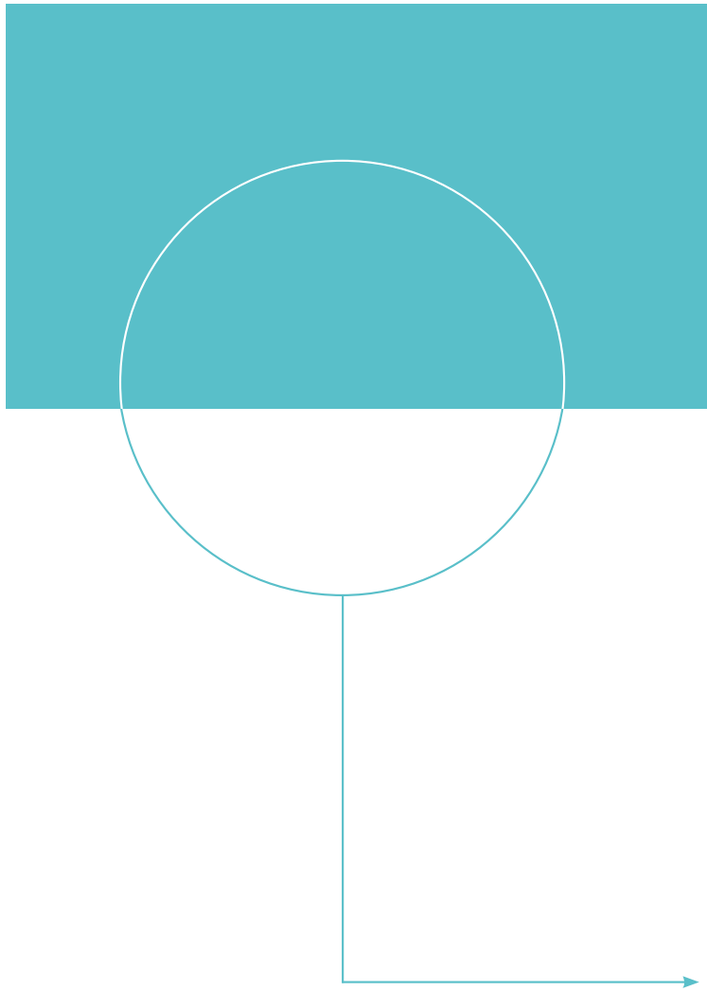


Master's thesis

2019

Pål Erik Lystad

NTNU
Norwegian University of
Science and Technology
Faculty of Information Technology and Electrical
Engineering
Department of Mathematical Sciences

Pål Erik Lystad

Signatures in Shape Analysis

An Efficient Approach to Motion Identification

June 2019

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Signatures in Shape Analysis

An Efficient Approach to Motion Identification

Pål Erik Lystad

Master of Science in Applied Physics and Mathematics
Submission date: June 2019
Supervisor: Elena Celledoni

Norwegian University of Science and Technology
Department of Mathematical Sciences

Preface

This thesis concludes a five-year integrated master's degree in Applied Physics and Mathematics at the Norwegian University of Science and Technology (NTNU), with Industrial Mathematics as the main profile. The work done with this thesis is part of a collaboration with Professor Elena Celledoni and Dr. Nikolas Tapia for the Geometric Science of Information Conference (GSI 2019).

I would like to express my sincere gratitude towards my supervisor, Professor Elena Celledoni, for our weekly meetings, beneficial discussions, as well as the opportunity to partake in seminars and conferences with the researchers in our field.

I would also like to express my gratitude to Dr. Nikolas Tapia for the effort he has put into this joint project, for the hours spent discussing, and for the help proof-reading this thesis.

Pål Erik Lystad
Trondheim, Norway
June 2019

Sammendrag

Signaturer gir en konsis beskrivelse av visse geometriske egenskaper ved en kurve - på en reparameteriseringsinvariant måte. Vi foreslår en metode for å klassifisere former basert på signaturer, og sammenligner den med nåværende metodikker basert på SRV-transformasjonen og dynamisk programmering.

Abstract

Signatures provide a succinct description of certain features of paths in a reparametrization invariant way. We propose a method for classifying shapes based on signatures, and compare it to current approaches based on the SRV transform and dynamic programming.

Table of Contents

1 Introduction

Shape Analysis

The field of shape analysis revolves around the task of determining the similarity between shapes. Initially, applications were restricted to comparing geometric shapes, typically 2D or 3D models [35] of common objects, such as cars, chairs and chess pieces. The methods have however now proved prosperous in a wide range of diverse, and in a sense more abstract, applications. These range form structural bioinformatics [14, 37] to medical diagnostics [39] and motion identification and recognition [15, 6, 2, 21] as we discuss in this thesis. The importance of the field has increased tremendously with the development of computer vision, and with data sets growing to sizes too big for humans to manually process. See [19] for a short survey of some applications of shape analysis.

Traditionally one has compared shapes by carefully selecting feature points, or landmarks, on the boundary of the shape. In recent years, however, the focus has shifted more towards curves and surfaces as these better encapsulate the geometry of the objects that arise in applications. As a result, there has now been developed significant literature concerning the analysis of shapes as elements of infinite-dimensional Riemannian manifolds [37]. This notion was first introduced by Younes [40] who worked with shape spaces of planar curves with Riemannian metrics. We will refer to Bauer, Bruveris, and Michor for an overview of the work done in this field [1].

We will consider a popular approach [37, 15, 38, 6, 7, 30], in which one models shapes as unparameterized curves, by defining a certain quotient space, for which one builds a translation and scaling invariant Sobolev metric on the space using the *Square Root Velocity Transformation* [37]. We will further model curves as having a Lie group target space to better capture the underlying geometry of the shape, as introduced by Celledoni et al [6].

Signatures

Signatures, introduced by K.-T. Chen [8] for smooth paths and later generalized by Lyons [26] under the name of geometric rough paths, are an important tool for the study of the solutions of controlled differential equations, but have also proved useful for solving classification problems of time series, Machine Learning [11] and Topological Data Analysis [12]. The theory of rough paths has gained considerable importance as a toolbox both for mathematical analysis and for mathematical modeling and simulation in practical applications.

The signature of a path can be regarded as an infinite formal tensor series with coefficients defined by iterated integrals. This map provides a unique representation of a path, capturing its essential global properties, up to some irreducibility condition. A fundamental property of the signature is its invariance under reparameterization, surmising its importance for shapes.

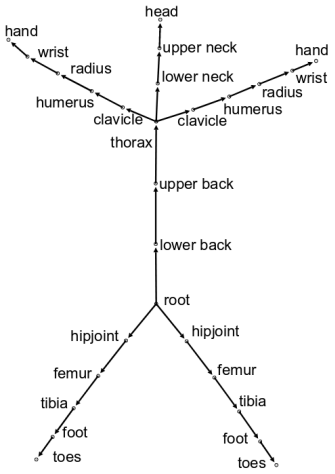


Fig. 1.1. Graph of joint angles used in the skeleton configuration used in our applications. This skeleton is based on the data taken from the CMU Graphics Lab Motion Capture Database [13]. Figure taken from [15].

Creating Skeleton Animations with Motion Capture

Creating realistic human movement remains a challenge in modern computer animation. Instead of artificially creating animations, many animators will, therefore, opt for a technique where one converts real human moment to digital animations. The most common way of doing this is called motion capture. Using multiple cameras one films a character moving in an environment. The movement is recorded by having the subject wear a special suit with easily distinguishable markers.

The placement of markers will vary depending on the application, but when capturing movement one usually uses a skeleton configuration. In a skeleton configuration, the character is modeled as a skeleton made of bones, connected by joints in a directed rooted tree. Assuming that bone length is fixed, the markers are placed on top of the joints, and the movement recorded as joint angles, the relative angle between joints, for every frame. When rendering the animation the movement of the character can be recreated by calculating the position of the character from the joints angles and the fixed bone length assumption. Textures are then applied on top of the skeleton to create a convincing animation. This technique has become popular because of its ability to capture real human movement while keeping data sizes low.

Representing the motions as joint angles will enable us to transform the motions into curves in a Riemannian manifold. This will let us utilize the before mentioned shape analysis framework. Joint angles being very geometric will let us exploit the structural similarities to Lie groups of rotations. Because human movement differs a lot from person to person, using real motion capture data will

allow us to test whether shape analysis is a valid approach for human movement identification.

In our applications we will use the skeleton configuration depicted in figure 1.1, using motion capture data from the CMU Graphics Lab Motion Capture Database [13]. The data used in this project was obtained from mocap.cs.cmu.edu. The database was created with funding from NSF EIA-0196217.

2 Overview

One of the central model assumptions in shape analysis is that shapes are unparameterized curves. This is usually done by defining shapes as equivalence classes of curves under reparameterization, that is, two curves $c_0, c_1 : [0,1] \to M$ determine the same shape if there exists a strictly increasing smooth bijection $\varphi : [0,1] \to [0,1]$ such that $c_1 = c_0 \circ \varphi$. This is an equivalence relation on curves, and for a given curve c, we denote by $[c]$ the corresponding shape.

The similarity between two shapes $[c_0], [c_1]$ is then defined by creating a distance function $d_{\mathcal{S}}$ on the space of shapes $\mathcal{S}$,

$$d_{\mathcal{S}}([c_0], [c_1]) := \inf_{\varphi} d_{\mathcal{P}}(c_0, c_1 \circ \varphi)$$

where $d_{\mathcal{P}}$ is a suitable reparameterization invariant distance on the manifold of parametrized curves.

Finding the optimal reparameterization φ can however be quite computationally demanding, and in many applications simply unnecessary. This is specifically the case of applications where the optimal parametrization is not explicitly used for further calculations, e.g. problems of identification and classification. Ways of circumventing this step are therefore of great interest.

In this thesis, we define a measure of similarity between shapes in $\mathcal{S}$ by means of the signature. We test the viability of this approach by using it to identify and classify motion capture animations obtained from the CMU motion capture database [13]. Indeed, this leads to an efficient technique that delivers results comparable to that obtainable with other more expensive methodologies.

3 Lie theory

3.1 Lie groups

In the following, G will denote a finite-dimensional Lie group, with identity element denoted by e. We let $\mathfrak{g}$ and $\mathcal{L}_R(G)$ denote the corresponding right Lie algebra of the group. For a fixed $g \in G$, left and right translation by g will be denoted $L_g(h) = g \cdot h$ and $R_g(h) = h \cdot g$ respectively.

In our applications, we will model the joints in the skeletons as a Cartesian product of Lie groups. The following result gives some important properties when describing this kind of structure.

Define component-wise multiplication by

$$(g_1, g_2) \cdot (h_1, h_2) := (g_1 h_1, g_2 h_2), \tag{3.1}$$

for $(g_1, g_2), (h_1, h_2) \in G \times G$. Then, $G \times G$ is a Lie group under component-wise multiplication, with Lie algebra $\mathfrak{g} \oplus \mathfrak{g}$.

The group structure is inherited from G by defining the identity as (e, e) and inverse map as $(g_1, g_2)^{-1} := (g_1^{-1}, g_2^{-1})$.

For a smooth manifold M, the Cartesian product $M \times M$ is a smooth manifold with

$$T_{(p,q)}(M \times M) \cong T_p(M) \oplus T_q(M)$$

for all $(p, q) \in M \times M$ [28, remark 9.20, p. 76]. We also know that the right Lie algebra[1] $\mathfrak{g}$ is isomorphic to the tangent space at the identity $T_e(G)$ [31, p.49]. From these observations, we conclude that $G \times G$ is a smooth Lie group with Lie algebra

$$\mathcal{L}_R(G \times G) \cong T_{(e,e)}(G \times G) \cong T_e(G) \oplus T_e(G) \cong \mathfrak{g} \oplus \mathfrak{g},$$

when endowed with component-wise multiplication.

3.2 Maurer–Cartan forms and the right logarithmic derivative

For a given Lie group G, its Maurer–Cartan form is the $\mathfrak{g}$-valued 1-form defined as the pushforward of right translation, that is, at a point $g \in G$ the Maurer–Cartan form is given by

$$\omega_g(X) = (R_{g^{-1}})_* X, \quad X \in T_g G.$$

If G is embedded in $GL(n)$, the group of invertible matrices, we obtain a nice expression for the Maurer–Cartan forms [31, p. 72]

$$\omega_g = \mathrm{d}g \cdot g^{-1}, \tag{3.2}$$

where the group actions are matrix multiplication and inversion.

[1] Both the left and the right lie algebras are isomorphic the the tangent space at the identity, the choice of right is one of convention.

For a parameterized curve $c \in C^\infty(I, G)$, Kriegl and Michor define the right logarithmic derivative by [22, p. 404] in terms of the Maurer–Cartan form

$$\begin{aligned} \delta^r &: C^\infty_*(I, G) \to C^\infty(I, \mathfrak{g}) \\ \delta^r c &:= (R_c^{-1})_*(\dot{c}). \end{aligned} \tag{3.3}$$

Again, we get a nice expression if G is a matrix group

$$\delta^r(c)(t) = \omega_c(\dot{c})(t) = \dot{c}(t) \cdot c(t)^{-1}. \tag{3.4}$$

4 Shape analysis on Lie groups

In this section, we present a framework for shape analysis, for curves with values in Lie groups, based on the *Square Root Velocity Transform(SRVT)*.

4.1 Shape space

We consider the space $C^\infty([0,1],G)$ of parameterized smooth curves on G, i.e. smooth maps $c:[0,1]\to G$. To model the curves as unparameterized, or independent of parameterization, we define the *shape space* $\mathcal{S}$ as the quotient space

$$\mathcal{S} = C^\infty([0,1],G)/\mathrm{Diff}^+, \tag{4.1}$$

where Diff^+ is the group of orientation preserving diffeomorphisms of the parameter space $[0,1]$. The elements of $\mathcal{S}$ are equivalence classes of curves which can be mapped to each other by changing their parameterization, that is, two curves $c_0, c_1 \in C^\infty(I,G)$ determine the same shape if there exists $\varphi \in \mathrm{Diff}^+$ such that $c_1 = c_0 \circ \varphi$. For a curve c we define the shape $[c]$ as the equivalence class containing c.

Intuitively, this projection can be viewed as syncing up the animations, removing disturbances due to small pauses, different periodicity, or asynchronous starting and stopping, by shifting the movement of one character to match the other as closely as possible.

4.2 Geodesic distances on shape space

Our goal is to introduce a meaningful and computable distance $d_\mathcal{S}$ on $\mathcal{S}$ to estimate the similarity between two shapes. We will restrict the space of curves to the space of immersions, i.e., curves with non-vanishing first derivative, which we denote by

$$\mathcal{P} = \mathrm{Imm}([0,1],G). \tag{4.2}$$

Let $d_\mathcal{P}$ be a pseudo-metric on $\mathcal{P}$. We define $d_\mathcal{S}$, for two shapes $[c_0],[c_1]\in\mathcal{S}$, by

$$d_S([c_0],[c_1]) := \inf_{\varphi\in\mathrm{Diff}^+} d_\mathcal{P}(c_0, c_1\circ\varphi). \tag{4.3}$$

As shown in [6, Lemma 3.4], d_S will be a pseudo-metric on $\mathcal{S}$ if $d_\mathcal{P}$ is a *reparameterization invariant* or, in other words, if for any two $c_0, c_1 \in \mathcal{P}$ and any $\varphi\in\mathrm{Diff}^+$ we have that

$$d_\mathcal{P}(c_0\circ\varphi, c_1\circ\varphi) = d_\mathcal{P}(c_0,c_1). \tag{4.4}$$

An obvious choice of metric on $\mathcal{P}$ is the familiar L_2-metric. However, as shown by Michor and Mumford [29], this metric leads to vanishing geodesic distance, which renders it useless. Michor and Mumford further show, in [30], that one solution to this problem is to consider metrics based on arch-length derivatives, creating a class of Sobolev-type metrics.

