# Curve Signatures in Theory and Applications

Isak Hammer*†
August, 2024
*isakhammer@gmail.com †Department of Mathematics, UiB

***Abstract*—What are Curve signatures? How can they be used and lastly**

## I. Introduction

### A. Context

### B. Outline of this report

Signatures has proved to be fundamental for several applications. The theory is quite applicable in the algebraic theory of rough paths [1]. This has showed to be a very interesting way to represent data because of the special properties of time reversial propertie and ivariance of time reparamerization of representing data via signatures [2]. In fact, in a 2013 competition focused on recognizing handwritten Chinese characters, the winner represented the characters as arrays based on a "signature" from rough path theory, then classified them using a convolutional neural network [3]. Recently have we sen that it is clear that signatures can be a great tool of extracting geometric shape of any path given missing data.

Recently, it has become clear that signatures can be a powerful representation in machine learning for extracting the geometric shape of any path, even when data is missing [2] and time series .

We will divide the report into two parts. One part for the fundamental properties of signatures and the second pat for the recent applications of this method.

### C. Signatures

The concept of the signature approach is to extract characteristic features from a data, that is a function or data points in a non parametric way. First we want to define the so-called path integral. Consider a two parameterized one dimensional paths $Y_t : [a,b] \to \mathbb{R}^d$ and $X_t : [a,b] \to \mathbb{R}$, then we say the path integral of $Y_t$ against $X_t$ is

$$\int_a^b Y_t dX_t = \int_a^b Y_t \dot{X}_t dt, \tag{1}$$

where we defined $\dot{X}_t = \frac{d}{dt}X(t)$. Following [4] and [2], will we explain the basic definition of a signature.

A fundamental piece in the definition of a signature is the so-called path integral. Lets consider the parametrized smooth path of $d$ dimensions be $X_t : [a,b] \to \mathbb{R}^d$ such that $X_t = \{X_t^1, X_t^2, \dots, X_t^d\}$. Now since each path is $X_t^i : [a,b] \to \mathbb{R}$ for $i \in \{1,\dots,d\}$ , we say define the integral

$$S(X)^i_{a,t} = \int_{a<s<t} dX^i = X_t^i - X_0^i \tag{2}$$

Similarly we define the double-iterated double integral

$$S(X)^{i,j}_{a,t} = \int_{a<s<t} S(X)^i_{a,s} dX_s^j = \int_{\substack{a<r<t \\ a<s<t}} dX_r^i dX_s^j$$

Continuing recursively we obtain the definition

$$S(X)^{i_1,\dots,i_k}_{a,t} = \int_{a<s<t} S(X)^{i_1,\dots,i_{k-1}} dX_s^{i_k}$$

where $i_1,\dots,i_k \in \{1,\dots,d\}$. Notice that we still obtain the mapping $S(X)^{i_1,\dots,i_k} : [a,b] \to \mathbb{R}$. Finally we have the tools rquired to define a signature.

**Definition I.1** (Signature)**.** *We say a signature of a path $X : [a,b] \to \mathbb{R}^d$, denoted by $S(X)_{a,b}$ is the collection of all the iterated integrals of $X$. Thus we, nor have the sequence of numbers*

$$S(X)_{a,b} = (1, S(X)^1_{a,b}, \dots, S(X)^d_{a,b}, S(X)^{1,1}_{a,b}, S(X)^{1,2}_{a,b}, S(X)^{2,1}_{a,b}, \dots). \tag{3}$$

*Here the first term is defined as 1. Keep in mind that we iterate over all multi-indexes, that is the set*

$$W = \left\{ \begin{array}{c} (i_1,\dots,i_k) \text{ where } k \geq 1, \\ \text{for all } i_1,\dots,i_k \in \{1,\dots,d\}. \end{array} \right\} \tag{4}$$

*We denote the set $W$ as words and $A = \{1,\dots,d\}$ as the alphabet of $d$ letters .*

