_j . These regions are specifically delimited by hyperplanes of $\mathbb{R}^n$ which are the mediatrices $H_{jl} = \{x \in \mathbb{R}^n : \|x - c_j\| = \|x - c_l\|\}$ between the different centers. As this algorithm is not guaranteed to converge to a global minimum of the dispersion (1), it is customary to run it with multiple initializations of the cluster centers and ultimately select the solution that yields the smallest dispersion.

Evaluation of the above dispersion and implementation of Lloyd's algorithm essentially requires two operations, namely the computation of distances and means. For that reason, it can be extended quite naturally to the setting of Riemannian manifolds provided distances and Fréchet means are well-defined. This is in particular the case of Cartan-Hadamard manifolds which are geodesically complete and for which the Fréchet mean of any set of points exists and is unique. There are however important differences with the Euclidean case when it comes to the understanding of k-means in such manifolds and the derivation of effective corresponding algorithms. We discuss a few of these in the following paragraphs as well as some open questions, by focusing specifically on the manifold $\operatorname{SDP}(n)$ of symmetric positive definite matrices of size n (with typically "large" n) equipped with the *affine-invariant metric*. This metric is defined for $P \in \operatorname{SDP}(n)$ and tangent

vectors $V, W \in \text{Symm}(n)$ by:

$$\langle V, W \rangle_P = \text{tr}(P^{-1}VP^{-1}W)$$

and its properties have been extensively studied (see [1]).

On a mathematical level, a first interesting distinction with the Euclidean situation appears when attempting to extend the concept of hyperplane to $\text{SDP}(n)$. Indeed, a first natural extension of hyperplanes are the sets $H(P, V) = \{\exp_P(W) : \langle V, W \rangle_P = 0\}$ i.e. the set of points obtained by shooting geodesics from P in directions orthogonal to V . We call these *Riemannian separators* of $\text{SDP}(n)$. Unlike in $\mathbb{R}^m$, Riemannian separators are generally not totally geodesic except for particular choices of P and V that can be characterized precisely. Aside from Riemannian separators, one can also consider *mediatrices* in $\text{SDP}(n)$ that we define, for two distinct matrices $A, B \in \text{SDP}(n)$, as $\Gamma(A, B) = \{P \in \text{SDP}(n) : d(A, P) = d(B, P)\}$ where $d(\cdot, \cdot)$ is the Riemannian distance corresponding to the affine-invariant metric. One can naturally ask whether, as with Euclidean spaces, $\Gamma(A, B)$ coincides with the Riemannian separator $H(P_{1/2}, V_{1/2})$ where $P_{1/2}, V_{1/2}$ denote respectively the midpoint between A and B and the tangent vector to the geodesic at that midpoint. This is however not the case in $\text{SPD}(n)$. A side open question is to determine more exactly what the intersection $\Gamma(A, B) \cap H(P_{1/2}, V_{1/2})$ is depending on the matrices A, B ; only partial results have been obtained so far.

From a computational point of view, the manifold case poses significant challenges for k-means clustering as each iteration of Lloyd's algorithm involves the computation of N Riemannian distances and k Fréchet means. In $\text{SPD}(n)$, the affine-invariant distance can be expressed in closed form but requires computing matrix square roots and logarithm for a numerical complexity of $O(n^3)$. Fréchet means, on the other hand, can only be approximated via optimization or iterative schemes thus demanding many evaluations of distances and their gradients. This can limit quite significantly the scope of application when considering SPD matrices of large dimension, such as the covariance matrices of high dimensional processes. In order to make k-means scale to these situations, we propose to investigate particular embeddings of $\text{SPD}(n)$ into some Euclidean spaces via so-called Fréchet maps, which can serve as an alternative to the log-Euclidean framework introduced in [2]. The (squared) Fréchet map $F : \text{SPD}(n) \rightarrow \mathbb{R}_+^p$ is defined based on a given set of p reference points $A_1, \dots, A_p \in \text{SPD}(n)$ by:

$$F(P) = (d(A_1, P)^2, \dots, d(A_p, P)^2)$$

Some empirical evidence suggests that applying Euclidean k-means in $\mathbb{R}^p$ to the images of the data points by such Fréchet maps can lead, in certain cases, to results comparable to the Riemannian k-means approach but at a significantly lower computational cost. This, however, strongly depends on the choice of the reference points $A_1, \dots, A_p$. The central question we wish to address is to derive some theoretical properties of the Fréchet map that could provide heuristics on how to select reference points.

Among the properties we consider is the injectivity of the Fréchet map. It is indeed easy to show that the Fréchet map with $m+1$ points where $m = n(n+1)/2$

is the dimension of $\text{SPD}(n)$ is injective as long as the $m + 1$ reference points do not lie on a common mediatix, and it is a immersion if the reference points do not lie on a common Riemannian separator of $\text{SPD}(n)$. Yet those conditions remain implicit and it is not clear how to test them in practice. We also ask the question of whether such configurations are generic, in the sense that for random choices of $m + 1$ reference points, the corresponding Fréchet map is an embedding into $\mathbb{R}^{m+1}$.

REFERENCES

- [1] X. Pennec. *Manifold-valued image processing with SPD matrices*, chap.3 of *Riemannian Geometric Statistics in Medical Image Analysis*. Elsevier, 2020.
- [2] V. Arsigny, P. Fillard, X. Pennec and N. Ayache. *Log-Euclidean metrics for fast and simple calculus on diffusion tensors*. Magnetic Resonance in Medicine, vol. 56, 2006.

Geodesic and stochastic completeness for landmark space

KAREN HABERMANN (UNIVERSITY OF WARWICK)

The manifold Q of $n \geq 2$ distinct landmarks in $\mathbb{R}^d$, for $d \geq 1$, given by

$$Q = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}^d \text{ with } x_i \neq x_j \text{ for } i \neq j\} \subset \mathbb{R}^{nd}$$

can be endowed with a Riemannian metric that descends from a right-invariant metric on a diffeomorphism group of $\mathbb{R}^d$. For a suitable positive definite kernel $K: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$, that depends on the chosen right-invariant metric upstairs, the above construction amounts to defining a metric g on the space Q through specifying its inverse by, for $i, j \in \{1, \dots, n\}$,

$$g^{ij}(x_1, \dots, x_n) = K(x_i, x_j).$$

Relevant questions which now arise naturally in the context of statistical shape analysis concern the geodesic as well as stochastic completeness of the Riemannian manifold that is the landmark configuration space (Q, g) .