There are multiple possible metrics in this class. One option is based on what is usually referred to as the SRVT. This transform and accompanying metric were first introduced, in the context of shape analysis, by Srivastava et al. [37], who used the transformation when working with shapes with values in Euclidian spaces. The transformation has later been adapted to more general shapes. Of particular interest is the formulation for shapes that are represented by Lie group valued curves [6, 7].

We define the SRVT $\mathcal{R} : \mathcal{P} \to C^\infty([0,1], \mathfrak{g} \setminus \{0\})$ by

$$\mathcal{R}(c)(t) \coloneqq \frac{\delta^r(c)}{\sqrt{\|\delta^r(c)\|}}. \tag{4.5}$$

This transformation has the following useful properties [6, Lemma 3.6]:

1. For every $c \in \mathcal{P}$ and $\varphi \in \mathrm{Diff}$, the following equivariant property holds:
$$\mathcal{R}(c \circ \varphi) = \mathcal{R}(c) \circ \varphi \cdot \sqrt{\dot{\varphi}}. \tag{4.6}$$
2. It is translation invariant: for all $c \in \mathcal{P}$ and $g \in G$
$$\mathcal{R}(R_g(c)) = \mathcal{R}(c).$$

Further, one can obtain a Riemannian metric $d_{\mathcal{P}*}$ that coincides with the geodesic distance on a submanifold $\mathcal{P}_* \subset \mathcal{P}$ by using the SRVT to pull back the L_2-metric on $C^\infty(I, \mathfrak{g} \setminus \{0\})$ [6]. Further restricting the immersion space to $\mathcal{P}_* = \{c \in \mathcal{P} : c(0) = e\}$, where e is the identity element in G, the distance $d_{\mathcal{P}_*}$ turns out to be reparameterization invariant.

This invariance implies, in particular, that it will also yield a geodesic distance on $\mathcal{S}_* \coloneqq \mathcal{P}_*/\mathrm{Diff}^+$ [4]. The restriction to $\mathcal{P}_*$ isn't very troublesome as any curve can be transferred to this space by right translation by the inverse of its initial value, that is $R_{c(0)^{-1}}$ [6].

4.3 Finding the optimal reparameterization

Using the equivariant property for the SRVT from equation (4.6) and defining $q_i = \mathcal{R}(c_i)$ for $i = 0, 1$, the problem of calculating the metric for the shape space S_* in equation (4.3) can be written as

$$d_{\mathcal{S}_*}(c_0, c_1) = \inf_{\varphi \in \mathrm{Diff}^+(I)} \sqrt{\int_I \|q_0(t) - q_1(\varphi(t)) \cdot \sqrt{\dot{\varphi}}\|^2 dt}. \tag{4.7}$$

Finding this infimum will generally be very difficult. The usual approach is, therefore, to discretize the curves and instead solve a finite dimensional optimization problem. The most common methods used to solve this problem in shape analysis [37] are based on either the gradient descent method or a dynamic programming algorithm (DP). We will opt for a DP approach based on the work by Bauer et al. [2]. This approach was inspired by the work done by Klein [34].

For more information on solving this optimization problem by gradient descent, see [37, 38].

Next, we give a short outline of the algorithm described by Bauer et al. [2]. We want to determine a piece-wise linear approximation to the orientation persevering diffeomorphism ϕ that minimizes $d_{\mathcal{P}_*}$. This is done by creating a discretization of the interval $[0,1]$, and then defining a local energy functional allowing us to create a cost matrix that reduces the problem to an optimal path problem.

Let $\mathcal{I} = \{s_1, \ldots, s_M\}$ be a discretization of the interval $[0,1]$, and Φ the set of piece-wise linear maps $\varphi_{k,l;i,j}$ satisfying $\varphi_{k,l;i,j}([k,i]) = [l,j]$ where $[k,i]$ and $[l,j]$ are increasing intervals with values in the discretization $\mathcal{I}$. We define the energy functional

$$\begin{aligned} E(k,l;i,j) := &\int_k^i \|q_0(t) - q_1(\varphi_{k,l;i,j}(t))\sqrt{\dot{\varphi}_{k,l;i,j}}\|^2 dt \\ &+ \lambda \sum_{k < s_m \leq i} |\varphi_{k,l;i,j}(s_m) - s_m|^2, \end{aligned} \tag{4.8}$$

where

$$\varphi_{k,l;i,j}(t) = l + (t-k)\frac{j-l}{i-k}$$

with

$$\sqrt{\dot{\varphi}_{k,l;i,j}} = \sqrt{\frac{j-l}{i-k}},$$

and $\lambda > 0$ denotes the usual step penalty.

Let A be a $M \times M$ matrix, where M is the number of discretizations in $\mathcal{I}$. We construct A recursively by

$$A_{i,j} = \min_{k,l \in \mathcal{I}, k<i; l<j} E(k,l;i,j) + A_{k,l}, \tag{4.9}$$

while keeping track of the minimizing indices (k,l) for every pair (i,j). The optimal path through the local cost matrix A will then give us the set of linear functions $\varphi_{k,l;i,j}$ best approximating the diffeomorphism that minimizes $d_{\mathcal{P}_*}$. Again, we refer to Bauer et al. [2] for further details and proof of correctness for this algorithm.

5 Curves in $SO(3)$ and $SO(3)^d$

Since the animations in our applications are represented by joint angles, the obvious choice of target Lie group is the group of 3D rotations $SO(3)$, or $SO(3)^d$ for animations consisting of d rotating joints. In this section, we discuss some important properties which will allow us to utilize the proposed frameworks in this setting.

5.1 Properties of the groups

The Lie group $SO(3)$, the group of orthonormal 3×3 matrices, can be defined as

$$SO(3) := \left\{ A \in \mathbb{R}^{3\times 3} \text{ s.t. } \det(A) = 1 \text{ and } A^T A = I \right\},$$

where $\mathbb{R}^{3\times 3}$ is the space of real 3×3-matrices and A^T denotes the usual matrix transpose.

$SO(3)$ has Lie algebra $\mathfrak{so}(3)$ consisting of skew symmetric matrices with Lie bracket given by the usual matrix commutator $[A, B] = AB - BA$. The Lie algebra $\mathfrak{so}(3)$ is isomorphic to $\mathbb{R}^3$. We denote this isomorphism with $\wedge$, the hat-map,

$$\begin{aligned} \wedge : \mathbb{R}^3 &\to \mathfrak{so}(3), \\ \omega = (a, b, c)^T &\mapsto \hat{\omega} = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}. \end{aligned} \tag{5.1}$$

We will let $\vee$ denote the inverse hat-map. For $\tau \in \mathfrak{so}(3)$, $\check{\tau} \in \mathbb{R}^3$, notice that $\hat{\check{\tau}} = \tau$.

Under this isomorphism, the Lie bracket will correspond to the cross product in $\mathbb{R}^3$.

We will compute the exponential map from $\mathfrak{so}(3)$ to $SO(3)$ using what is commonly referred to as the Euler-Rodrigues formula [16]

$$\begin{aligned} \exp : \mathfrak{so}(3) &\to SO(3) \\ \exp(\hat{\omega}) = I + \frac{\sin\theta}{\theta}\hat{\omega} + \frac{(1-\cos\theta)}{\theta^2}\hat{\omega}^2, &\quad \theta = \|\hat{\omega}\|_F, \end{aligned} \tag{5.2}$$

where $\|\hat{\omega}\|_F$ is the Frobenius norm, see equation (5.4).

For a rotation $X \in SO(3)$, it is well known [16] that we can retrieve the angle of rotation θ by

$$\theta = \arccos\left(\frac{\operatorname{tr}(X) - 1}{2}\right).$$

This angle is necessary to compute the inverse of the exponential [16, 36], that is $\log : SO(3) \to \mathfrak{so}(3)$, given by

$$\log(X) = \omega = \frac{\theta}{2\sin\theta}(X_{32} - X_{23}, X_{13} - X_{31}, X_{21} - X_{12})^T. \tag{5.3}$$

where X_{ij} denote row i, column j of the matrix X. If we instead wish to map the rotation X to the skew matrices of $\mathfrak{so}(3)$, the formula can be written in a more compact form by applying the inverse hat-map to the vector

$$\log(X) = \hat{\omega} = \frac{\theta}{2\sin\theta}(X - X^T).$$

For $\hat{\omega} \in \mathfrak{so}(3)$, we use the familiar Frobenius norm, scaled by a factor $1/2$, when calculating norms,

$$\|\hat{\omega}\|^2 = \frac{1}{2}\|\hat{\omega}\|_F^2 = \frac{1}{2}\mathrm{tr}(\hat{\omega} \cdot \hat{\omega}^T), \tag{5.4}$$

as this corresponds to the 2-norm of the vector representation $\omega \in \mathbb{R}^3$,

$$\begin{aligned}
\|\hat{\omega}\|^2 = \frac{1}{2}\|\hat{\omega}\|_F^2 &= \frac{1}{2}\mathrm{tr}\left(\begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix}\right) \\
&= \frac{1}{2}\mathrm{tr}\left(\begin{bmatrix} c^2+b^2 & -ba & -ca \\ -ba & c^2+a^2 & -cb \\ -ca & -cb & a^2+b^2 \end{bmatrix}\right) \\
&= \frac{1}{2}(2a^2 + 2b^2 + 2c^2) = \|\omega\|_2^2.
\end{aligned}$$

Next, we consider a target space with d rotating joints, represented by the Lie group $SO(3)^d$ which we define as

$$SO(3)^d := \underbrace{SO(3) \times SO(3) \times \cdots \times SO(3)}_{d\text{-times}}.$$

We showed, in section 3, that this is a Lie group with corresponding Lie algebra

$$\mathfrak{so}(3)^{\mathbf{d}} := \underbrace{\mathfrak{so}(3) \oplus \mathfrak{so}(3) \oplus \cdots \oplus \mathfrak{so}(3)}_{d\text{-times}}$$

when endowed with component-wise multiplication.

The Lie algebra consists of skew symmetric matrices, each isomorphic to $\mathbb{R}^3$. This allows us to extend the hat-map by mapping each element $v_k \to \hat{v}_k$, $k = 1 \ldots, d$ with the hat-map from (5.1)

$$\begin{aligned}
\mathbb{R}^3 \oplus \cdots \oplus \mathbb{R}^3 &\to \mathfrak{so}(3) \oplus \cdots \oplus \mathfrak{so}(3) \\
\omega = (v_1, \ldots, v_d) &\mapsto \hat{\omega} = (\hat{v}_1, \ldots, \hat{v}_d).
\end{aligned}$$

Since $\mathbb{R}^3 \oplus \ldots \oplus \mathbb{R}^3$ is isomorphic to $\mathbb{R}^{3d}$, by stacking the vectors

$$(v_1, \ldots, v_d) \mapsto \begin{bmatrix} v_1 \\ \vdots \\ v_d \end{bmatrix},$$

we observe that the hat-map will induce an isomorphism

$$\mathfrak{so}(3)^{\mathbf{d}} \cong \mathbb{R}^{3d}. \tag{5.5}$$

Mappings between Lie group and Lie algebra can be defined by applying the log and exp maps for $SO(3)$, equations (5.2) and (5.3), joint-wise. Let $\hat{\omega} = (\hat{\omega}_1, \ldots, \hat{\omega}_d) \in \mathfrak{so}(3)^{\mathbf{d}}$, the exponential map: $\exp : \mathfrak{so}(3)^{\mathbf{d}} \to SO(3)^d$ is

$$\exp(\hat{\omega}) := (\exp(\hat{\omega}_1), \ldots, \exp(\hat{\omega}_d)),$$

and similarly the logarithm: $\log : SO(3)^d \to \mathfrak{so}(3)^{\mathbf{d}}$, for $x = (X_1, \ldots, X_d) \in SO(3)^d$

$$\log(x) := (\log(X_1), \ldots, \log(X_d)).$$

Note that the angle of rotation is distinct for every joint.

We construct a norm by summing up the norm of the components in the Lie algebra. Let $\hat{\omega} = (\hat{\omega}_1, \ldots, \hat{\omega}_d) \in \mathfrak{so}(3)^{\mathbf{d}}$ be an element in the Lie algebra. The map

$$\|\hat{\omega}\|^2 := \frac{1}{2} \sum_{i=1}^{d} \|\hat{\omega}_i\|_F^2,$$

is a norm in $SO(3)^d$, which is also equivalent to the 2-norm for elements $\mathbb{R}^{3d}$. The first statement, that this is in fact a norm, follows from linearity of the Frobenius norm, and that the elements in the sum by definition are strictly positive which means the map can only be zero if all the elements are zero. Let $\omega = ((a_1, b_1, c_1)^T, \ldots, (a_d, b_d, c_d)^T) \in \mathfrak{so}(3)^{\mathbf{d}}$, it becomes apparent that this map is equivalent to the 2-norm of R^{3d} by writing out the expression

$$\|\hat{\omega}\|^2 = \sum_{i=0}^{d} \frac{1}{2} \|\hat{\omega}_i\|_F^2 = \sum_{i=1}^{d} \|\omega_i\|_2^2 = \sum_{i=1}^{d} (a_i^2 + b_i^2 + c_i^2) = \|\omega\|_2^2.$$

5.2 Geodesic Interpolation

As we saw in the last section, constructing the joint space as a cartesian product of rotations in $SO(3)$ allowed us to extend the theory by simply applying the transformations joint-wise. Similarly, the methods described in this subsection for curves in $SO(3)$ can easily be extended to $SO(3)^d$ utilizing the same joint-wise approach.

As mention in the introduction, the animations consist of discrete joint angles for every frame. To approximate the curve c, from which the joint angles are sampled, we wish to create an interpolation $\overline{c}$.

One common interpolation scheme [36] is to map the rotations to the Lie algebra, using the previously mentioned log map, linearly interpolate in the vector space, and then map back from the Lie algebra using the exponential. Let $s \in [0, 1]$, we call $\overline{c}$ the geodesic interpolation between sample points $c_0, c_1 \in SO(3)$

$$\overline{c}(s) = \exp\left(s \log(c_1 c_0^T)\right) c_0. \tag{5.6}$$

As noted by Shingel [36], the term "geodesic interpolation" is appropriate in this setting as the interpolation will coincide with the geodesic connecting the points $c_0, c_1 \in SO(3)$.

The interpolation scheme can be extended to data sets of n joint angles. Let s_k be the time stamp for sampling point c_k. The sampling times will be an increasing sequence of the form

$$s_0 = 0 < s_1 < s_2 < \ldots < s_{n-1} < s_n = 1. \tag{5.7}$$

Let $t \in I$, we approximate the original curve c with the interpolation $\overline{c}$

$$\overline{c}(t) = \sum_{k=0}^{n-1} \mathbb{1}_{[s_k, s_{k+1})}(t) \exp\left(\frac{t - s_k}{s_{k+1} - s_k} \log(c_{k+1} c_k^T)\right) c_k, \tag{5.8}$$

where for a set A, $\mathbb{1}_A(t)$ is the indicator function defined as

$$\mathbb{1}_A(t) = \begin{cases} 1 \text{ if } t \in A \\ 0 \text{ if } t \notin A \end{cases}.$$

Reparameterization We will also use this interpolation scheme to reparameterize curves. Let $r_1 = 0 < r_2 < \ldots < r_{n-1} < r_n = 1$ a new set of sampling time stamps. Given a curve $\overline{c}$ with sampling points c_k and sampling times s_k for $k = 1, \ldots, n$, the idea is to find which sub-interval the new sampling point belongs to by enforcing $s_k \leq r_i < s_{k+1}$, and then interpolating to find the new point d_i

$$d_i = \exp\left(\frac{r_i - s_k}{s_{k+1} - s_k} \log(c_{k+1} c_k^T)\right) c_k \quad i = 1, \ldots, n. \tag{5.9}$$

d_i together with r_k can then be used to create a new interpolation curve $\overline{d}$.