One of the most fundamental properties of the signature is its invariance under time reparameterization. This can easily be demonstrated using the definitions of the path integral. Consider two paths $X, Y : [a,b] \to \mathbb{R}$, which are real-valued. Now, consider two corresponding reparameterized paths $\widetilde{X}, \widetilde{Y} : [a,b] \to \mathbb{R}$, where $\widetilde{X}_t = X_{\psi(t)}$ and $\widetilde{Y} = Y_{\psi(t)}$, with some smooth reparameterization $\psi : [a,b] \to [a,b]$. From the chain rule it is clear that

$$\frac{d}{dt}\widetilde{X}_t = \dot{\widetilde{X}}_t \dot{\psi}(t)$$

, thus it follows that

$$\int_a^b \widetilde{Y}_t d\widetilde{X}_t = \int_a^b Y_{\psi(t)} \dot{X}_{\psi(t)} \dot{\psi}(t) = \int_a^b Y_u dX_u \tag{5}$$

Here we used the subsitution $u = \psi(t)$. This is of course applicable in the multidimensional case in the case of a signature. Let $\widetilde{X}, X : [a,b] \to \mathbb{R}d$ where $\widetilde{X}_t = X_t$.Then we see that

$$S(\widetilde{X})^{i_1,\dots,i_k}_{a,b} = S(X)^{i_1,\dots,i_k}_{a,b} \tag{6}$$

for any $i_1, \dots, i_k \in \{1, \dots, d\}$. Thus we see that the signature is, in fact, invariant under time re parameterization.

### D. Shuffle product

A fundamental property of the signature is that the product of two signature terms $S(X)_a^{i_1,\dots,i_k,b}$ and $S(X)_a^{j_1,\dots,j_m,b}$ can be expressed as a sum of terms depending on shuffled multi-indexes.

To formalize this, we define the shuffle product for two multi-indexes. A permutation $\sigma$ of the set $\{1, \dots, k+m\}$ is called a $(k,m)$-shuffle if $\sigma^{-1}(1) < \cdots < \sigma^{-1}(k)$ and $\sigma^{-1}(k+1) < \cdots < \sigma^{-1}(k+m)$. The list $(\sigma(1), \dots, \sigma(k+m))$ forms a shuffle of $(1, \dots, k)$ and $(k+1, \dots, k+m)$. Let Shuffles$(k,m)$ denote the collection of all such shuffles.

For multi-indexes $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_m)$ with $i_1, \dots, i_k, j_1, \dots, j_m \in \{1, \dots, d\}$, define the multi-index

$$I \sqcup\!\sqcup J = \{(\sigma(1), \dots, \sigma(k+m)) \mid \sigma \in \text{Shuffles}(k,m)\}.$$

Thus, we have that The shuffle product $I \sqcup\!\sqcup J$ is the set of multi-indexes of length $k+m$.

**Theorem I.1.** ***Shuffle product identity*** *For a path $X : [a,b] \to \mathbb{R}^d$ and multi-indexes $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_m)$, it holds that*

$$S(X)_a^b S(X)_a^b = \sum_{K \in I \sqcup\!\sqcup J} S(X)_a^{K,b}.$$

Let the terms $e_{i_1}, \dots, e_{i_k}$ be monomials. Then we can denote the representation $S(X)_{a,b}$ as a formal power series

$$S(X)_{a,b} = \sum_{k=0}^{\infty} \sum_{i_1,\dots,i_k} S(X)_{a,b}^{i_1,\dots,i_k} e_{i_1} \cdot \ldots \cdot e_{i_k} \tag{7}$$

The main reason why this is fundamental is because this is necessary to state the so-called Chens identity which states the relationship between concatenation and tensor product. First we define the concatenation of paths. That is. For two paths $X : [a,b] \to \mathbb{R}^d$ and $Y : [b,c] \to \mathbb{R}^d$, we define the concatenation as the path $X * Y : [a,c] \to \mathbb{R}^d$, where the first part for $t \in [a,b]$ is $(X * Y)_t = X_t$, and for $t \in [b,c]$ we define $(X * Y)_t = X_b + (Y_t - Y_b)$. And the tensor product is simply defined joining the monomials, that is $e_{i_1} \dots e_{i_k} \otimes e_{j_1} \dots e_{j_m} = e_{i_1} \dots e_{i_k} e_{j_1} \dots e_{j_m}$. Finally we can state the relationship between these operators.