For the subsequent discussion, we shall further assume that the kernel K is both rotationally and translationally invariant, so that there exists a continuous scalar function $k: [0, \infty) \rightarrow \mathbb{R}$ such that

$$K(x_i, x_j) = k(\|x_i - x_j\|)I_d,$$

with $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^d$ and I_d the $d \times d$ identity matrix. Under this assumption, which covers most important examples, and for the case of exactly two landmarks, the recent article [1] provides a sharp condition for the scalar function k to deduce stochastic completeness or stochastic incompleteness, respectively, of the landmark space (Q, g) . Note that Brownian motion on (Q, g) can fail to exist for all times due to the collision of two landmarks or as a result of one landmark running off to infinity in finite time with positive probability.

Ongoing joint work [2] with Stephen C. Preston and Stefan Sommer establishes a sharp criterion for any number of landmarks and in terms of the function k which guarantees geodesic completeness or geodesic incompleteness, respectively, of the landmark configuration space (Q, g) . As for stochastic completeness, geodesic completeness is ensured by ruling out both collision of landmarks and landmarks running off to infinity in finite time.

For a geodesically complete landmark space there is then further a potential road towards proving stochastic completeness by suitably controlling the volume growth of growing geodesic annuli.

REFERENCES

- [1] Karen Habermann, Philipp Harms, and Stefan Sommer. Long-time existence of Brownian motion on configurations of two landmarks. *Bulletin of the London Mathematical Society*, 56(5):1658–1679, 2024.
- [2] Karen Habermann, Stephen C. Preston, Stefan Sommer. Characterization of geodesic completeness on landmark space. *In preparation*, 2024.

Conic Gromov Wasserstein Distance for Unbalanced Measure Networks

EMMANUEL HARTMAN (UNIVERSITY OF HOUSTON)

We review of the Gromov-Wasserstein (GW) distance for measure networks [1] and introduce an extension to unbalanced measures. The GW distance between two measure networks $H = (X, \omega_X, \mu_X)$ and $H' = (Y, \omega_Y, \mu_Y)$, where μ_X and μ_Y are probability measures on X and Y respectively, is defined as:

$$\text{GW}(H, H')^2 = \inf_{\gamma \in \Gamma(\mu_X, \mu_Y)} \int_{(X \times Y)^2} |\omega_X(x, x') - \omega_Y(y, y')|^2 d\gamma(x, y) d\gamma(x', y').$$

To address measure networks with unbalanced measures, we define the **conic Gromov-Wasserstein (CGW)** distance, formulated as:

$$\begin{aligned} \text{CGW}(H, H') = \inf_{\gamma \in \bar{\Gamma}(\mu_X, \mu_Y)} & \int_{X \times Y} \int_{X \times Y} \frac{d\gamma_0}{d\gamma}(x, y) \frac{d\gamma_0}{d\gamma}(x', y') + \frac{d\gamma_1}{d\gamma}(x, y) \frac{d\gamma_1}{d\gamma}(x', y') \\ & - 2\overline{\cos}(|\omega_X(x, x') - \omega_Y(y, y')|) \sqrt{\frac{d\gamma_0}{d\gamma}(x, y) \frac{d\gamma_1}{d\gamma}(x', y')} d\gamma(x, y) d\gamma(x', y'). \end{aligned}$$

When ω_X and ω_Y are distance functions, this formulation recovers the CGW metric as defined in [2]. Additionally, we review the concept of **weak isomorphism** between measure networks from [1]. Two measure networks $H = (X, \omega_X, \mu_X)$ and $H' = (Y, \omega_Y, \mu_Y)$ are weakly isomorphic (denoted $H \sim H'$) if there exists a measure space Z and two functions $f : Z \rightarrow X$ and $g : Z \rightarrow Y$ such that:

$$\begin{aligned} g_*\mu_Z &= \mu_Y, \\ f_*\mu_Z &= \mu_X, \\ \|f^*\omega_X - g^*\omega_Y\|_\infty &= 0. \end{aligned}$$

We show that the CGW distance defines a valid metric on the space of measure networks modulo weak isomorphism, analogous to the result of [1] with showed GW forms a distance on the space of probability networks modulo weak isomorphism. However, the structure and geometry of this metric space, such as whether it forms a geodesic space, remain open areas of investigation.

REFERENCES

- [1] Samir Chowdhury, Facundo Mémoli, The Gromov-Wasserstein distance between networks and stable network invariants, *Information and Inference: A Journal of the IMA*, Volume 8, Issue 4, December 2019, Pages 757–787, <https://doi.org/10.1093/imaiai/iaz026>
- [2] Séjourné, T., Vialard, F., and Peyré, G. (2020). The Unbalanced Gromov Wasserstein Distance: Conic Formulation and Relaxation. *Neural Information Processing Systems*.

Helicity and domino tilings

BORIS KHESIN (UNIVERSITY OF TORONTO)

Domino tilings of 3D regions traditionally have two invariants associated to them, flux and twist. Flux is understood as the homology class associated with a certain cycle constructed for the tiling. For a cubulated region M (a topological 3-manifold with boundary) the flux $Flux(t)$ of a tiling t is an element of the first homology group of the region, defined up to an additive constant. The ambiguity can be removed by considering the *relative flux* $RFlux(t) \in H_1(M, \partial M; \mathbb{R})$. For tilings with zero relative flux there is an integer invariant, the *twist*, associated with the tiling and measuring the "mutual linking of tiles" around each other. It is invariant with respect to *flips*, local moves which consist of removing two neighboring parallel dominoes and placing them back after a rotation. The twist changes under another move, a *trit*, which replaces a frame-like triple of tiles to the one pointing in the opposite way.

While there have been pointed out similarities of the twist invariant with the Hopf invariant, the correspondence remained at either an intuitive level or in the continuous limit for turning tiles into vector fields via a broadly understood tiling's refinement.