Notice that the criteria $s_1 = 0, s_n = 1$ together with $r_1 = 0, r_n = 1$ imply that the start and end points will remain fixed, that is $c_1 = d_1$ and $c_n = d_n$.

Right logarithmic derivative and SRV transformation Next we compute the right logarithmic derivative, which is exactly the Maurer–Cartan form, for the interpolation curves. Let

$$\eta_k = \frac{\log\left(c_{k+1} c_k^T\right)}{s_{k+1} - s_k}, \tag{5.10}$$

we denote by κ the interpolation curve $\overline{c}$ restricted to the time segment $[s_k, s_{k+1})$ for some $k \in 1, \ldots, (n-1)$. Then, κ can be written as

$$\kappa(t) = \exp((t - s_k)\eta_k) c_k$$

Since the interpolation curves take values in a matrix group, $SO(3)$, we can utilize the explicit formula for the right logarithmic derivative (3.4)

$$\begin{aligned}\delta^r(\kappa)(t) &= \dot{\kappa}\cdot\kappa^T = \eta_k \exp((t-s_k)\eta_k)c_k \cdot (\exp((t-s_k)\eta_k)c_k)^T \\ &= \eta_k \exp((t-s_k)\eta_k)\underbrace{c_k c_k^T}_{=I}\exp(-(t-s_k)\eta_k) \\ &= \eta_k \exp((t-s_k)\eta_k - (t-s_k)\eta_k) = \eta_k \exp(0\cdot\eta_k) \\ &= \eta_k\end{aligned}$$

Notice that the right logarithmic derivative is constant for the curve segment κ. Inserting into (4.5) yields the SRV transformation of κ

$$q_k = \frac{\delta^r(\kappa))}{\sqrt{\|\dot{\kappa}(t)\|}} = \frac{\eta_k}{\sqrt{\|\eta_k\|}}. \tag{5.11}$$

The right logarithmic derivative $\delta^r(\overline{c})$ and SRVT $\mathcal{R}(\overline{c})$ of the interpolation curve $\overline{c}$ will therefore be piece-wise constant and will respectively take the form

$$\mathcal{R}(\overline{c})(t) := \overline{q}(t) = \sum_{k=1}^{n-1} \mathbb{1}_{[s_k,s_{k+1})}(t)q_k. \tag{5.12}$$

$$\delta^r(\overline{c})(t) := \overline{\eta}(t) = \sum_{k=1}^{n-1} \mathbb{1}_{[s_k,s_{k+1})}(t)\eta_k. \tag{5.13}$$

According to [6], the SRVT will have a well defined and smooth inverse $\mathcal{R}^{-1} : C(I,\mathfrak{g}\backslash 0) \to \mathcal{P}_*$. Using the formula [6, Lemma 3.9] will let us reconstruct the sampling point $c_{k+1} = \overline{c}(s_{k+1})$ from a SRV transformed curve $\overline{q} = \mathcal{R}(\overline{c})$, satisfying $\overline{q}(s_k) = q_k$, from the recursion $c_1 = I$ and

$$c_{k+1} = \mathcal{R}^{-1}(\overline{q})|_{s_k} = \exp\left((s_{k+1}-s_k)\|q_k\|q_k\right)c_k \text{ for } k = 1,\dots,(n-1).$$

Inserting the SRV transformed curve segment q_k from equation (5.11) yields a straight forward computation to confirm that this will indeed reconstruct the sampling point c_{k+1}

$$\begin{aligned}\exp\left((s_{k+1}-s_k)\|q_k\|q_k\right)c_k &= \exp\left((s_{k+1}-s_k)\left\|\frac{\eta_k}{\sqrt{\|\eta_k\|}}\right\|\frac{\eta_k}{\sqrt{\|\eta_k\|}}\right)c_k \\ &= \exp\left((s_{k+1}-s_k)\frac{\|\eta_k\|}{\|\eta_k\|}\eta_k\right)c_k \\ &= \exp\left((s_{k+1}-s_k)\frac{\log(c_{k+1}c_k^T)}{(s_{k+1}-s_k)}\right)c_k \\ &= \exp\big(\big(\log(c_{k+1}c_k^T)\big)c_k \\ &= c_{k+1}.\end{aligned}$$

Having an inverse of the SRVT makes it possible to interpolate between two curves $c_a, c_b \in \mathcal{P}_*$. Let $s \in [0,1]$, the interpolation scheme

$$c_{\text{int}} = \mathcal{R}^{-1}(s\mathcal{R}(c_a) + (1-s)\mathcal{R}(c_b)) \tag{5.14}$$

will be well defined due to the vector space structure of the Lie algebra [6, Prep 3.15]. Note that this is the geodesic with respect to the elastic metric between the points c_a and c_b in the infinite dimensional Riemannian manifold $\mathcal{P}_*$ [6].

6 Identifying movement using the square root velocity framework

In this section, we attempt to identify different types of movement by using the proposed metric (4.7) as a similarity measure for the shapes of animations.

6.1 Motion capture data

As mentioned, we use animation data from the CMU Motion Capture Database [13]. The data in the motion capture database is separated into `XX.asf/XX_YY.amc`-files. These describe subject `XX` and animation `YY`, respectively. The subject file describes which joints are connected and the length of the bones between them, while the animation file contains the rotation of every joint in every frame. The animation files can be parsed and transformed to discrete sample points in $SO(3)^d$, with accompanying sample time stamps $s_1, ..., s_n \in [0, 1]$. The animations are recorded with cameras capable of a frame rate of 120hz. The high frame rate will make the geodesic interpolation scheme suggested in equation (5.8) a good approximation of the recorded movement. This is because a high frame rate implies short time steps between samples. The interpolation, which in practice is piece-wise linear, will then accurately approximate curvature of the movement. We will let $\overline{c}$ denote the geodesic interpolation connection the sample points in $SO(3)^d$.

6.2 Computing the metric

For shapes of animations $[\overline{c}_0], [\overline{c}_1] \in \mathcal{S}_*$, computing the metric $d_{\mathcal{S}_*}([\overline{c}_0], [\overline{c}_1])$ is done in four steps:

1. Compute SRVT representation of the curves $\overline{c}_0$ and $\overline{c}_1$.
2. Find the optimal reparameterization φ^{opt} satisfying (4.7) with the DP algorithm from section 4.3.
3. Reparameterize the curve $\overline{c}_1$ with the reparameterization φ^{opt} using the approach described in equation (5.9) to create a new curve $\overline{c}_1^{\text{opt}}$.
4. Approximate $d_{\mathcal{S}_*}([\overline{c}_0], [\overline{c}_1]) \approx d_{\mathcal{P}_*}(\overline{c}_0, \overline{c}_1^{\text{opt}})$.

From an array of sampling points in $SO(3)^d$ we calculate the SRV representation of the geodesic interpolation using (5.11). This will yield an array of SRV values $q_1, ..., q_k$, where q_k determines the values of the SRVT representation on the interval $[s_k, s_{k+1})$. Note that since the SRV representations $\overline{q}_0 = \mathcal{R}(\overline{c}_0)$ and $\overline{q}_1 = \mathcal{R}(\overline{c}_1)$ are piece-wise constant, computing the L_2-metric in step 4 is trivial

$$\int_0^1 \|\overline{q}_0 - \overline{q}_1\|^2 dt = \sum_{k=1}^{n-1} \|\overline{q}_0(s_k) - \overline{q}_1(s_k)\|^2 \cdot (s_{k+1} - s_k).$$

An important implementation detail to note when computing this integral is that the curves might have different parameterizations. This could either stem

from the new parameterization gained from the optimization step, or from the underlying animations having a different number of frames. Either way, this can be overcome by creating a shared set of sampling timestamps, i.e. $s_k \in I$, and then filling in the missing sampling points for each curve by interpolating to the previous value in the array.

Given a reparameterization φ^{opt} we create a new set of sample timestamps $r_1, .., r_n$. This set is used to reparameterize the curve c_1 as described in equation (5.9), completing step 3.

Finding the optimal reparameterzation Based on the DP algorithm outlined in section 4.3, we use the following approach to find the optimal set of sampling time stamps I_1^{opt} corresponding to φ^{opt} in step 2. Again let, $\overline{c}_0$ and $\overline{c}_1$ denote geodesic interpolations with sampling time stamps I_0, I_1 and SRVT representatives $\overline{q}_0, \overline{q}_1$.

We create a shared discretization I by merge joining the sampling time stamps I_0 and I_1, while removing any duplicate values. Next, we create the local cost matrix, while keeping track of the minimizing indices, that is the indices (k, l) minimizing the energy functional in equation (4.9). To find the optimal reparameterization φ^{opt} we backtrack through the minimizing indices, starting at (M, M), where M is the number of points in the discretization I. This yields a mapping $k \hookrightarrow l$ of optimal reparameterizations for $\forall k \in I$. Using this mapping, we create the piece-wise linear map φ^{opt}, approximating the optimal diffeomorphism, by linearly interpolating such that $\varphi^{\text{opt}}(k) = l$ for $\forall k \in I$. Lastly, we map the time stamps in the original parameterization I_1 to the new, optimal, parameterization I_1^{opt} which we utilize in step 3 to create the optimal curve c_1^{opt}.

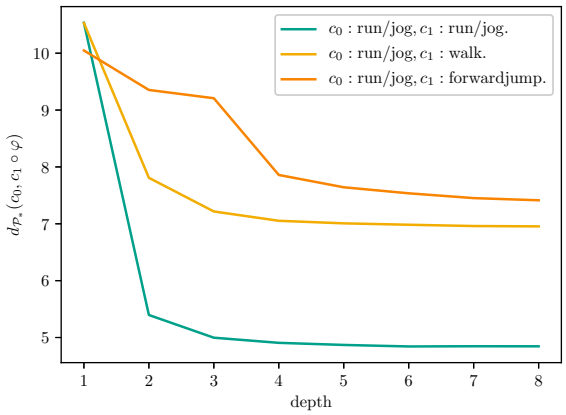


Fig. 6.1. Observed $d_{\mathcal{S}_*}([c_0], [c_1]) = d_{\mathcal{P}_*}(c_0, c_1 \circ \varphi^{\text{opt}})$, for animations $c_0, c_1 \in \mathcal{P}_*$, with minimizing reparameterization φ^{opt} calculated with a DP algorithm for different search depth parameters D.

Calculating the L_2-metric when making the local cost matrix will be very computationally demanding. For a discretization consisting of M points, where we consider N predecessors, the algorithm will have an asymptotic run time of $\mathcal{O}(M^2N)$. Bauer et al. [2] suggest restricting the number of predecessors to indices (k, l) close to (i, j) to speed up the computation. Let the depth D denote how far away from the (i, j) we search, that is we only allow predecessors (k, l) such that $(i - k) \leq D$ and $(j - l) \leq D$.

Next we test how the DP algorithm behaves for different depths D. In figure 6.1 we plot the measured similarity $d_{\mathcal{P}_*}(c_0, c_1 \circ \varphi^{opt})$ for the minimizing reparameterization φ^{opt} found for varying depth constraints.

For this small selection of animations, we observe that optimal reparameterization seems to be within a depth $D < 6$. The sample also shows nice convergence properties, both in terms of which animations the method deems most similar ("run/jog" and "run/jog") and in terms of stability.

6.3 Visualizing similarity measures

An effective way of testing whether a similarity measure produces reasonable results is to test if it is able to classify animations. To do this, we calculate the similarity between every pair of animations and then collect the result in a similarity matrix. For the SRVT framework, the similarity used is the geodesic distance between shapes of animations, this could, however, be any similarity measure. Note that the similarity matrix will be symmetric with zeros along the diagonal.

There are a few different methods for visualizing the contents of a similarity matrix. One is called multi dimensional scaling. This method creates a scatter plot that preserves the distance between the nodes, meaning that if one was to create a similarity matrix from the plot using a Euclidian distance measure one would get the same similarity matrix. The advantage of this method is that it quickly lets the reader get an intuitive understanding of the dissimilarities in a data set. Similar data points will be close, while dissimilar data will be apart. The downside is that it is difficult to read out exact data from the plot. See the book by Kruskal for more information on this visualization technique [23].

We will also use dendrograms to visualize the similarity matrix. A dendrogram plots the hierarchical distance between nodes in a tree. These are less intuitive, but have the advantage that they expose the underlying data to the reader.

6.4 Identification of basic movement

We would like to test whether the proposed framework is effective at identifying different types of motion. The motions should be similar in length, and there should be enough trials to yield statistically significant results. The selection of animations should also reflect that similarity is a fundamentally subjective type of measure. That is, the animations should be of a type such that it is easy to determine whether the results conform with a human notion of similarity.

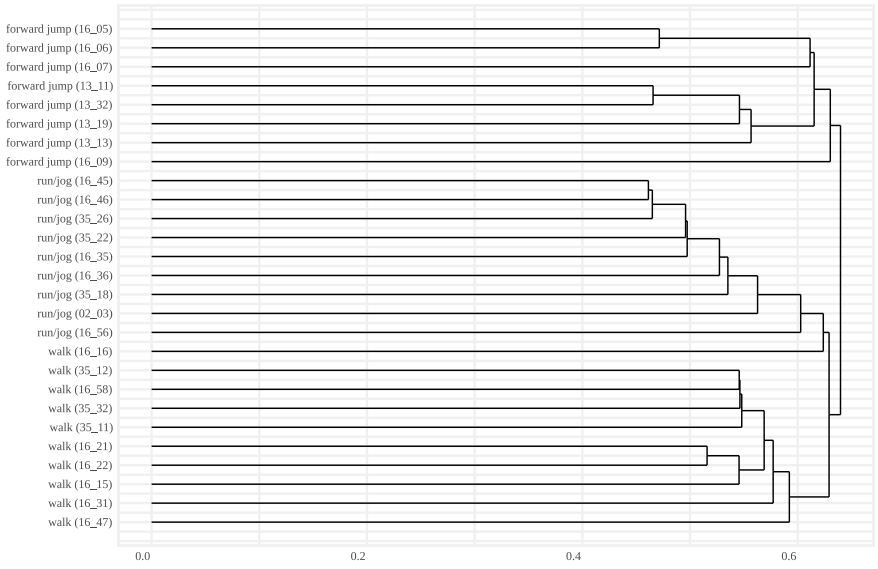


Fig. 6.2. Dendrogram of a similarity matrix for animations projected to the space $\mathcal{S}_*$ using the metric $d_{\mathcal{S}_*}(4.7)$ calculated with a DP algorithm with search depth $D = 9$. In this plot we have taken animation with descriptions "run/jog", "forward jump" and "walk" from the CMU Motion Capture Database [13].

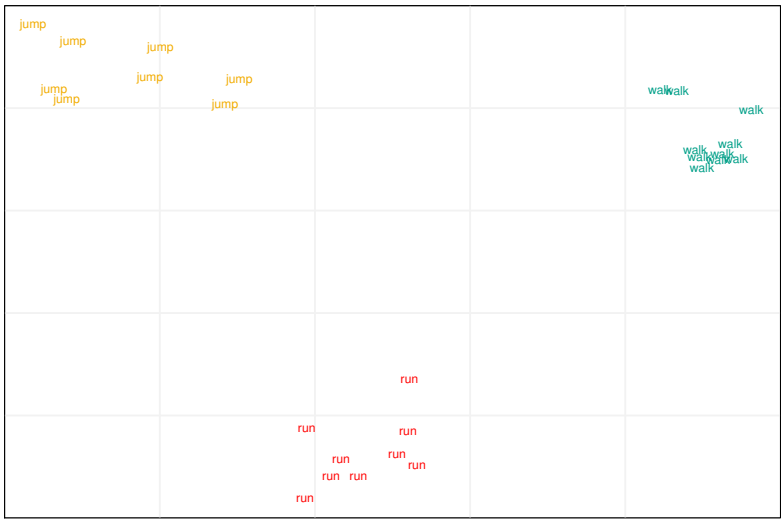
The animations best fitting of these criteria in the CMU motion capture database [13] are the animations with descriptions "walk", "run/jog", and "forward jump". These are both numerous and similar in length, ranging from approximately 130 to 420 frames. They should also produce results coinciding with human intuition. "forward jump"-animations consists of characters jumping once from a standstill. There are only 8 of these so they will limit the number of trails in this experiment. Some of the animations had long pauses in the beginning and/or end where the character just stood still, seemingly waiting for instructions, these have been roughly cropped to an appropriate length. Other than that the animations are unchanged from how they appear in the database.