In this paper we present a construction of a smooth divergence-free vector field associated to an arbitrary tiling ("5-pipe construction") so that *the twist invariant becomes, up to a factor, the relative helicity* of that vector field. Furthermore, we extend the notion of relative helicity to divergence-free vector fields on arbitrary three-manifolds and not necessarily tangent to their boundaries. This allows one to compare relative helicity with twists of tilings in non-simply-connected regions. The toolbox includes an introduction of "an isolating shell" for a cubulated region, the use of refinements and appropriate connectivity of the spaces of tilings. Finally, we relate *the flux invariant of a tiling to the rotation class* of the associated vector field. This is joint work with Nicolau Saldanha, see [1].

REFERENCES

- [1] B. Khesin, and N. C. Saldanha. Relative Helicity and Tiling Twist. arXiv preprint arXiv:2408.00522 (2024).

RNA velocity fields

ELODIE MAIGNANT (ZUSE INSTITUTE BERLIN)

In this talk, we discuss a new trajectory inference problem arising from recent advances in single-cell RNA sequencing. Single-cell sequencing is a technique for simultaneously recovering RNA sequences from a large population of cells. The data collected are available as a set of vectors $x_i \in \mathbb{R}^n$ where x_i represents how much each of n given genes is expressed in the RNA sequence of the i -th cell. The trajectory inference problem is that of recovering from such transcriptomic data the dynamic process (e.g. the differentiation process) that governs the evolution of the cell population. More precisely, it consists in ordering the cells according their stage of the process. Indeed, since single-cell sequencing is a destructive method, the evolution stage t_i at which each cell was observed is not known. However, assuming that the changes undergone by the cells are governed by a common process, one should be able to recover such stages from the set of all vectors x_i . Now, in addition to these expression vectors, we investigate what are known as RNA velocities. These are other vectors $v_i \in \mathbb{R}^n$ quantifying the infinitesimal change in expression of each gene in the RNA sequence of the i -th cell after stage t_i . Eventually, our data is given as a discrete vector field (x_i, v_i) .

Our problem is thus to recover the dynamics of the cell population from such a vector field. To the best of our knowledge, current methods incorporating RNA velocities do not fully leverage the nature of the data, that is that of a coupling between points and tangent vectors. In any case, such data present two main challenges. The first one is their high dimension (the number of genes being sometimes as large as the number of cells). The second is their underlying tree structure, a natural hypothesis for modeling the process of cell differentiation. Indeed, the combination of the two makes learning tasks as simple as interpolating and integrating the discrete vector field (x_i, v_i) – a first approach mentioned in this discussion – tricky. While the literature has addressed the first limitation by applying dimension reduction methods, the tree structure questions in fact the efficiency of such methods, a priori optimized for more uniformly distributed data sets. In the absence of any assumption on the regularity of the vector field, we identify the need to take full advantage of the tree structure. Finally, an additional challenge typically encountered in data analysis is the noisy nature of the data. To answer our problem, we conclude on several approaches covering tree inference, stochastic processes and optimal transport. This is based on ongoing work with Christoph von Tycowicz.

**The orbit of the diffeomorphism group through the Euclidean metric
on $\mathbb{R}^d$**

PETER MICHOR

Theorem. *The orbit of the pullback action of $\text{Diff}_{\mathcal{A}}(\mathbb{R}^d)$ through the Euclidean metric in $\text{Met}_{\mathcal{A}}(\mathbb{R}^d)$ consists of all flat metrics in $\text{Met}_{\mathcal{A}}(\mathbb{R}^d)$.*

Here $\mathcal{A} \in \{\mathcal{S}, H^\infty, \mathcal{D}\}$ among other choices, where $\mathcal{S}$ denotes any space of rapidly decreasing functions, $H^\infty = \bigcap_{k \in \mathbb{N}} H^k$ is the intersection of all Sobolev spaces, and $\mathcal{D}$ denotes any space of functions with compact support. Moreover, we put

$$\begin{aligned} \text{Diff}_{\mathcal{A}}(\mathbb{R}^d) &= \{\varphi = \text{Id} + f : f \in \mathcal{A}(\mathbb{R}^d, \mathbb{R}^d), \det(df) > -1\} \\ \text{Met}_{\mathcal{A}}(\mathbb{R}^n) &= \{g = \bar{g} + h : h \in \mathcal{A}(\mathbb{R}^n, L^2_{\text{sym}}(\mathbb{R}^n, \mathbb{R})), g \text{ positive}\} \end{aligned}$$

the Lie group of diffeomorphisms of $\mathbb{R}^d$ which differ from the identity by an element in $\mathcal{A}$, and the smooth manifold of all Riemannian metrics which differ from the Euclidean metric $\bar{g}$ by an element in $\mathcal{A}$.

Proof. For the curvature we have $R^{\varphi^* \bar{g}} = \varphi^* R^{\bar{g}} = \varphi^* 0 = 0$, so the orbit consists of flat metrics. It remains to prove that each flat metric is in the orbit. So let $g \in \text{Met}_{\mathcal{A}}^{\text{flat}}(\mathbb{R}^n)$ be a flat metric. Considering g as a symmetric positive matrix, let $s := \sqrt{g}$. We search for an orthogonal matrix valued function $u \in C^\infty(\mathbb{R}^n, SO(n))$ such that $u.s = d\varphi$ for a diffeomorphism φ . For the following see [1, Section 25].

Let $\sigma_i := \sum_j s_{ij} dx^j$ be the rows of s . Then for the metric we have $g = \sum_i \sigma_i \otimes \sigma_i$, thus the column vector $\sigma = (\sigma_1, \dots, \sigma_n)^t$ of 1-forms is a global orthonormal coframe. We want $u.\sigma = d\varphi$, so the 2-form $d(u.\sigma)$ should vanish. But

$$0 = d(u.\sigma) = du \wedge \sigma + u.d\sigma \iff 0 = u^{-1}.du \wedge \sigma + d\sigma$$

This means that the $\mathfrak{o}(n)$ -valued 1-form $\omega := u^{-1}.du$ is the connection 1-form for the Levi-Civita connection of the metric g . Since g is flat, the curvature 2-form $\Omega = d\omega + \omega \wedge \omega$ vanishes. We consider now the trivial principal bundle $\text{pr}_1 : \mathbb{R}^n \times SO(n) \rightarrow \mathbb{R}^n$ and the principal connection form $\text{pr}_1^* \omega$ on it which is flat, so the horizontal distribution is integrable. Let $L(u_0) \subset \mathbb{R}^n \times SO(n)$ be the horizontal leaf through the point $(0, u_0) \in \mathbb{R}^n \times SO(n)$, then the restriction $\text{pr}_1 : L(u_0) \rightarrow \mathbb{R}^n$ is a covering map and thus a diffeomorphism whose inverse furnishes us the required $u \in C^\infty(\mathbb{R}^n, SO(n))$ which is unique up to right multiplication by $u_0 \in SO(n)$. The function $u : \mathbb{R}^n \rightarrow SO(n)$ is also called the Cartan development for ω .

Thus $u.\sigma = d\varphi$ for a column vector $\varphi = (\varphi^1, \dots, \varphi^n)$ of functions which defines a smooth map $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Since $d\varphi = u.\sigma$ is everywhere invertible, φ is locally a diffeomorphism. Since g falls to $\bar{g} = \mathbb{I}_n$ as a function in $\mathcal{A}$, the same is true for σ and thus also for $u.\sigma$ since u is bounded. So $d(\varphi - \text{Id}_{\mathbb{R}^n})$ is asymptotically 0, thus $\varphi - \text{Id}_{\mathbb{R}^n}$ is asymptotically a constant matrix A ; here we need $n \geq 2$. Replacing φ by $\varphi - A$ we see that φ then falls asymptotically towards $\text{Id}_{\mathbb{R}^n}$. Thus φ is a

proper mapping and thus has closed image, which is also open since φ is still a local diffeomorphism. Thus $\varphi \in \text{Diff}_{\mathcal{A}}(\mathbb{R}^n)$.

Finally, $d\varphi^t.d\varphi = (u.\sigma)^t.u.\sigma = \sigma^t.u^t.u.\sigma = \sigma^t.\sigma = g$. Note that φ is unique in $\text{Diff}_{\mathcal{A}}(\mathbb{R}^n)$. This is also clear from the fact, that the fiber of $\text{Pull}^{\bar{g}}$ over $\varphi^*\bar{g}$ consists of all isometries of $\varphi^*\bar{g}$ which is the group $\varphi^{-1} \circ (\mathbb{R}^n \ltimes O(n)) \circ \varphi \subset \text{Diff}(\mathbb{R}^n)$ whose intersection with $\text{Diff}_{\mathcal{A}}(\mathbb{R}^n)$ is trivial. $\square$

REFERENCES

- [1] P. W. Michor. *Topics in differential geometry*, volume 93 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.

A Zeitlin discretization for axisymmetric Euler on S^3

KLAS MODIN (CHALMERS UNIVERSITY OF TECHNOLOGY)

The Hopf fibration $\pi: S^3 \rightarrow S^2$ gives rise to “axisymmetry” for Euler’s incompressible equations on S^3 . Namely, a divergence free vector field u on S^3 is axisymmetric if it is equivariant under the S^1 action on S^3 . The infinitesimal version of this conditions is that u commutes with the Hopf field. Together with the divergence free condition $\operatorname{div} u = 0$, we can then describe u in terms of two functions ψ and σ on S^2 . The “2-D stream function” ψ corresponds to the vector field on the base S^2 , whereas the “swirl function” σ corresponds to the speed along the fiber directions. Lichtenfelz, Misiołek, and Preston [1] showed that the axisymmetric Euler equations on S^3 , expressed in these variables, have the beautiful form

$$\Delta \dot{\psi} + \{\psi, \Delta \psi + \sigma\} = 0, \quad \dot{\sigma} + \{\psi, \sigma\} = 0.$$

This form of the equations immediately reveals how to discretize them via Zeitlin’s approach: apply quantization on S^2 to the functions ψ and σ , which yields corresponding matrices $P \in \mathfrak{su}(n)$ and $\Sigma \in \mathfrak{u}(n)$. The corresponding matrix hydrodynamic version of the equations then reads

$$\Delta_n \dot{P} + \frac{1}{\hbar} [P, \Delta_n P + \Sigma] = 0, \quad \dot{\Sigma} + \frac{1}{\hbar} [P, \Sigma] = 0,$$

where Δ_n is the Hoppe-Yau quantized Laplacian $\mathfrak{u}(n) \rightarrow \mathfrak{su}(n)$, $[\cdot, \cdot]$ is the matrix commutator, and $\hbar = 1/\sqrt{n(n-1)}$.

The great expectation is that these matrix equations shall provide a useful numerical tool for guiding the analysis of finite time blow up. As a first step, in a recent preprint together with Preston [2], we studied their underlying Euler-Arnold structure and gave some results on sectional curvature.

REFERENCES

- [1] L. Lichtenfelz, G. Misiołek, and S. C. Preston. Axisymmetric diffeomorphisms and ideal fluids on Riemannian 3-manifolds. *Int. Math. Res. Not.*, 2022(1):446–485, 2022.
- [2] K. Modin and S. C. Preston. Zeitlin’s model for axisymmetric 3-D Euler equations. *arXiv:2408.11204*, 2024.