To reiterate, we create a similarity matrix by calculating the geodesic distance between animations projected to the metric space $(\mathcal{S}_*, d_{\mathcal{S}_*})$. From this similarity matrix, we produce multidimensional scaling scatter plots and dendrograms depicting the similarities between animations. We also test the impact of modeling the animations as unparameterized by repeating the same experiment for animations projected to the non-parameterization invariant space $(\mathcal{P}_*, d_{\mathcal{P}_*})$.

In figure 6.3 we have plotted multi dimensional scaling plots of the similarity matrix created by projecting 27 animations with description "walk", "run/jog" and "forward jump" to the two spaces. Figure 6.3a shows animations projected to $(\mathcal{P}_*, d_{\mathcal{P}_*})$, while figure 6.3b shows animations projected to $(\mathcal{S}_*, d_{\mathcal{S}_*})$. In figure 6.3b we used a search depth $D = 9$ as this was well beyond the observed limit of convergence for the sample set in figure 6.1.

Looking at figure 6.3b, we observe that modeling the animations as parameterization invariant yields three easily distinguishable clusters of animations. The dendrogram of the similarity matrix for the animations projected to $(\mathcal{S}_*, d_{\mathcal{S}_*})$ in figure 6.2 confirms that the methods successfully identify the three types of movement. Compared to figure 6.3a, we see a big benefit from this model assumption.

(a) Animations projected to $\mathcal{P}_*$ with similarity matrix calculated with the metric $d_{\mathcal{P}_*}$.



(b) Animations projected to $\mathcal{S}_*$ with similarity matrix calculated with metric $d_{\mathcal{S}_*}$ using a DP algorithm with search depth $D = 9$.

Fig. 6.3. Multi dimensional scaling plots of similarity matrix based on geodesic distances calculated in $\mathcal{P}_*$ and $\mathcal{S}_*$, figure (a) and (b) respectively. In this plot we have taken animation with descriptions "run/jog", "forward jump" and "walk" from the CMU Motion Capture Database [13].

7 Signatures

In this section, we introduce the signature representation of a smooth path. This representation is interesting in the context of shape analysis as its uniqueness characteristics are closely related to the equivalence classes of shapes. Before we justify this claim, we have to define the mapping and discuss some of its most important properties.

As mentioned in the introduction, this representation can be seen as a formal tensor series, with coefficients determined by iterated integrals. For a finite-dimensional vector space V of dimension $d = \dim V$, we define the tensor algebra over V as

$$T(V) \coloneqq \bigoplus_{n \geq 0} V^{\otimes n} = \mathbb{R} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots .$$

We observe that this space is infinite-dimensional. Its dual space is denoted by $T(\!(V)\!) \coloneqq T(V)^*$, and it may be identified with the ring of formal power series in d non-commuting variables $\{e_1, \ldots, e_d\}$.

For a d dimensional vector space, we may think of these tensor series as infinite vectors indexed by *words* over the alphabet $\{1, \ldots, d\}$. In this context, the word $w = i_1 \cdots i_n$ where $i_1, \ldots, i_n \in \{1, \ldots d\}$, correspond to the basis element

$$e_w \coloneqq e_{i_1} \otimes \cdots \otimes e_{i_n}.$$

Let $x : [0, 1] \to \mathbb{R}^d$ be a smooth path. We adopt the notation commonly used in stochastic analysis where $x_t \coloneqq x(t)$ and superscript i denotes the i-th component of the path $x_t = (x_t^1, ..., x_t^d)^T$. Now, for the word $w = i_1 \cdots i_n$, we define the *n-fold iterated integral* over the interval $[s, t] \subset [0, 1]$ as

$$\langle S(x)_{s,t}, e_w \rangle = \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} dx_{u_1}^{i_1} \cdots dx_{u_{n-1}}^{i_{n-1}} dx_{u_n}^{i_n}.$$

These integrals will describe different geometric properties of the path x. If we, for example, look at the one-letter word $w = i$, the integral

$$\langle S(x)_{0,1}, e_i \rangle = \int_0^1 dx_u^i = x_1^i - x_0^i$$

describes the increment of the i-th component. For two-letter words, the integral can be interpreted as a signed area generated by the path. We will refer to [18] for a more in-depth discussion of these interpretations.

Given a smooth path $x : [0, 1] \to \mathbb{R}^d$, we define its *signature* over the interval $[s, t] \subset [0, 1]$ as the formal tensor series

$$S(x)_{s,t} = 1 + \sum_{|w| \geq 1} \langle S(x)_{s,t}^w, e_w \rangle e_w \in T(\!(\mathbb{R}^d)\!). \tag{7.1}$$

Note that letting $\langle S(x)_{s,t}^w, e_w \rangle$ denote the n-fold integral with respect to the word w is consistent with the duality pairing of the tensor series $S(x)_{s,t}$ with

the basis element e_w: Let u, v denote two words. Then $\langle e_u, e_v \rangle = 1$ if $u = v$ and $\langle e_u, e_v \rangle = 0$ otherwise. The duality pairing between the tensor series and the basis element e_w is therefore zero for every element except for the element corresponding to the n-fold integral over the word w.

In practise one has to truncate the signature to obtain a finite dimensional object. The usual approach is to truncate it to include all the elements corresponding to words of length less than or equal to some length n. We call this the *signature truncated to degree* n, and denote it with a superscript (n)

$$S(x)_{s,t}^{(n)} = 1 + \sum_{1 \leq |w| \leq n} \langle S(x)_{s,t}, e_w \rangle e_w.$$

For a path x, with values in R^d, the number of elements M in the signature truncated to degree n is given by the formula [33]

$$M = \frac{n(n^d - 1)}{n - 1}.$$

Due to the factorial decay of iterated integrals little information is lost when truncating the signature. Still, some level has to be chosen and usually this is determined by running experiments.

7.1 Signatures of linear paths

Signatures may be efficiently computed for some restricted classes of paths. For example, if x is a straight line with base point $b \in \mathbb{R}^d$ direction $a \in \mathbb{R}^d$, i.e. $x_t = at + b$ for $t \in [0, 1]$, then the n-fold iterated integral of the word $w = i_1 \cdots i_n$ equals

$$\begin{aligned} \langle S(x)_{s,t}, e_w \rangle &= \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} a_{i_1} du_1 \cdots a_{i_{n-1}} du_{n-1} a_{i_n} du_n \\ &= \prod_{k=1}^{n} a_{i_k} \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} du_1 \cdots du_{n-1} du_n \\ &= \frac{(t-s)^n}{n!} \prod_{k=1}^{n} a_{i_k}. \end{aligned} \tag{7.2}$$

Using this, we observe that the signature of a linear path can be expressed as the exponential of a certain vector in $\mathbb{R}^d$

$$\begin{aligned} S(x)_{s,t} &= 1 + \sum_{|w| \geq 1} \frac{(t-s)^n}{n!} \prod_{k=1}^{n} a_{i_k} e_w \\ &= 1 + (t-s)a + \frac{(t-s)^2}{2} a \otimes a + \frac{(t-s)^3}{6} a \otimes a \otimes a + \cdots \\ &= \exp_{\otimes}((t-s)a). \end{aligned} \tag{7.3}$$

One important property of the signature is what is usually referred to as **Chen's rule** [8]: For any three $0 \leq s < u < t \leq 1$ we have

$$S(x)_{s,u} \otimes S(x)_{u,t} = S(x)_{s,t}. \tag{7.4}$$

For a general piece-wise linear path x, we can use Chen's rule and the signature for a linear path (7.3) to deduce that

$$S(x)_{s,t} = \exp_{\otimes}(\Delta t_1 a_1) \otimes \exp_{\otimes}(\Delta t_2 a_2) \otimes \cdots \otimes \exp_{\otimes}(\Delta t_m a_m)$$

where $\Delta t_k = t_k - t_{k-1}$ are the length of the time intervals where the path is sampled and $a_1, \ldots, a_m \in \mathbb{R}^d$ are the slopes of the path in each of these intervals. The entries of this expression can be efficiently computed using a Baker–Campbell–Hausdorff-type formula [5], for example.

7.2 Paths with values in Lie groups

Chen originally proposed the signature for paths with values in a differentiable manifold [8]. This definition is quite general and relies on the selection of a frame bundle. For Lie groups, there is a canonical choice: the Maurer–Cartan form.

The signature will again be a tensor series, now in the dual tensor algebra $T(\!(\mathfrak{g})\!)$. For a smooth curve $\alpha : [0,1] \to G$, we define the n-fold iterated integral recursively, in terms of the Maurer–Cartan form ω, by $\langle S(\alpha)_{s,t}, 1\rangle \coloneqq 1$ and

$$\langle S(\alpha)_{s,t}, e_{i_1 \cdots i_n}\rangle \coloneqq \int_s^t \langle S(\alpha)_{s,u}, e_{i_1 \cdots i_{n-1}}\rangle \, \omega^{i_n}_{\alpha(u)}(\dot{\alpha}(u)) \, \mathrm{d}u.$$

In this definition, the notation $\omega^j_g(v)$ denotes the j-th component of the vector $\omega_g(v) \in \mathfrak{g}$ in a basis of the Lie algebra $\mathfrak{g}$ of the finite-dimensional Lie group G.

The signature defined with these integrals has the same characteristics and uniqueness conditions as the signature of paths with values in Euclidean space.

Paths in $SO(3)$ As with paths with values in Euclidean space, there are some classes of paths which are especially simple to compute. An important example in this class is the geodesic interpolation between rotations in $SO(3)$. Note that for Lie groups embedded in $GL(n)$, the right Maurer–Cartan form will be the same as the right logarithmic derivative.

Let $\kappa : [s_k, s_{k+1}] \to SO(3)$ be the geodesic interpolation (5.8) between the rotations $c_k, c_{k+1} \in SO(3)$. Define

$$\hat{\eta}_k = \frac{\log\left(c_{k+1} c_k^T\right)}{s_{k+1} - s_k},$$

from equation (5.13) the right logarithmic derivative of the path κ is constant on the interval $[s_k, s_{k+1}]$

$$\omega_\kappa(\dot{\kappa})(t) = \delta^r(\kappa)(t) = \hat{\eta}_k.$$

Using the inverse hat map, we map the skew matrix $\hat{\eta}_k$ to the vector $\eta_k = (\eta_k^1, \eta_k^2, \eta_k^3) \in \mathbb{R}^3$. Since the derivative is constant in time, the n-fold integrals closely resemble that of linear paths:

$$\langle S(\alpha)_{s_k,s_{k+1}}, e_{i_1 \cdots i_n} \rangle = \frac{(s_{k+1} - s_k)^n}{n!} \prod_{j=1}^{n} \eta_k^{i_j}$$

where $i_1, \ldots, i_n \in \{1, 2, 3\}$ and $p \geq 1$. Intuitively this resemblance makes sense as both path types can be regarded as being first order approximations. One would expect a similar result to be true in general for geodesic interpolations in finite-dimensional compact, connected Lie groups, we will however not investigate this any further.

Writing out the signature with the n-fold integrals above would again reveal that the signature can be regarded as the exponential of a vector in $\mathbb{R}^d$

$$S(\alpha)_{s_k,s_{k+1}} = \exp_{\otimes}((s_{k+1} - s_k)\eta_k)).$$

As with linear paths, extending this result to a piece-wise geodesic interpolation with multiple sample points can be done by applying Chen's rule to every piece. One can also extend the results discussed to paths with values in $SO(3)^d$. The main difference being that because $\mathfrak{so}(3)^d$ is isomorphic to $\mathbb{R}^{3d}$, the letters $i, \ldots, i_n$ take values in the alphabet $\{1, 2, \ldots, 3d\}$.

7.3 The space of signatures

According to Lyons, the signatures of paths take values in a curved subspace of the tensor algebra [24]. This subspace is closed under tensor multiplication, with a well-defined inverse and identity element. Since the subspace satisfies the group axioms it is common to refer to signatures as the group-like elements of the tensor algebra.

For a path $x : [0, 1] \to \mathbb{R}^d$, the group structure will have identity element $S(x)_{s,s} = 1$, and inverse given by the *reversal operator* $S_{s,t}^{-1}(x) = S_{s,t}(\overleftarrow{x})$, where the reversed path is defined as

$$\overleftarrow{x}(t) = x(1 - t). \tag{7.5}$$

That this is in fact an inverse becomes apparent from what is usually referred to as **Chen's Identity** [8]:

$$S(x) \otimes S(\overleftarrow{x}) = 1. \tag{7.6}$$

For two paths, $x, y : [0, 1] \to \mathbb{R}^d$, we define concatenation[2] $*$, by

$$(x * y)(t) = \begin{cases} x(2t) & t \leq \frac{1}{2} \\ x(1) - y(0) + y(2t - 1) & t > \frac{1}{2} \end{cases}. \tag{7.7}$$

[2] Note that due to the reparameterization invariance the mid point of the concatenation can be chosen to be any real number in (0,1).

Using Chen's rule we can deduce an interesting property, namely that the signature is a homomorphism from path space with concatenation to the dual tensor algebra [10]. This means that if we are given two paths $x, y : [0,1] \to \mathbb{R}^d$, and we concatenate them to form a new path $x * y$, then

$$S(x * y)_{s,t} = S(x)_{s,t} \otimes S(y)_{s,t}.$$

Using this property we can also give an alternative formulation of Chen's Identity:

$$S(x * \overleftarrow{x})_{0,1} = 1. \tag{7.8}$$

In a Lie group setting, it is natural to define this notion in terms of the group actions. For paths $u, v : [0,1] \to G$, we therefore define concatenation $*$, by

$$(u * v)(t) = \begin{cases} u(2t) & t \leq \frac{1}{2} \\ v(2t-1) \cdot v(0)^{-1} \cdot u(1) & t > \frac{1}{2} \end{cases}. \tag{7.9}$$

The previously discussed homomorphism property also holds in this setting [3], that is for paths $u, v : [0,1] \to G$, the signature is a homomorphism under concatenation

$$S(u * v)_{s,t} = S(u)_{s,t} \otimes S(v)_{s,t}.$$

7.4 Uniqueness for the signature representation of a path

The signature is a *unique* representation of a path, up to some condition. Chen proved that for smooth paths [10], this condition states that the signature is unique up to translation, parameterization, and irreducibility[3].

The invariance to translation is because the iterated integrals are of the form $dx_t = \frac{dx}{dt}dt$, that is translation is lost when differentiating. The invariance to parameterization is because the signature is *reparameterization invariant*; for any orientation-preserving diffeomorphism φ on $[s,t]$ we have that

$$S(x \circ \varphi)_{s,t} = S(x)_{s,t}. \tag{7.10}$$

The proof of the reparameterization invariance follows from computing the change of variable $t \to t \circ \varphi$ for the iterated integrals.

The irreducibility condition can be viewed as a consequence of Chen's Identity. Consider the three paths $x, y, z : [0,1] \to \mathbb{R}^d$, and consider the concatenation

$$x * y * \overleftarrow{y} * z.$$

[3] For rough paths, Hambly and Lyons [20] showed that this condition is what they call 'tree-like equivalence'. This work is one of the important contributions that reignited the interest in signatures as a mathematical tool.