The Z -Gromov-Wasserstein Distance

TOM NEEDHAM (FLORIDA STATE UNIVERSITY)

The Gromov-Wasserstein (GW) distances, introduced by Mémoli in [4], comprise a family of metrics on the space of *metric measure spaces* (*mm-spaces*), or triples of the form $\mathcal{X} = (X, d_X, \mu_X)$, where (X, d_X) is a separable, complete metric space and μ_X is a probability measure on X . For $p \in [1, \infty)$, the *Gromov-Wasserstein p -distance* between mm-spaces $\mathcal{X}$ and $\mathcal{Y}$ is

$$\text{GW}_p(\mathcal{X}, \mathcal{Y})^p = \inf_{\pi} \frac{1}{2} \iint_{(X \times Y)^2} |d_X(x, x') - d_Y(y, y')|^p \pi(dx \times dy) \pi(dx' \times dy'),$$

where the infimum is over *couplings*—probability measures π on $X \times Y$ whose left and right marginals are μ_X and μ_Y , respectively. This defines a notion of optimal transport between measures supported on *different* metric spaces, and has applications in shape analysis and machine learning, where an optimal coupling provides a soft registration between a priori incomparable spaces such as point clouds or Riemannian manifolds. The formula for GW distance still provides a meaningful pseudometric between more general structures of the form (X, ω_X, μ_X) , where $\omega_X : X \times X \rightarrow \mathbb{R}$ is only assumed to be a measurable map, rather than a metric [3]. For example, one can model a graph as a structure of this form by taking X to be its node set, μ_X to be the uniform measure, and ω_X to be some standard graph kernel, such as the graph’s (weighted) adjacency function or Laplacian. In this setting, computing the GW distance amounts to solving a relaxed graph matching problem.

In applications, a graph frequently comes with additional structure such as node or edge features. For example, edges in a social network are typically attributed with vectors in $\mathbb{R}^n$ encoding user interaction data. As another example, there has been significant recent interest in *shape graphs*, a formalism for modeling filamentary structures, such as arterial networks, as graphs with edges attributed in a space of curves (endowed with some metric coming from the field of shape analysis), encoding the ‘shape’ of each edge [1].

In recent work with Bauer, Mémoli and Nishino [2], we define a GW-like distance which is designed to handle edge-attributed graphs, as well as more general structures. For a fixed separable metric space (Z, d_Z) , we define a *Z -network* to be a triple of the form $\mathcal{X} = (X, \omega_X, \mu_X)$, where X is a Polish space, μ_X is a Borel probability measure, and $\omega_X : X \times X \rightarrow Z$ is a measurable Z -valued kernel. The *Z -Gromov-Wasserstein p -distance* between Z -networks $\mathcal{X}$ and $\mathcal{Y}$ is defined as

$$\text{GW}_p^Z(\mathcal{X}, \mathcal{Y})^p = \inf_{\pi} \frac{1}{2} \iint_{(X \times Y)^2} d_Z(\omega_X(x, x'), \omega_Y(y, y'))^p \pi(dx \times dy) \pi(dx' \times dy'),$$

where the infimum is once again over couplings of μ_X and μ_Y . For graphs with edge attributes valued in Z , this notion of distance gives a method of comparison whose computation results in an attribute-aware soft registration (i.e., graph matching). Taking Z to be $\mathbb{R}^n$ or a space of curves, respectively, recovers the concrete settings

described in the previous paragraph. This construction is closely related to that of the recent work [5]; please refer to our paper for a precise comparison.

We derive several results in [2] on the metric properties of the Z -GW distance, summarized in the following theorem.

Theorem 1 ([2]). *The Z -Gromov-Wasserstein p -distance defines a pseudometric on the space of Z -networks. Pairs of Z -networks with distance zero have an explicit geometric characterization. The induced metric space is complete if and only if Z is complete, is always path connected, and is geodesic if Z is geodesic.*

We additionally show that the Z -GW distance encompasses several variants of GW distance which have appeared previously in the literature. There has been a trend in the recent machine learning literature to introduce a variant of GW distance for handling a particular data structure, and to then derive its properties via proofs which follow a standard structure. A motivation for our work was to provide a general framework for reasoning about metric properties of GW-like metrics, in order to reduce such redundancies. Please refer to our paper for a description of these connections.

This talk posed several open questions about the Z -GW distance:

- Is GW_1^Z always a geodesic space? This would mirror the situation for Wasserstein distances on arbitrary metric spaces.
- Let $p > 1$. If GW_p^Z is geodesic, does this imply that Z is geodesic?
- Our proof that geodesicity of Z implies geodesicity of GW_p^Z is constructive. In this case, are all GW_p^Z geodesics of the form given by our construction? An affirmative answer would allow us to derive Alexandrov curvature bounds for GW_p^Z .
- Are there interesting precompact families of Z -networks (cf. Gromov's celebrated precompactness theorem)?

REFERENCES

- [1] Aditi Basu Bal, Xiaoyang Guo, Tom Needham, Anuj Srivastava. Statistical Analysis of Complex Shape Graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024 (available online; awaiting publication data).
- [2] Martin Bauer, Facundo Mémoli, Tom Needham, Mao Nishino. The Z -Gromov-Wasserstein Distance. *arXiv preprint arXiv:2408.08233*, 2024.
- [3] Samir Chowdhury, Facundo Mémoli. The Gromov-Wasserstein distance between networks and stable network invariants. *Information and Inference: A Journal of the IMA*, Volume 8, Issue 4, 2019, pp. 757–787.
- [4] Facundo Mémoli. On the use of Gromov-Hausdorff Distances for Shape Comparison. *Eurographics Symposium on Point-Based Graphics*, 2007, pp. 1811–7813.
- [5] Junjie Yang, Matthieu Labeau, and Florence d'Alché-Buc. Exploiting Edge Features in Graphs with Fused Network Gromov-Wasserstein Distance. *arXiv preprint arXiv:2309.16604*, 2023.

Curvatures of Sobolev metrics on diffeomorphism groups

STEPHEN C. PRESTON (BROOKLYN COLLEGE)

The original motivation for studying geometry of diffeomorphism groups was to use curvature to understand the observed instability of fluid mechanics: an ideal fluid with no external forces follows a geodesic trajectory in the group of volume-preserving diffeomorphisms, under a right-invariant metric determined by the kinetic energy. Hence negative curvature implies instability via exponential divergence of nearby geodesics. However in the original paper of Arnold [1], it was immediately discovered that although the sectional curvature is negative in many sections, it can also sometimes be positive. Furthermore the sectional curvature is not bounded, which is essentially due to the fact that the Riemannian metric is too weak to generate the topology that makes the diffeomorphism group a smooth manifold (L^2 topology for the metric vs. at least C^1 topology for the manifold structure). Furthermore, the sectional curvature on a general Lie group with a left- or right-invariant metric is given by a rather complicated formula, about which it has been difficult to draw general conclusions.

In particular this makes it difficult to apply the most straightforward theorems of Riemannian geometry, which are typically of the form that a uniform bound on all sectional curvatures has topological implications for the manifold. In the case of a strong metric, such as a Sobolev metric of order s on the manifold M , where $s > \frac{1}{2}\dim(M) + 1$, the metric generates a topology where the diffeomorphism group is a smooth manifold, and for this reason the curvature and all other geometric quantities are uniformly bounded. Our question is when this works for weaker metrics.

The simplest situation is the degenerate $\dot{H}^1$ metric on $\text{Diff}(S^1)$, which descends to a true Riemannian metric on the quotient space $\text{Diff}(S^1)/S^1$ (modulo rotations). Here the quotient space is isometric to a portion of the round infinite-dimensional sphere [2], and so the curvature is a positive constant (and in particular it is bounded). This is the only known situation where the curvature takes only one sign, illustrating the difficulty of analyzing curvature for diffeomorphism groups.

The simplest case is the one-dimensional case, either $\text{Diff}(S^1)$ or $\text{Diff}(\mathbb{R})$ with some reasonable decay conditions in the latter case. Here the Sobolev metric H^s becomes a strong metric (and all boundedness results are automatic) when $s > \frac{3}{2}$. Meanwhile the geodesic equation is not well-posed unless $s \geq \frac{1}{2}$, so it is unreasonable to expect any boundedness for $s < \frac{1}{2}$. Our conjecture is the following:

Conjecture 1. *On the diffeomorphism group $\text{Diff}(S^1)$ with the right-invariant Sobolev H^s or degenerate $\dot{H}^s$ metric, let $K(u, v)$ be the sectional curvature of the section spanned by vectors $u, v \in T_{\text{id}} \text{Diff}(S^1)$.*

- *For $s < \frac{1}{2}$, the sectional curvature $K(u, v)$ is unbounded as a function of either u or v , even if one vector is held fixed.*
- *For $\frac{1}{2} \leq s < 1$ and for each fixed smooth u , the sectional curvature $K(u, v)$ is bounded as a function of v , but unbounded if both u and v can vary.*

- For $s \geq 1$, the sectional curvature $K(u, v)$ is uniformly bounded in both u and v .

Unboundedness is easy to prove: one simply constructs examples with explicit formulas. Boundedness is substantially harder to prove. Known cases for boundedness include $s = 1$ for the $\dot{H}^1$ metric mentioned above, along with the case $s > \frac{3}{2}$. I discussed a strategy that Alice Le Brigant, Martin Bauer, and I are pursuing to prove boundedness in the $\dot{H}^s$ case in general by considering the Sobolev metric as a quotient of the group of horizontal planar diffeomorphisms. The Lie algebra here is the space of vector fields $u(x, y) \frac{\partial}{\partial x}$ corresponding to diffeomorphisms that fix each horizontal line, and the right-invariant metric is given by

$$\langle u, u \rangle_\alpha = \int_{\mathbb{R}^2} \frac{u_x(x, y)^2 + u_y(x, y)^2}{y^\alpha} dx dy.$$

It can be shown that if $\alpha = 2s - 1$ for $\frac{1}{2} \leq s < 1$, the quotient map, obtained by projecting a horizontal diffeomorphism to the restricted diffeomorphism of the x -axis, is a Riemannian submersion. Hence one can compute the curvature in the simpler α -metric and use the O'Neill formula for curvature of a Riemannian submersion to hopefully prove both parts of this conjecture.

REFERENCES

- [1] V.I. Arnold, On the differential geometry of infinite-dimensional Lie groups and its application to the hydrodynamics of perfect fluids. Vladimir I. Arnold-Collected Works: Hydrodynamics, Bifurcation Theory, and Algebraic Geometry 1965-1972, pp.33-69, 2014.
- [2] J. Lenells, The Hunter-Saxton equation describes the geodesic flow on a sphere. Journal of Geometry and Physics, 57(10), pp.2049-2064, 2007.

Conditionally positive kernels and landmark shape spaces with induced metrics

STEFAN SOMMER (UNIVERSITY OF COPENHAGEN)

Joint work with Peter Michor and Sarang Joshi.

We present a unification of the landmark shape spaces of Kendall, that remove translation, rotations, and scale from landmark configurations with initial Euclidean geometry, with landmark configuration spaces equipped with Riemannian metrics descending from right-invariant Sobolev metrics on the diffeomorphism group. This is possible when the metric is invariant to the adjoint action of the rigid motion group and potentially rescaling, so that the Riemannian metric descends first to a quotient of the diffeomorphism group, and subsequently to quotients of the landmark space. We explore the resulting geometry that is intrinsically linked to conditionally positive kernels and Beppo-Levi spaces. The construction will allow the use of landmark configuration spaces with induced metrics in applications including evolutionary morphology and medical imaging while still keeping the shape invariances that are a hallmark of Kendall's shape spaces.

Riemannian metrics on Quotient manifolds

ALICE BARBARA TUMPACH (INSTITUT CNRS PAULI, VIENNA)

The following is part of a joint work with S.C. Preston (see [2], [3]) and was inspired by considerations in shape analysis (see [1]).

Consider a (possibly infinite-dimensional) Lie group G and K an immersed Banach–Lie subgroup of G with Banach–Lie subalgebra $\mathfrak{k} \subset \mathfrak{g}$. Suppose that the quotient space G/K for the right action of K on G

$$G/K = \{gK, g \in G\}$$

is a Hausdorff Banach manifold. This is for instance the case when K is a split Banach–Lie subgroup (i.e the Lie algebra $\mathfrak{k}$ is closed in $\mathfrak{g}$ and has a closed complement), but we will consider the more general case. We will denote by π the quotient map: $\pi(g) = gK$.

Definition 1. The normal bundle of G/K is the fiber bundle over G whose fiber at $g \in G$ is given by

$$\text{Nor}_g = T_g G / \ker T_g \pi,$$

where $T_g \pi$ denotes the differential of the canonical projection π at $g \in G$.