The signature of this concatenated path can be reduced using the homomorphism property and Chen's identity

$$\begin{aligned} S(x * y * \overleftarrow{y} * z)_{s,t} &= S(x)_{s,t} \otimes \underbrace{S(y)_{s,t} \otimes S(\overleftarrow{y})_{s,t}}_{=1} \otimes S(z)_{s,t} \\ &= S(x)_{s,t} \otimes S(z)_{s,t} \\ &= S(x * z)_{s,t}. \end{aligned}$$

That is, the path $x * y * \overleftarrow{y} * z$ will have the same signature as the reduced path $x * z$, even though that paths are vastly different. We say that a path is *irreducible* if it cannot be expressed on the form $x * y * \overleftarrow{y} * z$, and that the signature is a unique representation of an irreducible path, up to translation and parameterization.

7.5 Log signatures

Since the space of signatures can be given a group structure, an interesting question is whether it also has a corresponding "Lie algebra" and if there exists a logarithmic mapping to this algebra. Chen [9] showed that such a map exists and that the image is a closed subspace of the tensor algebra which, according to Lyons and Sidorova, can be "regarded as some sort of formal Lie algebra for this group." [27]. The algebra has Lie bracket [24] given by

$$[p, q] = p \otimes q - q \otimes p.$$

To define the logarithm we take the same approach as in [11]: Consider the power series

$$p = \lambda_0 + \sum_{k=1}^{\infty} \sum_{\substack{|w|=k \\ w=i_1 \cdots i_k}} \lambda_w e_{i_1} \otimes \cdots \otimes e_{i_k} \in T((\mathbb{R}^d)),$$

where $\lambda_0 > 0$. We can define the logarithm of the power series as

$$\log p = \log \lambda_0 + \sum_{k=0}^{\infty} \frac{(-1)^k}{k} \left(1 - \frac{p}{\lambda_0}\right)^{\otimes k}.$$

As discussed in [11], this transformation is well-defined because for every word $w = i_1 \cdots i_k$ the coefficient of $e_1 \otimes \cdots \otimes e_k$ only depend on finitely many elements of the power series p.

Let $x : [0, 1] \to \mathbb{R}^d$ be a path, we define the log signature of x as the formal power series $\log S_{s,t}(x)$. Since the logarithm is injective [20], the log signature will be unique representation of a path up to the same conditions as the signature of the path.

An advantage of working with the space of log signatures is that, in contrast to the usual space of signatures, it is flat [24]. This will make computing distances easy as one can rely on the usual Euclidean measures.

Another problem with regular signatures is that they contain a lot of *redundancy*. To motivate what this means we give a quick example: Consider the n-fold integral corresponding to the word w

$$\langle S(x)_{s,t}, e_w \rangle.$$

For words only made up of one letter, $w \in \{i, ii, iii, ...\}$, the signature will only depend on the increment of the path x

$$\begin{aligned}
\langle S(x)_{s,t}, e_i \rangle &= x_t^i - x_s^i \\
\langle S(x)_{s,t}, e_{ii} \rangle &= \frac{(x_t^i - x_s^i)^2}{2!} \\
\langle S(x)_{s,t}, e_{iii} \rangle &= \frac{(x_t^i - x_s^i)^3}{3!} \\
&\vdots
\end{aligned}$$

These are all just different powers of the increment, and give no additional information about the geometry of the path.

To understand the full extent of this redundancy we need to look at at what is known as the *shuffle product* over signatures. First introduced by Ree [32], the shuffle product states how the product of two signature terms can be written as a sum of higher order signature terms. The reason for the name is that this sum will be exactly the multi-indices generated by the (p, q)-shuffles of the words corresponding to the product terms. We define the set of (p, q)-shuffles of the words $i = i_1 \cdots i_p$ and $j = j_1 \cdots j_q$: For words i, j, the (p, q)-shuffle $\mathrm{Sh}(p, q)$ is the set of all permutations σ of the word

$$k := k_1 \cdots k_{p+q} = i_1 \cdots i_p j_1 \cdots j_q$$

that satisfy the following conditions

$$\sigma(1) < \cdots < \sigma(p) \text{ and } \sigma(p+1) < \cdots < \sigma(p+q),$$

that is, preserve the order of the letters in the words. The product of the signature terms corresponding to the words i and j is equal to the sum over all possible (p, q)-shuffles of the combined word k

$$\langle S(x)_{s,t}, e_{i_1 \cdots i_p} \rangle \cdot \langle S(x)_{s,t}, e_{j_1 \cdots j_q} \rangle = \sum_{\sigma \in \mathrm{Sh}(p,q)} \langle S(x)_{s,t}, e_{k_{\sigma(1)} \cdots k_{\sigma(p+q)}} \rangle. \quad (7.11)$$

Using this formula we can generate more intuitive examples of redundancy in the signature. For example, the term corresponding to the word ij

$$\langle S(x)_{s,t}, e_{ij} \rangle = \langle S(x)_{s,t}, e_{ji} \rangle - \langle S(x)_{s,t}, e_i \rangle \cdot \langle S(x)_{s,t}, e_j \rangle,$$

really only contain the same geometric information as the reversed 'area' ji and the increments i and j.

The log signatures do however not possess this redundancy - it is the most compact description of the geometric properties of the signature of a path. This can, according to Lyons and Sidorova [27], be deduced from the classical Rashevski-Chow Theorem.

8 Using signatures to define similarity measures for shapes

Consider a shape $[c] \in \mathcal{S}_*$, because of the translation and parameterization invariance, every representative of the equivalence class $[c]$ will have the same signature. It, therefore, becomes very natural to define the signature of a shape $[c] \in \mathcal{S}_*$ as

$$S([c]) := S(c),$$

where c is an arbitrary representative of the shape $[c]$. It is, however, important to note that due to the irreducibility condition, the signature of a shape is not a unique representation.

Parameterization invariance proved to be the most important model assumption when identifying character movements in section 6. Since the signature is a faithful representation of the geometric properties of a shape - in a parameterization invariant way - it would be interesting to investigate whether similarity measures defined using signatures yield any additional insight into the structure of the shape space. In this section, we will, therefore, present some commonly used metrics defined on the subspace of signatures.

As far as the author can tell, considering signatures in the context of shape analysis has not been given substantial attention. One of the more notable examples, however, comes from Geng [17] who worked with reconstructing a representative of a shape from its signature.

8.1 Group metric

One of the most commonly used metrics in the study of rough paths is what we will refer to as the "group metric" on signatures. This metric was introduced by T. Lyons and N. Victoir when studying geometric rough paths [25]. They proved that the metric is well-defined on the curved structure of the subspace of signatures.

By utilizing the group structure of the space, we know that reversing a path yields a group-like inverse in the tensor algebra of signatures. A natural measure of similarity arises from using this property to induce a metric. Let $x, y : [0,1] \to \mathbb{R}^d$, we define the *group metric* ρ^n as

$$\rho^n(x,y) = \left\| S(y)_{0,1}^{-1} \otimes S(x)_{0,1} \right\|_{G^n}, \tag{8.1}$$

where $\|\cdot\|_{G^n}$ is the norm on the elements in the group-like tensor algebra which Lyons and Victoir [25] define as

$$\|g\|_{G^n} = \max_{i=1,..,n} \left(i! \sqrt{\sum_{|w|=i} \langle g, e_w \rangle^2} \right)^{1/i} + \max_{i=1,..,n} \left(i! \sqrt{\sum_{|w|=i} \langle g^{-1}, e_w \rangle^2} \right)^{1/i}. \tag{8.2}$$

To give some verification that this is in fact a metric we note the following:

1. Since the signature has a unique inverse and $\|g\|_{G^n} = 0$ if and only if $g = 1$, then $\rho(x, y) = 0$ if and only if $S(x)_{s,t} = S(y)_{s,t}$.
2. Because
$$\left(S(y)_{0,1}^{-1} \otimes S(x)_{0,1}\right)^{-1} = S(x)_{0,1}^{-1} \otimes S(y)_{0,1},$$
the second term in the norm, where we maximize over $\langle g^{-1}, e_w \rangle$, ensures that the metric is symmetric: $\rho(x, y) = \rho(y, x)$.

Using the homomorphism property of the signature map, this metric has an equivalent formulation by concatenating the curves

$$\rho^n(x, y) = \left\| S(\overleftarrow{y} * x)_{0,1} \right\|_{G^n}. \tag{8.3}$$

Equivalently, we define the distance function for the shapes generated by the paths x, y by

$$\rho^n([x], [y]) = \rho^n(x, y).$$

8.2 Convergence for the group metric for linear paths

In applications, it is common to consider signatures of linear and piece-wise linear paths as these are efficient to implement and give a good description of the data. It would, therefore, be revealing to examine how the group metric behaves when we perturb linear paths. Define the paths $x, x_\epsilon : [0, 1] \to \mathbb{R}^2$ by

$$\begin{aligned} x(t) &:= at \\ x_\epsilon(t) &:= (a + \epsilon \mathbf{1})t, \end{aligned}$$

where $a = (a_1, a_2)$ and ϵ is some arbitrary constant. Clearly $\lim_{\epsilon \to 0} x_\epsilon = x$, and since ρ^n is a metric this implies $\lim_{\epsilon \to 0} \rho^n(x_\epsilon, x) = 0$. To better understand the metric we will calculate the order of convergence for the limit.

We start by calculating the signatures for the two paths. Since both paths are linear these are trivial to calculate using equation (7.3)

$$\begin{aligned} S(x)_{0,1} &= 1 + a_1 e_1 + a_2 e_2 + \frac{a_1^2}{2} e_1^{\otimes 2} + \frac{a_2^2}{2} e_2^{\otimes 2} + \frac{a_1 a_2}{2} e_1 \otimes e_2 + \frac{a_2 a_1}{2} e_2 \otimes e_1 + ... \\ S(x_\epsilon)_{0,1} &= 1 + (a_1 + \epsilon) e_1 + (a_2 + \epsilon) e_2 + \frac{(a_1 + \epsilon)^2}{2} e_1^{\otimes 2} + \frac{(a_2 + \epsilon)^2}{2} e_2^{\otimes 2} \\ &\quad + \frac{(a_1 + \epsilon)(a_2 + \epsilon)}{2} e_1 \otimes e_2 + \frac{(a_2 + \epsilon)(a_1 + \epsilon)}{2} e_2 \otimes e_1 + ..., \end{aligned}$$

as well as the inverse of the signature of x

$$S(x)_{0,1}^{-1} = 1 - a_1 e_1 - a_2 e_2 + \frac{a_1^2}{2} e_1^{\otimes 2} + \frac{a_2^2}{2} e_2^{\otimes 2} + \frac{a_1 a_2}{2} e_2 \otimes e_1 + \frac{a_2 a_1}{2} e_1 \otimes e_2 - .... \tag{8.4}$$

To get a sense of how the norm $\|\cdot\|_{G^n}$, and by extension the metric ρ^n, behave we calculate n-fold integral

$$\langle S(x)_{0,1}^{-1} \otimes S_{0,1}(x_\epsilon), e_w \rangle$$

for words w of length $n \geq 1$. Starting with $n = 1$, that is $w = i_1$, where $i_1 \in \{1, 2\}$, the duality pairing is equal to the increment

$$\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_{i_1} \rangle = -a_{i_1} + (a_{i_1} + \epsilon) = \epsilon$$

There are only two words of length one, $i_1 = 1$ and $i_1 = 2$, so calculating the group norm gives

$$\left\| S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1} \right\|_{G^1} = \sqrt{\epsilon^2 + \epsilon^2} + \sqrt{(-\epsilon)^2 + (-\epsilon)^2} = 2\sqrt{2}\epsilon$$

That is,

$$\left\| S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1} \right\|_{G^1} = \mathcal{O}(\epsilon).$$

To examine level two signatures we look at the word $w = i_1 i_2$,

$$\begin{aligned}
&\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_{i_1 i_2} \rangle \\
&= \frac{a_{i_2} a_{i_1}}{2} - a_{i_1}(a_{i_2} + \epsilon) + \frac{(a_{i_2} + \epsilon)(a_{i_1} + \epsilon)}{2} \\
&= \left(\frac{a_{i_2} a_{i_1}}{2} + \frac{a_{i_1} a_{i_2}}{2} - a_{i_2} a_{i_1} \right) + \left(-a_{i_1} + \frac{a_{i_2} + a_{i_1}}{2} \right) \epsilon + \frac{\epsilon^2}{2} \\
&= \frac{\epsilon}{2}(-a_{i_1} + a_{i_2}) + \frac{\epsilon^2}{2}
\end{aligned}$$

we observe that for off-diagonal words, that is words in which $i_1 \neq i_2$, the integral has a non-zero ϵ-term of order one. If we assume $\epsilon << 1$ and $a_1 \neq a_2$, then this term will dominate the convergence

$$\left\| S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1} \right\|_{G^2} = \mathcal{O}(\epsilon^{1/2}).$$

Note that for the special case $a_1 = a_2$, the first order term vanishes, making the metric linearly proportional to ϵ.

Finally, we look at words of length three $w = i_1 i_2 i_3$, the elements in the group norm will then be elements on the form

$$\begin{aligned}
&\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_{i_1 i_2 i_3} \rangle \\
&= -\frac{a_{i_3} a_{i_2} a_{i_1}}{6} + \frac{a_{i_2} a_{i_1}(a_{i_3} + \epsilon)}{2} - \frac{a_{i_1}(a_{i_2} + \epsilon)(a_{i_3} + \epsilon)}{2} + \frac{(a_{i_1} + \epsilon)(a_{i_2} + \epsilon)(a_{i_3} + \epsilon)}{6} \\
&= \left(\frac{a_{i_1} a_{i_2} a_{i_3}}{6} - \frac{a_{i_3} a_{i_2} a_{i_1}}{6} + \frac{a_{i_2} a_{i_1} a_{i_3}}{2} - \frac{a_{i_1} a_{i_2} a_{i_3}}{2} \right) \\
&+ \epsilon \left(\frac{a_{i_2} a_{i_1}}{2} - \frac{a_{i_1} a_{i_2} + a_{i_1} a_{i_3}}{2} + \frac{a_{i_1} a_{i_3} + a_{i_2} a_{i_3} + a_{i_1} a_{i_2}}{6} \right) \\
&+ \epsilon^2 \left(-\frac{a_{i_1}}{2} + \frac{a_{i_3} + a_{i_1} + a_{i_2}}{6} \right) \\
&+ \frac{\epsilon^3}{6} \\
&= \frac{\epsilon}{6}(a_{i_1} a_{i_2} + a_{i_2} a_{i_3} - 2 a_{i_1} a_{i_3}) + \frac{\epsilon^2}{6}(-2 a_{i_1} + a_{i_2} + a_{i_3}) + \frac{\epsilon^3}{6}
\end{aligned}$$

Again, we observe that off-diagonal words will yield non-zero coefficients for the first order term in the ϵ polynomial, which implies

$$\left\|S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}\right\|_{G^3} = \mathcal{O}(\epsilon^{1/3}$$

for sufficiently small ϵ, if we assume $a_1 \neq a_2$. Again, for $a_1 = a_2$, the lower order terms cancel, giving

$$\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_{i_1 i_2 i_3}\rangle = \frac{\epsilon^3}{6}$$

From these examples there seem to be a pattern where off-diagonal terms dominate the convergence when $a_1 \neq a_2$. This could be useful for use cases where we are trying to determine an exact match. For noisy data, however, this type of convergence behaviour is unwanted as it would amplify small, and muffle big, differences. In animations, where the technique and execution of a movement usually vary quite a bit between subjects/animations, this characteristic could make the group metric unsuitable for identifying motions.