Proposition 1. ([1]) *The tangent bundle $T(G/K)$ of the homogeneous space G/K is canonically isomorphic to the quotient Nor/K of the normal bundle Nor by the right action of K . In particular, for any $g \in G$ and any $X \in T_g G$, the K -orbit of $[X] \in T_g G / \ker T_g \pi$ is identified with $T_g \pi(X) \in T_{\pi(g)} G/K$.*

Proof. Note that $\ker T_g \pi$ is the vector space generated by the infinitesimal action of K on $g \in G$. In particular, $R_k \ker T_g \pi = \ker T_{gk} \pi$, where R_k denotes the differential of the right translation by $k \in K$. Hence the right action of K on TG induces a right action of K on the normal bundle Nor . Since π is a submersion, for any $g \in G$ such that $\pi(g) = p$, the quotient Banach space $T_g G / \ker T_g \pi$ is isomorphic to $T_p(G/K)$. Moreover $\pi(g) = \pi(h) = p$ if and only if $h = gk$ for some $k \in K$. Hence $T(G/K)$ is isomorphic to the quotient Nor/K . $\square$

Remark 1. For $g \in G$, the dual space to the normal space $\text{Nor}_g = T_g G / \ker T_g \pi$ at $g \in G$ is canonically isomorphic to $(\ker T_g \pi)^0$ where

$$(\ker T_g \pi)^0 = \{f \in (T_g G)^*, f|_{\ker T_g \pi} = 0\}.$$

Notation 1. We will denote by $(\cdot, \cdot)_{\text{Nor}/K, (\text{Nor})^*/K}$ the duality pairing between Nor/K and its dual bundle $(\text{Nor})^*/K$.

Remark 2. A Riemannian metric on the quotient space G/K is a family of positive-definite bilinear forms $g : T_{[g]}(G/K) \times T_{[g]}(G/K) \rightarrow \mathbb{R}$ depending smoothly on the foot point $[g] \in G/K$ and is therefore equivalent to a vector bundle map

$$A : \text{Nor}/K \rightarrow (\text{Nor}/K)^* \simeq (\ker T\pi)^0 / K,$$

such that

$$\begin{aligned} \text{Nor}/K \times \text{Nor}/K &\rightarrow \mathbb{R} \\ (X, Y) &\mapsto (X, AY)_{\text{Nor}/K, \text{Nor}^*/K} \end{aligned}$$

is symmetric positive definite.

Proposition 2. *If $\pi : G \rightarrow G/K$ admits a globally defined section $s : G/K \rightarrow G$, then $A : \text{Nor}/K \rightarrow (\text{Nor}/K)^*$ can be lifted to a K -invariant vector bundle map $\tilde{A} : \text{Nor} \rightarrow \text{Nor}^* \simeq (\ker T\pi)^0$, by defining $\tilde{A}$ on the range of the section s as*

$$\tilde{A}_{s(\pi(g))}(X_{s(\pi(g))}) = A_{\pi(g)}T_g\pi(X_{s(\pi(g))}),$$

for $X_{s(\pi(g))} \in \text{Nor}_{s(\pi(g))} = T_{s(\pi(g))}G/\ker T_{s(\pi(g))}\pi$ and extending it by K -invariance to the whole fiber bundle Nor over G .

Definition 2. Gauge invariant metrics were introduced in [4] (see also [5]) and consist of degenerate metrics $\tilde{g} : TG \times TG \rightarrow \mathbb{R}$ on a fiber bundle $\pi : G \rightarrow G/K$, whose kernel at $g \in G$ coincides with the vertical space $\ker T_g\pi$ and which descend to a Riemannian metric on G/K .

Theorem 2. *Consider a quotient manifold G/K endowed with a Riemannian metric g and suppose that $\pi : G \rightarrow G/K$ admits a globally defined smooth section $s : G/K \rightarrow G$. Denote by $\mathcal{S}$ the smooth manifold $\mathcal{S} = s(G/K) \subset G$. Then*

- (1) *the Riemannian metric g naturally induces a Riemannian metric $g^{\mathcal{S}}$ on $\mathcal{S}$ defined by*

$$g^{\mathcal{S}}(X_{s(\pi(g))}, Y_{s(\pi(g))}) = g(T_{s(\pi(g))}\pi(X_{s(\pi(g))}), T_{s(\pi(g))}\pi(Y_{s(\pi(g))})),$$

where $X_{s(\pi(g))}, Y_{s(\pi(g))} \in T_{s(\pi(g))}\mathcal{S}$.

- (2) *the Riemannian metric g naturally induces a gauge invariant metric $\tilde{g}$ on G defined by*

$$\tilde{g}(X_g, Y_g) = g(T_{s(\pi(g))}\pi(P(R_{k^{-1}}X_g)), T_{s(\pi(g))}\pi(P(R_{k^{-1}}Y_g)))$$

where $g \in G$ decomposes as $g = s(\pi(g))k$ for a unique $k \in K$, and $X_g, Y_g \in T_gG$, and $P : T_{s(\pi(g))}G \rightarrow T_{s(\pi(g))}\mathcal{S}$ is the projection on $T_{s(\pi(g))}\mathcal{S}$ relative to the decomposition:

$$T_{s(\pi(g))}G = T_{s(\pi(g))}\mathcal{S} \oplus \ker T_{s(\pi(g))}\pi.$$

Note that $\ker T_g\pi$ is exactly the kernel of $\tilde{g}$ at $g \in G$.

REFERENCES

- [1] I. Ciuclea, A.B. Tumpach, C. Vizman, Shape spaces of nonlinear flags, International Conference on Geometric Science of Information, 41–50, 2023, Springer, DOI:10.1007/978-3-031-38271-0_5
- [2] A.B. Tumpach and S.C. Preston, *Three methods to put a Riemannian metric on Shape Space*, Geometric Science of Information, 6th International Conference, GSI 2023, Proceedings, Part I, 3–11, Springer, 2023. DOI:10.1007/978-3-031-38271-0_1
- [3] A.B. Tumpach and S.C. Preston, *Riemannian metrics on shape spaces: comparison of different constructions*, in preparation.

- [4] A.B. Tumpach, H. Drira, M. Daoudi, A. Srivastava, *Gauge Invariant Framework for Shape Analysis of Surfaces*. IEEE Transactions on Pattern Analysis and Machine Intelligence, January 2016, Volume 38, Number 1. DOI:10.1109/TPAMI.2015.2430319.
- [5] A.B. Tumpach, *Gauge Invariance of degenerate Riemannian metrics*, Notices of American Mathematical Society, April 2016.

Entropic regularization and specific relative entropy

FRANÇOIS-XAVIER VIALARD

In this talk, I presented the joint work in [1]. The starting point of this work was the following question:

How to use the well-known entropic regularization and alternate optimization (i.e. the associated Sinkhorn-type algorithms) to solve variational problems defined on diffusion processes?

As such, this question is fairly vague. However, a well-studied framework gives sense to the entropic regularization of optimal transport in continuous time. This is the problem of Schrödinger bridge which can be formulated as a convex optimization problem on the path space $\Omega = C([0, 1], \mathbb{R}^d)$ ($\text{Ent}(\mu, \nu)$ is the Kullback-Leibler divergence, also called relative entropy),

$$(1) \quad \min_{\pi \in \mathcal{P}(\Omega)} \text{Ent}(\pi, \pi_0),$$

under the constraints that (recall that $T_{\#}(\mu)$ is the pushforward of the measure μ by T) $[\text{ev}_0]_{\#}(\pi) = \mu \in \mathcal{P}(\mathbb{R}^d)$ and $[\text{ev}_1]_{\#}(\pi) = \nu \in \mathcal{P}(\mathbb{R}^d)$ where $\text{ev}_t : \Omega \rightarrow \mathbb{R}$ is the evaluation at time t . The reference measure π_0 can be taken as the standard Wiener measure. Note this implies that a minimizer has a finite entropy with respect to this measure, which implies charging the same set of paths. In particular, one cannot hope to optimize on the set of paths that is charged by using straight Schrödinger bridges. Another formulation of Schrödinger bridges consists in solving

$$(2) \quad \inf_{\alpha(t,x)} \int_0^1 \mathbb{E}_{\mathbb{P}}[|\alpha^2(t, X_t)|^2] dt,$$

under the constraint that $dX_t = \alpha(t, X_t)dt + \bar{\sigma}dW$ where the initial distribution of X_t at time 0 is μ and its distribution at time 1 is ν . Here $\bar{\sigma}$ is a volatility that is fixed and it is not a parameter that is optimized, only the drift is.

A more general problem (written in 1d hereafter) consists of optimizing

$$(3) \quad \inf_{\alpha(t,x)} \int_0^1 \mathbb{E}_{\mathbb{P}}[F(\alpha(t, X_t)) + G(\sigma(t, X_t)^2)] dt,$$

under the constraint that $dX_t = \alpha(t, X_t)dt + \sigma(t, X_t)dW$ under similar constraints, where F, G are convex functions. Guo and Loeper and co-authors put forward such a direction, see for instance [2]. The corresponding methods' main computational issue consists in optimizing a *non-strongly* convex functional using PDE solvers, *limited to low dimensions*. To circumvent these two issues, entropic regularization comes as a natural tool. However, due to the issues mentioned above, it is not completely clear how to use it. Our method is based on the following important observation:

Each marginal at time t of a diffusion process has finite entropy with respect to the Lebesgue measure. In particular, discretizing the time and computing the relative entropy with a well-chosen

reference measure is finite. Using a proper renormalization with the number of timesteps leads to the notion of *specific relative entropy*, as introduced by N. Gantert in [3].

Let us define the specific relative entropy, when it exists for a probability measure P on the path space

$$\mathcal{S}(P|\bar{P}) := \lim_{h \searrow 0} h \operatorname{Ent}(P^h|\bar{P}^h),$$

where the P^h is the time discretization of P with timestep h and similarly for the reference measure $\bar{P}$. When P and $\bar{P}$ is the probability measure of a diffusion process in dimension d with smooth coefficients and with an elliptic condition on the diffusion coefficients, one has

$$(4) \quad \mathcal{S}(P|\bar{P}) = \mathbb{E}[\operatorname{Tr}(\bar{a}^{-1}(a - \bar{a})) - \log \frac{\det(a)}{\det \bar{a}}],$$

where $a = \sigma \sigma^\top$.

Our result is a Gamma-convergence type of result: The continuous time problem

$$(5) \quad \mathcal{I}^0(P) := \mathbb{E}_P \left(\int_0^1 F(\alpha_t^P(X_t), a_t^P(X_t)) + \mathcal{S}(a_t^P(X_t)|\bar{a}) dt \right) + \operatorname{Dist}(P_0, \rho_0) + \operatorname{Dist}(P_1, \rho_1)$$

where Dist is a divergence on the space of measures that is bounded by above by the Wassertein distance squared, can be discretized by

$$(6) \quad \mathcal{I}^h(P^h) := h \sum_{i=0}^{N-1} \mathbb{E}_{P_i^h} (F(\alpha_i^h(X_i^h), a_i^h(X_i^h))) + h \operatorname{Ent}(P^h|\bar{P}^h) + \operatorname{Dist}(P_0^h, \rho_0) + \operatorname{Dist}(P_1^h, \rho_1),$$

where the coefficients are defined by

$$(7) \quad \begin{cases} \alpha_i^h(x_i) := \frac{1}{h} \mathbb{E}_{P^h} [X_{i+1}^h - x_i], \\ a_i^h(x_i) := \frac{1}{h} \mathbb{E}_{P^h} [(X_{i+1}^h - x_i)(X_{i+1}^h - x_i)^\star]. \end{cases}$$

We need to include strong constraints on the diffusion coefficients to obtain this result to guarantee that the limit process is also a diffusion.

REFERENCES

- [1] J.-D. Benamou, G. Chazareix, M. Hoffmann, G. Loeper, and F.-X. Vialard. Entropic Semi-Martingale Optimal Transport. ArXiv preprint arXiv:2408.09361.
- [2] Ivan Guo, Gregoire Loeper, Jan Obloj, Shiyi Wang. Optimal transport for model calibration. ArXiv preprint. arXiv:2107.01978
- [3] N. Gantert. Einige große Abweichungen der Brownschen Bewegung. Universität Bonn, 1991.

Editors: M. Bauer & N. Charon