Proposition 1. *Consider linear paths $x, x_\epsilon : [0,1] \to \mathbb{R}^d$ of the form*

$$x(t) := at + b \quad x_\epsilon(t) := (a + \epsilon\mathbf{1})t + b_\epsilon,$$

where $a, b, b_\epsilon \in \mathbb{R}^d$ and ϵ is some arbitrary constant. Let a_i denote the i-th component of the vector a. Then, for sufficiently small ϵ, the convergence of the limit $\lim_{\epsilon\to 0} \rho^n(x, x_\epsilon)$ will have order of convergence given by

$$\rho^n(x, x_\epsilon) = \begin{cases} \mathcal{O}(\epsilon) & \text{if } a_i = a_j \forall i, j \in \{1, .., d\}, \\ \mathcal{O}(\epsilon^{1/n}) & \text{otherwise} \end{cases}.$$

Proof. Consider the word $w = i_1 \cdots i_n$ of length n, the element of the tensor product corresponding to this word can be written as a sum

$$\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_{i_1 \dots i_n}\rangle = \sum_{w_1 w_2 = i_1 .. i_n} \langle S(x)_{0,1}^{-1}, e_{w_1}\rangle \cdot \langle S(x_\epsilon)_{0,1}, e_{w_2}\rangle,$$

where the sum runs over all ways to split the word $i_1 \cdots i_n$ in the two parts w_1 and w_2 such that $w_1 = i_1 \cdots i_k$ and $w_2 = i_{k+1} \cdots i_n$ for $0 \leq k \leq n$. Inserting the expression for the n-fold integral corresponding to a length n word for a linear paths (7.2) gives

$$\begin{aligned} &\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_{i_1 \dots i_n}\rangle \\ &= \sum_{\substack{w_1 w_2 = i_1 .. i_n \\ |w_1| = k}} \left(\frac{(-1)^k}{k!} \prod_{j=1}^{k} a_{i_j} \cdot \frac{1}{(n-k)!} \prod_{j=k+1}^{n} (a_{i_j} + \epsilon) \right) \\ &= \sum_{\substack{w_1 w_2 = i_1 .. i_n \\ |w_1| = k}} \left(\frac{(-1)^k}{k!(n-k)!} \prod_{j=1}^{k} a_{i_j} \cdot \prod_{j=k+1}^{n} (a_{i_j} + \epsilon) \right). \end{aligned}$$

Like we saw for words of length one, two, and three and paths in $\mathbb{R}^2$, this will again take the form of polynomials of ϵ. The second product can be written to reflect the polynomial structure. Let

$$\mathcal{C}_k^n(r) := \binom{\{i_{k+1}, ..., i_n\}}{r},$$

be the set of all possible ways to choose r elements from the set $\{i_{k+1}, ..., i_n\}$ without repetition. We can write the product as a polynomial in ϵ

$$\prod_{j=k+1}^{n} (a_{i_j} + \epsilon) = \sum_{l=0}^{n-k} \epsilon^l \left(\sum_{\mathcal{I} \in \mathcal{C}_k^n(n-k-l)} \prod_{i \in \mathcal{I}} a_i \right).$$

Inserting this into the duality pairing we get

$$\begin{aligned}
&\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_{i_1 \dots i_n} \rangle \\
&= \sum_{\substack{w_1 w_2 = i_1 .. i_n \\ |w_1| = k}} \left(\frac{(-1)^k}{k!(n-k)!} \prod_{j=1}^{k} a_{i_j} \cdot \sum_{l=0}^{n-k} \epsilon^l \left(\sum_{\mathcal{I} \in \mathcal{C}_k^n(n-k-l)} \prod_{i \in \mathcal{I}} a_i \right) \right).
\end{aligned}$$

This way of writing will make it easier to isolate the order of the terms, making it possible to check which terms contribute to the convergence. We start by calculating the zero order term of the polynomial:

$$\begin{aligned}
\sum_{\text{zero order}} &= \frac{(-1)^n}{n!} \prod_{j=1}^{n} a_{i_j} + \sum_{k=0}^{n-1} \left(\frac{(-1)^k}{k!(n-k)!} \prod_{j=1}^{k} a_{i_j} \cdot \epsilon^0 \left(\sum_{\mathcal{I} \in \mathcal{C}_k^n(n)} \prod_{i \in \mathcal{I}} a_i \right) \right) \\
&= \frac{(-1)^n}{n!} \prod_{j=1}^{n} a_{i_j} + \sum_{k=0}^{n-1} \left(\frac{(-1)^k}{k!(n-k)!} \prod_{j=1}^{k} a_{i_j} \cdot \prod_{j=k+1}^{n} a_{i_j} \right) \\
&= \frac{(-1)^n}{n!} \prod_{j=1}^{n} a_{i_j} + \sum_{k=0}^{n-1} \left(\frac{(-1)^k}{k!(n-k)!} \prod_{j=1}^{n} a_{i_j} \right) \\
&= \sum_{k=0}^{n} \left(\frac{(-1)^k}{k!(n-k)!} \prod_{j=1}^{n} a_{i_j} \right) \\
&= \prod_{j=1}^{n} a_{i_j} \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!(n-k)!},
\end{aligned}$$

which can be simplified using the binomial theorem

$$\begin{aligned}
\sum_{\text{zero order}} &= \frac{1}{n!} \prod_{j=1}^{n} a_{i_j} \cdot \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} (-1)^k \\
&= \frac{1}{n!} \prod_{j=1}^{n} a_{i_j} \cdot \sum_{k=0}^{n} \binom{n}{k} (-1)^k (1)^{n-k} = \frac{1}{n!} \prod_{j=1}^{n} a_{i_j} \cdot (1-1)^n = 0.
\end{aligned}$$

As expected the zero order term is zero. This could also be deduced from the fact that ρ^n is a metric, but will be left in for completeness. Now to the first order term:

$$\begin{aligned}\sum_{\text{first order}} &= \sum_{k=0}^{n-1}\left(\frac{(-1)^k}{k!(n-k)!}\prod_{j=1}^{k}a_{i_j}\cdot\epsilon^1\left(\sum_{\mathcal{I}\in\mathcal{C}_k^n(n-1)}\prod_{i\in\mathcal{I}}a_i\right)\right)\\ &= \epsilon^1\sum_{k=0}^{n-1}\left(\frac{(-1)^k}{k!(n-k)!}\prod_{j=1}^{k}a_{i_j}\cdot\left(\sum_{j=k+1}^{n}\prod_{\substack{l=k+1\\ l\neq j}}^{n}a_{i_l}\right)\right)\end{aligned}$$

For a general word, it is difficult to calculate the value of this sum, its close resemblance to a binomial sum might, however, give an indication that it is usually non-zero.

Assume that there exists a a pair of indices (i,j) such that

$$a_i \neq a_j,$$

This implies that either

1. $a_i \neq 0$ and $a_j \neq 0$, or
2. $a_i \neq 0$ and $a_j = 0$.

We start with the latter case. In terms of paths, this could be considered the special case where the path only has one non-zero slope component. Define the word

$$w = \underbrace{ii\ldots i}_{n-1}j.$$

For $a_i \neq 0$ and $a_j = 0$, the first order component of the innter product equals:

$$\begin{aligned}\sum_{\text{first order}} &= \epsilon^1\sum_{k=0}^{n-1}\left(\frac{(-1)^k}{k!(n-k)!}\prod_{j=1}^{k}a_{i_j}\cdot\left(\sum_{j=k+1}^{n}\prod_{\substack{l=k+1\\ l\neq j}}^{n}a_{i_l}\right)\right)\\ &= \epsilon^1\sum_{k=0}^{n-1}\left(\frac{(-1)^k}{k!(n-k)!}a_i^k\cdot a_i^{n-k-1}\right)\\ &= \epsilon^1\frac{a_i^{n-1}}{n!}\sum_{k=0}^{n-1}\left(\frac{n!}{k!(n-k)!}(-1)^k(1)^{n-k}\right)\\ &= \epsilon^1\frac{(-1)^{n-1}a_i^{n-1}}{n!}\end{aligned}$$

where the last step, again, is a consequence of the binomial theorem. Since $a_i \neq 0$ the first order term is non-zero for $\epsilon > 0$. We now look at the other case,

where $a_i \neq a_j$ and both a_i, a_j are non-zero:

$$\begin{aligned}
\sum_{\text{first order}} &= \epsilon^1 \sum_{k=0}^{n-1} \left(\frac{(-1)^k}{k!(n-k)!} \prod_{j=1}^{k} a_{i_j} \cdot \left(\sum_{j=k+1}^{n} \frac{a_{i_j}}{a_{i_j}} \prod_{\substack{l=k+1 \\ l \neq j}}^{n} a_{i_l} \right) \right) \\
&= \epsilon^1 \sum_{k=0}^{n-1} \left(\frac{(-1)^k}{k!(n-k)!} \prod_{j=1}^{k} a_{i_j} \cdot \left(\sum_{j=k+1}^{n} \frac{1}{a_{i_j}} \prod_{l=k+1}^{n} a_{i_l} \right) \right) \\
&= \epsilon^1 \prod_{j=1}^{n} a_{i_j} \sum_{k=0}^{n-1} \left(\frac{(-1)^k}{k!(n-k)!} \cdot \left(\sum_{j=k+1}^{n} \frac{1}{a_{i_j}} \right) \right) \\
&= \epsilon^1 a_i^{n-1} a_j \sum_{k=0}^{n-1} \left(\frac{(-1)^k}{k!(n-k)!} \cdot \left(\frac{n-k-1}{a_i} + \frac{1}{a_j} \right) \right) \\
&= \epsilon^1 a_i^{n-1} a_j \left(\sum_{k=0}^{n-1} \frac{(-1)^k}{k!(n-k)!} \cdot \frac{n-k-1}{a_i} + \sum_{k=0}^{n-1} \frac{(-1)^k}{k!(n-k)!} \frac{1}{a_j} \right) \\
&= \epsilon^1 a_i^{n-1} a_j \left(\frac{(-1)^n}{n!} \frac{1}{a_i} + \frac{(-1)^{n+1}}{n!} \frac{1}{a_j} \right) \\
&= \epsilon^1 \frac{(-1)^n}{n!} a_i^{n-2} (a_j - a_i)
\end{aligned}$$

which is non-zero.

That is, if we assume that there exists a pair of indices (i, j) such that $a_i \neq a_j$, then there exists a word w such that the element of the tensor product corresponding to that word has a non-zero first-order element when written as a polynomial in ϵ

$$\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_w \rangle = C\epsilon + \mathcal{O}(\epsilon^2).$$

For sufficiently small ϵ this term will dominate the convergence of the group norm

$$\left\| S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1} \right\|_{G^n} = \mathcal{O}(\epsilon^{1/n}),$$

and by extension the group metric

$$\rho^n(x, x_\epsilon) = \mathcal{O}(\epsilon^{1/n}),$$

since it is induced from the group norm.

For the other case, where $a_i = a_j$ for all $i, j \in \{1, ..., d\}$. We observed that the lower order terms cancelled for words of length $n \leq 3$ for linear paths in $\mathbb{R}^2$. This is also true for the general case. Let $a := a_i$, the duality pairing can be

simplified using the binomial theorem

$$\begin{aligned}
\langle S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1}, e_w \rangle &= \sum_{\substack{w_1 w_2 = i_1..i_n \\ |w_1|=k \\ a_0=a_1}} \left(\frac{(-1)^k}{k!} \prod_{j=1}^{k} a \cdot \frac{1}{(n-k)!} \prod_{j=k+1}^{n} (a+\epsilon) \right) \\
&= \sum_{k=0}^{n} \frac{(-1)^k}{k!(n-k)!} a^k (a+\epsilon)^{n-k} \\
&= \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} (-1)^k a^k (a+\epsilon)^{n-k} \\
&= \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (-a)^k (a+\epsilon)^{n-k} \\
&= \frac{1}{n!} (-a+a+\epsilon)^n \\
&= \frac{\epsilon^n}{n!}.
\end{aligned}$$

Which implies

$$\left\| S(x)_{0,1}^{-1} \otimes S(x_\epsilon)_{0,1} \right\|_{G^n} = \mathcal{O}(\epsilon),$$

when $a_i = a_j$ for all $i, j \in \{1, ..., d\}$. □

Numerical experiments To verify the result of proposition 1 we check the order of convergence by numerically calculating the value of the metric $\rho^n(x, x_\epsilon)$. Let $x, x_\epsilon : [0, 1] \in \mathbb{R}^4$ be linear paths with slope

$$a = (0.1, 0.5, 0.0, -0.2)^T.$$

In applications one usually truncates the signature to degree three or four, we have therefore restricted the experiments to $n \in [1, 8]$. In figure 8.2 we have plotted the metric $\rho^n(x, x_\epsilon)$ as a functions of ϵ. Figures 8.2a–8.2b show log-log plots for $n \in [1, 4]$ and $n \in [5, 8]$ respectively. From these we see that the metric converges we the expected $\epsilon^{1/n}$-rate.

In figure 8.1 we have repeated the experiment for paths where the slope is equal in all directions: $a = (0.1, 0.1, 0.1, 0.1)$. The metrics converge with the expected linear rate in this case. Looking closely, the metric seems to be numerically unstable for small ϵ for the higher order metrics. This can probably be attributed to machine precision as $\epsilon \approx 0.001$ will mean values of order $0.001^8 = 10^{-24}$ when calculating the n-fold integrals for the norm.

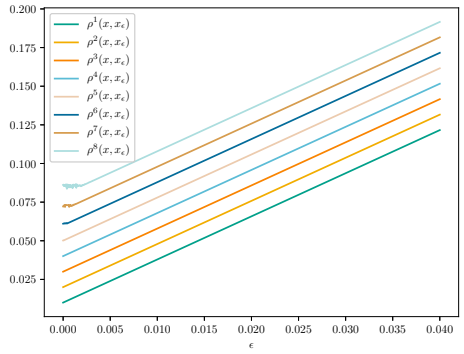
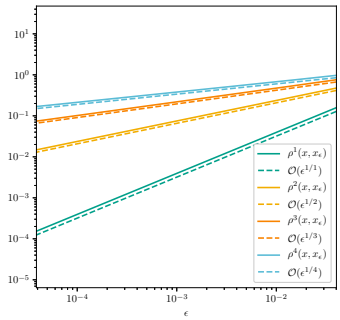
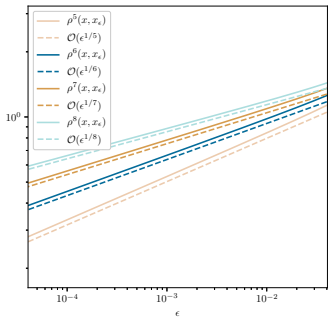


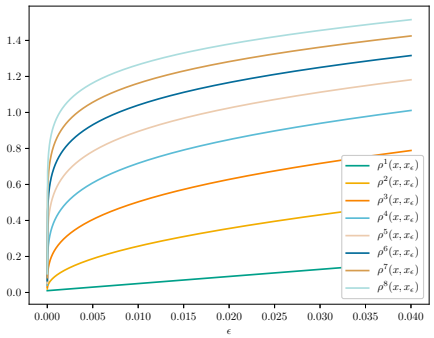
Fig. 8.1. Convergence experiment for the group metric ρ^n for linear paths in $\mathbb{R}^2$ with $a_1 = a_2$. The lines have been shifted by a small constant to make it possible to discern the different metrics.



(a) Log-log plot of the group metric $\rho^n(x, x_\epsilon)$ for varying ϵ, and $n \in [1, 4]$.



(b) Log-log plot of the group metric $\rho^n(x, x_\epsilon)$ for varying ϵ, and $n \in [5, 8]$.



(c) Plot of the group metric $\rho^n(x, x_\epsilon)$ for varying ϵ, and $n \in [1, 8]$.

Fig. 8.2. Convergence experiment for the group metric ρ^n for linear paths in $\mathbb{R}^4$ with $a_1 \neq a_2$. The two top log-log plots confirm that the metric convergences with a rate $1/n$ for $n \in [1, 8]$. The bottom plot is of the metric ρ^n for varying ϵ.

8.3 Linear metrics on the space of log signatures

In this section, we propose some potential similarity measures on the space of log signatures. The space of log signatures can be regarded as a formal Lie algebra for the group of signatures, and will, therefore, be a flat subspace of the dual tensor algebra $T((\mathbb{R}^d))$. This linearity will make calculating distances less involved, compared to the case for the curved subspace of signatures.

Concatenation metric The first similarity measure we consider is again based on the group actions of the signature space, that is that reversing a path yields an inverse and that the signature map is a homeomorphism over concatenation. Let $x, y : [0,1] \to \mathbb{R}^d$ be paths, we define the *log-concatenation metric* as

$$\sigma_*^n(x,y) := \left\| \log S(x * \overleftarrow{y})_{0,1}^{(n)} \right\|_2 + \left\| \log S(y * \overleftarrow{x})_{0,1}^{(n)} \right\|_2, \tag{8.5}$$

where the superscript (n) again denotes the signature truncated to length n words.

That this, in fact, is a metric can be deduced from the similar group metric (8.3) being a metric, and that the log map is injective and the L_2-norm well-defined.

Linear metric With the flat geometry of the space of log signatures, it is also possible to define similarity measures independent of the group structure of the signature space. We will refer to the similarity measure introduced in this section as the "linear metric". Let $x, y : [0,1] \to \mathbb{R}^d$ be paths, we define the linear metric as the normed difference

$$\sigma_L^n(x,y) = \left\| \log S(x)_{0,1}^{(n)} - \log S(y)_{0,1}^{(n)} \right\|_2. \tag{8.6}$$

Similarly, this will also be a metric due to the uniqueness and the L_2-norm being well-defined.

8.4 Convergence for the proposed metrics on log signatures for linear paths

The last section revealed some quite surprising convergence results for the group metric defined on the space of signatures. Proving anything similar for metrics on the space of log signatures is however quite difficult, as the bases needed to describe this space are more complex. We, therefore, have to rely on numerical experiments to better understand the proposed metrics. Again, we consider ϵ-perturbations of the linear paths $x, x_\epsilon : [0,1] \to \mathbb{R}^4$

$$x(t) := at, \quad x_\epsilon(t) := (a + \epsilon \mathbf{1})t.$$

And again, we investigate the rate of convergence for the limits $\lim_{\epsilon \to 0} \sigma_*^n(x, x_\epsilon) = 0$ and $\lim_{\epsilon \to 0} \sigma_L^n(x, x_\epsilon) = 0$ for small ϵ. In figure 8.3 and figure 8.4 we examine

the behaviour of the log-concatenation metric and log-linear metric, respectively. The linear paths in the experiment have slope

$$a = (0.1, 0.5, 0.0, -0.2)^T.$$

Both metrics seem to convergence linearly with respect to ϵ for small perturbations of linear paths. This could indicate that the metrics are more robust, compared to the previously discussed group metric.

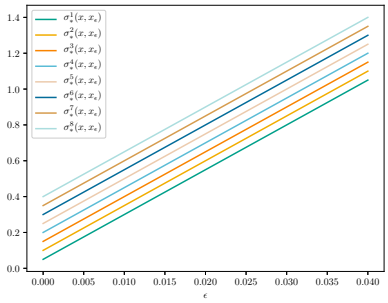


Fig. 8.3. Convergence experiment for the log-concatenation metric σ_*^n for linear paths in $\mathbb{R}^4$. The lines have been shifted by a small constant to make it possible to discern the different metrics.

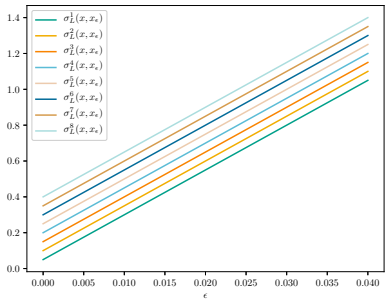


Fig. 8.4. Convergence experiment for the linear-log metric σ_L^n for linear paths in $\mathbb{R}^4$. The lines have been shifted by a small constant to make it possible to discern the different metrics.

9 Using signature-based similarity measures on the shape space of animations

To get a better sense of how the proposed similarity measures compare to the shape analysis framework used in section 6, we will in this section test their ability to identify animations. Since projecting the animations to the space $(\mathcal{S}_*, d_{\mathcal{S}_*})$ let us successfully identify movement, we use the properties of the space as a reference and benchmark for our experiments. That is, if the proposed similarity measures show similar characteristics, we take it as an indication that they give a suitable description of the geometric properties of shapes of animations.

A lot of research has been put into efficient methods for numerically calculating signatures and finding BCH coefficients, for example, [5]. We will use the 'iisignature' package developed by Jeremy Reizenstein and Benjamin Graham [33]. Given an array of data points in $\mathbb{R}^d$, this implementation calculates the truncated signature to degree k of the piece-wise linear interpolation connecting the data. The package also contains methods for calculating log signatures.

Earlier, we observed that the signature of a geodesic interpolation curve in $SO(3)$ can be regarded as the signature of a linear path in $\mathbb{R}^{3d}$. After parsing the animation data to a geodesic interpolation curves in $SO(3)^d$, in the same manner as discussed in section 6, we calculate the right logarithmic derivative using (5.13). By mapping the derivative to $\mathbb{R}^{3d}$ with the hat map, we can calculate the signature and log signatures of the animations using the 'iisignature' package.

The implementation is available on the author's github profile.[4]

9.1 Geodesic interpolation between animations

Having an inverse mapping to the SRVT allows us to interpolate between curves $c_a, c_b \in \mathcal{P}_*$. Let $\overline{c}(s,t)$ denote the linear interpolation defined in equation (5.14), satisfying $\overline{c}(s=0,t) = c_a(t)$ and $\overline{c}(s=1,t) = c_b(t)$. Since we are interpolating linearly between the curves, and since the distance $d_{\mathcal{P}}$ coincides with the geodesic distance in $\mathcal{P}_*$, plotting $d_{\mathcal{P}_*}(\overline{c}(s), c_b)$ should produce a straight reference line with respect to the interpolation parameter s.

An interesting follow-up question to this is then how the proposed similarity measures compare. Whether they deviate from the linearity of the interpolation scheme could give some insight into the geometry of these measures. We will therefore create an interpolation $\overline{c}(s,t)$ between two animations c_a and c_b with descriptions "run/jog" and "walk", which we use to compare the different similarity measures to the geodesic of the Riemannian metric $d_{\mathcal{P}_*}$. Note that we normalize the measures to make it easier to compare their properties. Since we only conduct the experiment for one pair of animations c_a and c_b, the experiment will not give conclusive information about the measures, but rather an indication of their properties.

[4] url: `https://github.com/paalel/Signatures-in-Shape-Analysis`

Group metric In figure 9.1, we have calculated the metric $\rho^n(\overline{c}(s,t), c_b(t))$ for increasing values of the interpolation parameter $s \in [0,1]$. The metric seems to behave with the same convergence characteristics discussed in proposition 1. This gives an indication that the result for linear paths might also be true for piece-wise linear paths. As discussed, having this characteristic could make the proposed metric unsuitable for classification of animations and noisy data.

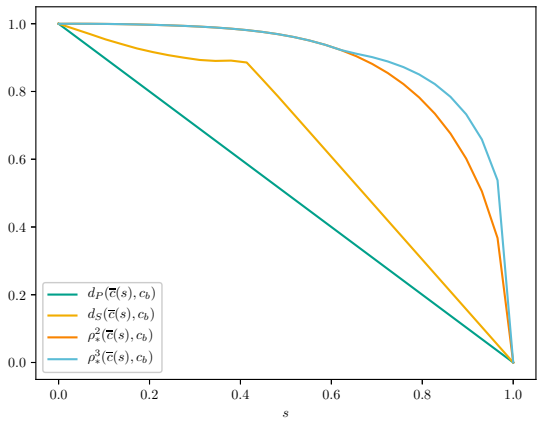


Fig. 9.1. Group metric $\rho^n(\overline{c}(s), c_b)$, for the linear interpolation $\overline{c}(s,t)$ between a walking animation (c_a) and a running animation (c_b).

Metrics on log signatures In figure 9.2, we have repeated the experiment for the metrics $\sigma^n_*(\overline{c}(s,t), c_b(t))$ and $\sigma^n_L(\overline{c}(s,t), c_b(t))$, respectively. Both metrics follow a parabolic arc, compared to the reference - the geodesic distance $d_{\mathcal{P}_*}$. Even though the curvature of the log metrics looks more relaxed, compared to the group metric in plot 9.1, they both deviate quite a bit from the geodesic distance. This could indicate that the metrics yield a different description of the space, compared to the shape metric.

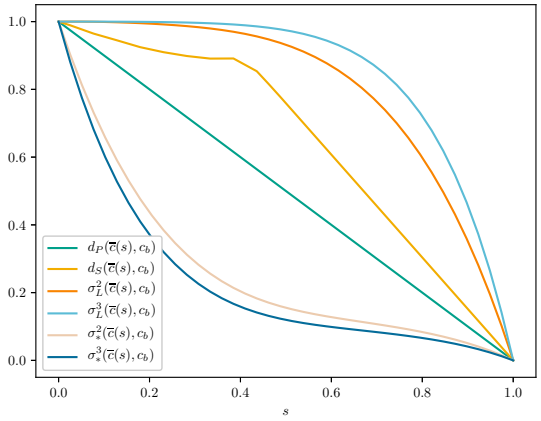


Fig. 9.2. The log linear metric $\sigma^n_L(\overline{c}(s), c_b)$ and log concatenation metric $\sigma^n_*(\overline{c}(s), c_b)$, for the linear interpolation $\overline{c}(s,t)$ between a walking animation (c_a) and a running animation (c_b).

9.2 Identifying movement with signatures

In this section we attempt to identify movement using the proposed metrics on signatures and log signatures. We repeat the experiment from section 6 where we calculate the similarity between pairs of animations with descriptions "walk", "run/jog", "forward jump", as this will make it easier to asses how the methods compare.

When computing the similarity matrix for the shape metric $d_{\mathcal{S}_*}$ in section 6, we had to compute the optimal reparameterization between every pair of animations. In contrast, for signature measures, we only have to compute the signature once for every animation, before calculating the similarity by inexpensive norms. And still, computing the signatures and log signatures is computationally less demanding, compared to the optimal reparameterizations, as the signature computation does not require the expensive optimization step.

In figure 9.3 we create a multi dimensional scaling plot from the similarity matrix generated by the group metric ρ^n for $n = 3$. Even though the three types of movement indeed are separate, the unevenness and asymmetry of the distribution make the results unsatisfying compared to the clusters produced by the shape metric $d_{\mathcal{S}_*}$ in figure 6.3b.

In figure 9.4 we repeat the experiment for the log linear metric σ_L^n with log signatures truncated to degree $n = 3$. Again, the distribution is quite unsatisfying, being both asymmetric and uneven.

And in figure 9.5, we use the concatenation metric on log signatures σ_*^n, with $n = 3$, to generate the similarity matrix for the multi dimensional scaling plot. This metric also suffers from the problems discussed for the other experiments.

It looks as if the way the signature encodes the geometric properties of a path, in a reparameterization invariant way, has potential for describing the equivalence classes of shapes. The proposed signature metrics does, however, not convincingly measure the similarity between shapes. In addition to being both uneven and asymmetric, there is a quite big difference in the internal variance of the clusters, especially compared with the shape metric which displayed a pretty uniform variance.

Having a wide range of values for the metric between animations could be indicative of there being some sort of scaling issue. Either that the n-fold integral coefficients, or the metrics themselves should include some sort of scaling factor to make the comparisons more accurate.

One could see this a parallel to the case for the SRVT, where the transformation indeed does include a scaling to induce a Riemannian metric on the manifold of shapes.

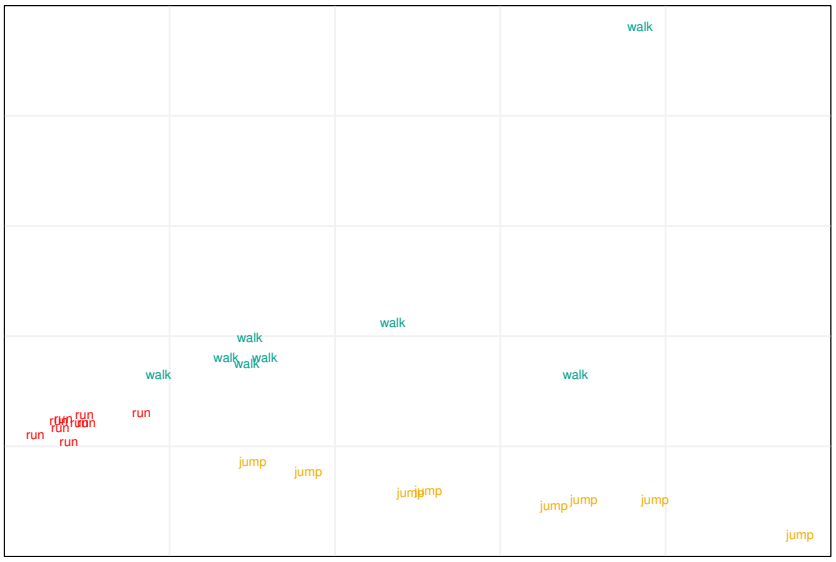


Fig. 9.3. Multi dimensional scaling plots of similarity matrix based on the group metric ρ^n with $n = 3$. In this plot we have taken animation with descriptions "run/jog", "forward jump" and "walk" from the CMU Motion Capture Database [13].

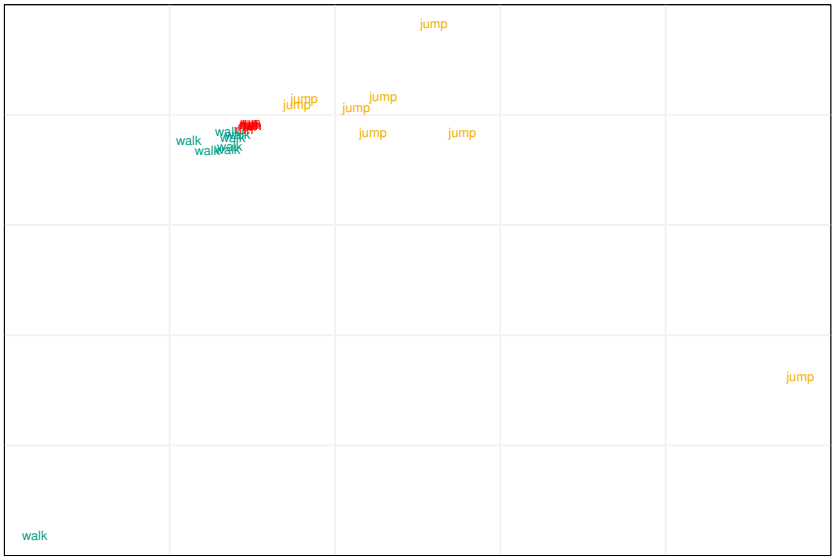


Fig. 9.4. Multi dimensional scaling plots of similarity matrix based on log linear metric σ^3_L. In this plot we have taken animation with descriptions "run/jog", "forward jump" and "walk" from the CMU Motion Capture Database [13].

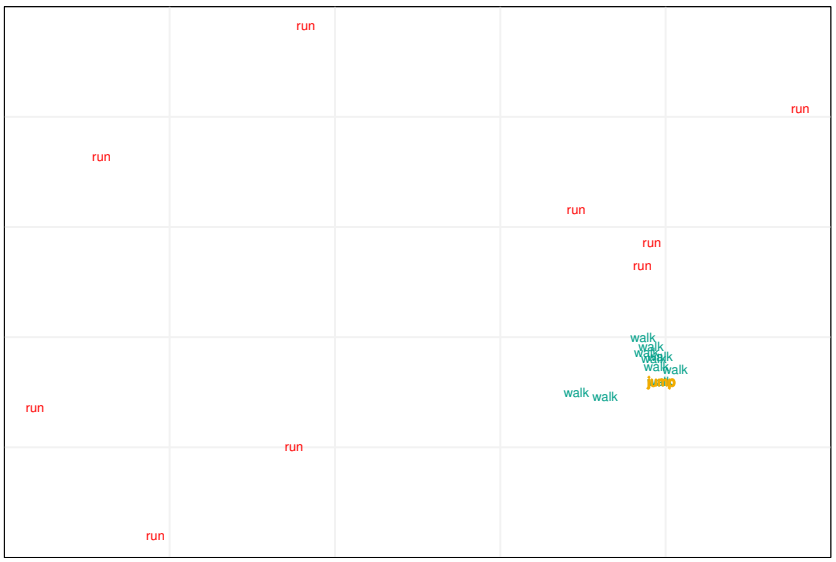


Fig. 9.5. Multi dimensional scaling plots of similarity matrix based on log concatenation metric σ_*^3. In this plot we have taken animation with descriptions "run/jog", "forward jump" and "walk" from the CMU Motion Capture Database [13].

9.3 A normalized linear difference on log signatures

A possible explanation to the observed unevenness in section 9.2 was that the methods fail to scale the similarity measures correctly. To test this hypothesis, we introduce a normalized version of the linear log metric

$$\mathrm{sim}^n(x,y) = \left\| \frac{\log S(x)^{(n)}_{0,1}}{\left\|\log S(x)^{(n)}_{0,1}\right\|_2} - \frac{\log S(y)^{(n)}_{0,1}}{\left\|\log S(y)^{(n)}_{0,1}\right\|_2} \right\|_2, \tag{9.1}$$

and repeat the experiments from earlier with this adapted similarity measure. In figure 9.6, we compare the normalized log difference to the geodesic distance $d_{\mathcal{P}_*}$ in the manifold $\mathcal{P}_*$.

In figure 9.7, we visualize the similarity matrix generated by the similarity measure, for the same set of animations as in section 9.2.

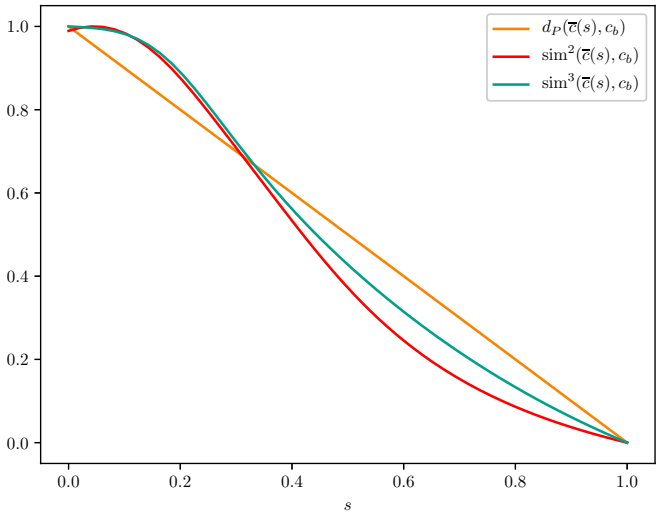


Fig. 9.6. Similarity measure $\mathrm{sim}^n(x,y)$ for the linear interpolation between a walking animation (c_a) and a running animation (c_b).

The normalized log similarity measure produces the most convincing results thus far. Even though one "run"-animation is miss-classified in the "walk" cluster, in figure 9.7, the similarity measure produces a very similar classification to that observed for the shape metric in figure 6.3b. Overall the produced classification is satisfying.

The distance between the interpolation $\overline{c}(s,t)$ and curve $c_b(t)$ in 9.6 aligns close to the metric $d_{\mathcal{P}_*}$, resembling that of the geodesic distance in $(\mathcal{P}_*, d_{\mathcal{P}_*})$.

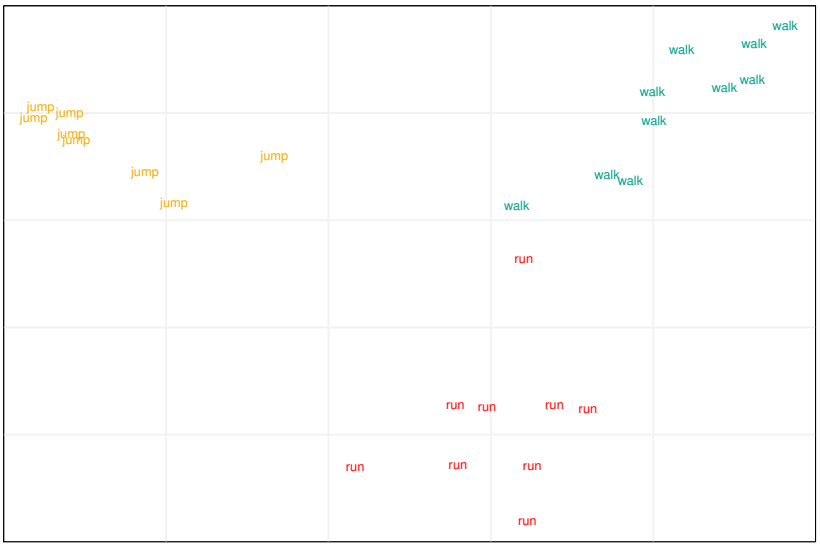


Fig. 9.7. Multi dimensional scaling plots of similarity matrix based on similarity measure sim^3. In this plot we have taken animation with descriptions "run/jog", "forward jump" and "walk" from the CMU Motion Capture Database [13].

In future work, it would be interesting to investigate why the normalization is necessary, and if there exist even better ways of scaling the signatures. One could, for example, consider scaling the n-fold integral coefficients of the signature, or look into more precise metrics for the space of the signatures. Still, the experiments show that signatures do give a useful description of shapes in the context of shape analysis, and could prove useful in understanding this branch of mathematics.

10 Conclusion

In our experiments, we show that it is indeed possible to use a distance function based on signatures to classify animations projected to the shape space $\mathcal{S}_*$. The proposed method is computationally very efficient, even though somewhat less accurate than known methods in shape analysis.

When comparing multiple animations, the Riemannian metric (4.3) requires calculating the very expensive optimal reparameterization between every pair. The proposed signature method instead only requires calculating the signature once for every animation, and then compare animations by computing inexpensive norms. The signatures themselves being cheap to compute, combined with the fact that the proposed method requires no optimization, makes the approach very efficient.

Further increasing the accuracy of the signature method might also be possible by defining a more precise similarity measure. Nonetheless, our results can be seen as proof of concept for using signatures as an efficient way of classifying shapes.

Our experiments also serve as evidence that signatures encode the essential information needed to describe shapes. A better understanding of the connection between signatures and shape analysis could be an important step to better understand the geometry of shapes, with applications potentially reaching beyond classifying animations.

10.1 Acknowledgements

Some of the work done in this thesis will be published as part of the conference proceedings of the 2019 Geometric Science of Information conference (GSI 2019) held at École Nationale de l'Aviation Civile(ENAC) in Toulouse, France[5]. The paper submitted to the conference is a joint work with Professor Elena Celledoni(NTNU) and Dr. Nikolas Tapia(WIAS)[6]. N.T. acknowledges that part of this work was carried out during his tenure of an ERCIM 'Alain Bensoussan' Fellowship Programme. This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie, grant agreement No. 691070.

The data used in this project was obtained from `http://mocap.cs.cmu.edu`. The database was created with funding from NSF EIA-0196217.

The implementation of the numerical experiments in this project is a continuation of work done by Markus Eslitzbichler [15, 2, 7, 6] while working as a PHD candidate at the Department of Mathematical Sciences, NTNU.

[5] url: `https://www.see.asso.fr/en/GSI2019`
[6] url: `http://www.wias-berlin.de/people/tapia/`

Bibliography

[1] Bauer, M., Bruveris, M., Michor, P.: Overview of the geometries of shape spaces and diffeomorphism groups. Journal of Mathematical Imaging and Vision **50** (2013). https://doi.org/10.1007/s10851-013-0490-z

[2] Bauer, M., Eslitzbichler, M., Grasmair, M.: Landmark-guided elastic shape analysis of human character motions. Inverse Problems and Imaging **11**, 601–621 (2015). https://doi.org/10.3934/ipi.2017028

[3] Boedihardjo, H., Geng, X., Lyons, T., Yang, D.: The signature of a rough path: Uniqueness. Advances in Mathematics **293**, 720 – 737 (2016). https://doi.org/https://doi.org/10.1016/j.aim.2016.02.011, `http://www.sciencedirect.com/science/article/pii/S0001870816301104`

[4] Bruveris, M., Michor, P., Mumford, D.: Geodesic completeness for Sobolev metrics on the space of immersed plane curves. Forum of Mathematics, Sigma [electronic only] **2** (2013). https://doi.org/10.1017/fms.2014.19

[5] Casas, F., Murua, A.: An efficient algorithm for computing the baker-campbell-hausdorff series and some of its applications. Journal of Mathematical Physics **50** (11 2008). https://doi.org/10.1063/1.3078418

[6] Celledoni, E., Eslitzbichler, M., Schmeding, A.: Shape Analysis on Lie Groups with Applications in Computer Animation. Journal of Geometric Mechanics Vol. 8 No. 3 (2016), pp. 273-304 (2015)

[7] Celledoni, E., Eidnes, S., Eslitzbichler, M., Schmeding, A.: Shape analysis on Lie groups and homogeneous spaces. In: Nielsen, F., Barbaresco, F. (eds.) Geometric Science of Information. pp. 49–56. Springer International Publishing, Cham (2017)

[8] Chen, K.T.: Iterated integrals and exponential homomorphisms. Proceedings of the London Mathematical Society **s3-4**(1), 502–512 (1954). https://doi.org/10.1112/plms/s3-4.1.502

[9] Chen, K.T.: Integration of paths, geometric invariants and a generalized baker- hausdorff formula. Annals of Mathematics **65**(1), 163–178 (1957), `http://www.jstor.org/stable/1969671`

[10] Chen, K.T.: Integration of paths–a faithful representation of paths by noncommutative formal power series. Transactions of the American Mathematical Society **89**(2), 395–407 (1958), `http://www.jstor.org/stable/1993193`

[11] Chevyrev, I., Kormilitzin, A.: A primer on the signature method in machine learning. CoRR **abs/1603.03788** (2016)

[12] Chevyrev, I., Nanda, V., Oberhauser, H.: Persistence paths and signature features in topological data analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence (2018). https://doi.org/10.1109/TPAMI.2018.2885516, to appear

[13] CMU Graphics Lab: CMU Graphics Lab Motion Capture Database (url visit date: 2018-12-10), `http://mocap.cs.cmu.edu/`

[14] Ellingson, L.A.: Statistical shape analysis on manifolds with applications to planar contours and structural proteomics (2011), http://purl.flvc.org/fsu/fd/FSU_migr_etd-0053
[15] Eslitzbichler, M.: Modelling character motions on infinite-dimensional manifolds. The Visual Computer **31** (2014). https://doi.org/10.1007/s00371-014-1001-y
[16] Gallego, G., Yezzi, A.J.: A compact formula for the derivative of a 3-d rotation in exponential coordinates. Journal of Mathematical Imaging and Vision **51**, 378–384 (2015). https://doi.org/10.1007/s10851-014-0528-x
[17] Geng, X.: Reconstruction for the signature of a rough path. Proceedings of the London Mathematical Society **114**(3), 495–526 (2017). https://doi.org/10.1112/plms.12013
[18] Gergely Gyurkó, L., Lyons, T., Kontkowski, M., Field, J.: Extracting information from the signature of a financial data stream (07 2013)
[19] Godil, A.: Applications of 3d shape analysis and retrieval. In: 2009 IEEE Applied Imagery Pattern Recognition Workshop (AIPR 2009) (2009). https://doi.org/10.1109/AIPR.2009.5466293
[20] Hambly, B., Lyons, T.: Uniqueness for the signature of a path of bounded variation and the reduced path group. Annals of Mathematics **171**(1), 109–167 (2010), http://www.jstor.org/stable/27799199
[21] Huang, Z., Wan, C., Probst, T., Van Gool, L.: Deep Learning on Lie Groups for Skeleton-based Action Recognition. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 1243 – 1252 (2017). https://doi.org/10.3929/ethz-b-000184741, 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Conference Location: Honolulu, HI, USA; Conference Date: July 21-26, 2017
[22] Kriegl A., Michor, P.: The Convenient Setting of Global Analysis. American Mathematical Society (1997). https://doi.org/http://dx.doi.org/10.1090/surv/053
[23] Kruskal, J., Wish, M.: Multidimensional Scaling. Sage Publications (1978)
[24] Lyons, T.: Rough paths, signatures and the modelling of functions on streams. In: Proceedings of the International Congress of Mathematicians 2014 (2014)
[25] Lyons, T., Victoir, N.: An extension theorem to rough paths. Annales de l'I.H.P. Analyse non linéaire **24**(5), 835–847 (2007). https://doi.org/10.1016/j.anihpc.2006.07.004, http://www.numdam.org/item/AIHPC_2007__24_5_835_0
[26] Lyons, T.J.: Differential equations driven by rough signals. Revista Matemática Iberoamericana **14**(2), 215–310 (1998). https://doi.org/10.4171/RMI/240
[27] Lyons, T.J., Sidorova, N.: On the radius of convergence of the logarithmic signature. Illinois J. Math. **50**(1-4), 763–790 (2006). https://doi.org/10.1215/ijm/1258059491, https://doi.org/10.1215/ijm/1258059491
[28] Madsen, I., Tornehave, J.: From calculus to cohomology. de Rham cohomology and characteristic classes. Cambridge University Press, Cambridge (1997)

[29] Michor, P., Mumford, D.: Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. Documenta Mathematica **10** (2004)

[30] Michor, P., Mumford, D.: D.: An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. Applied and Computational Harmonic Analysis **23**, 74–113 (2006). https://doi.org/10.1016/j.acha.2006.07.004

[31] Olver, P.J.: Equivalence, Invariants, and Symmetry. Cambridge University Press (1995)

[32] Ree, R.: Lie elements and an algebra associated with shuffles. Annals of Mathematics **68**(2), 210–220 (1958), http://www.jstor.org/stable/1970243

[33] Reizenstein, J., Graham, B.: The iisignature library: efficient calculation of iterated-integral signatures and log signatures. CoRR **abs/1802.08252** (2016)

[34] Sebastian, T.B., Klein, P.N., Kimia, B.B.: On aligning curves. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(1), 116–125 (2003). https://doi.org/10.1109/TPAMI.2003.1159951

[35] Shilane, P., Min, P., Kazhdan, M., Funkhouser, T.: The Princeton shape benchmark. In: Shape Modeling International (2004)

[36] Shingel, T.: Interpolation in special orthogonal groups. IMA Journal of Numerical Analysis **29**(3), 731–745 (2009). https://doi.org/10.1093/imanum/drn033

[37] Srivastava, A., Klassen, E., Joshi, S.H., Jermyn, I.H.: Shape analysis of elastic curves in euclidean spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence **33**(7), 1415–1428 (2011)

[38] Su, Z., Klassen, E., Bauer, M.: The Square Root Velocity Framework for Curves in a Homogeneous Space. Conference: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) pp. 680–689 (2017). https://doi.org/10.1109/CVPRW.2017.97

[39] Tan, M., Qiu, A.: Multiresolution diffeomorphic mapping for cortical surfaces. Information processing in medical imaging : proceedings of the conference **24**, 315—326 (2015), http://europepmc.org/abstract/MED/26221683

[40] Younes, L.: Computable elastic distances between shapes. SIAM J. of Applied Math pp. 565–586 (1998)