**Theorem I.2.** ***Chens Identity.*** *Let $X : [a,b] \to \mathbb{R}^d$ and $Y : [b,c] \to \mathbb{R}^d$ be two paths. Then we have the following identity,*

$$S(X * Y)_{a,c} = S(X)_{a,b} \otimes S(Y)_{b,c} \tag{8}$$

A very interesting property with the signature (7) is in fact time-reversible. For a path $X : [a,b] \to \mathbb{R}^d$ we can define the time-reverse path $\overleftarrow{X} : [a,b] \to \mathbb{R}^d$, for which $\overleftarrow{X} = X_{a+b-t}$. It can be shown that

$$S(X)_{a,b} \otimes S(\overleftarrow{X}) = 1. \tag{9}$$

This can be understood that that the time-reverse is the "tensor" inverse for the signature $S(X)_{a,b}$.

### E. Log signature

For a path $X : [a,b] \mapsto \mathbb{R}^d$, the log signature of $X$ is defined as the formal power series $\log S(X)_{a,b}$.

For two formal power series $x$ and $y$, let us define their Lie bracket by

$$[x,y] = x \otimes y - y \otimes x. \tag{1.55}$$

It is clear that the first few terms of the log signature are given by

$$\begin{aligned} \log S(X)_{a,b} &= \sum_{i=1}^{d} S(X)_{a,b}^{i} e_i \\ &+ \sum_{1 \le i < j \le d} \frac{1}{2} \left( S(X)_{a,b}^{i,j} - S(X)_{a,b}^{j,i} \right) [e_i, e_j] \\ &+ \dots . \end{aligned} \tag{10}$$

### F. Rough paths

One of the key properties of the shuffle product is that it enables the representation of the signature of a nonlinear function as a linear combination of iterated integrals. Moreover, we observe that the time reversal property, Chen's identity, and the invariance of the signature under time reparametrizations are preserved.

A natural question arises: what kind of information is captured by the signature? Due to the invariance of time reparametrization, it is clear that we cannot reconstruct exact speed at which the path is traversed [2]. However, a another interesting propertie is that if we have a paramterized picture of a path $X$ which never crossing itself, then we can completely describe the image and diraction of traversal of path. This implies that one can certainly reconstruct geometrical properties, see [5, 6, 7]

## II. The Signature Uniqueness Theorem

Recent work combines signature transforms from rough path theory with neural networks to tackle problems in data science (cf. [**LZL19**], [**WIN19**], [**XSJ17**]). The central task is to learn the nonlinear relationship between an input data stream (a sequence of points in a vector space) and an output. For instance, atmospheric data from the last 24 hours could be used to predict weather for the next 24 hours.

An input stream is naturally ordered, allowing it to be represented as a continuous path in a vector space $V$. The signature transform $S(x)$ maps a path

$x : [0,T] \to V$ to a sequence of global iterated integrals:

$$S(x) = \left( \int_0^T dx_t \otimes dx_t, \int_0^{t_1} \cdots \int_0^{t_n} dx_{t_1} \otimes \cdots \otimes dx_{t_n}, \ldots \right)$$

In practice, a truncated signature $S_N(x)$ is used as a feature set for learning the relationship between the path and output. This approach can leverage traditional statistical methods or deep learning techniques.

The theoretical foundation of this approach is the *signature uniqueness theorem*, which asserts that every rough path is uniquely determined by its signature, except for tree-like segments. This makes the signature a powerful tool for encoding all relevant information in the data stream. The effectiveness of this method is captured in two key properties:

**Property 1.** The truncated signature $S_N(x)$ converges to the full signature $S(x)$ rapidly, even for moderate $N$, ensuring minimal information loss.

**Property 2.** Polynomial functions on the signature space are linear, as implied by the shuffle product formula ([**formula115**]). Thus, lifting the problem to the signature space linearizes it, making the solution robust and model-free.

## III. Rough paths

## References


[1] Xi Geng. "An Introduction to the Theory of Rough Paths". In: *Lecture Notes* (2021), p. 9.

[2] Ilya Chevyrev and Andrey Kormilitzin. "A primer on the signature method in machine learning". In: *arXiv preprint arXiv:1603.03788* (2016).

[3] Fei Yin et al. "Chinese handwriting recognition competition". In: *International Journal on Document Analysis and Recognition (ICDAR)* 12.4 (2013), pp. 1464–1470.

[4] Jeremy Reizenstein. "Calculation of Iterated-Integral Signatures and Log Signatures". In: *Centre for Complexity Science, University of Warwick* (2016).

[5] Terry J Lyons and Weijun Xu. "Hyperbolic development and inversion of signature". In: *Journal of Functional Analysis* 272.7 (2017), pp. 2933–2955.

[6] Jiawei Chang and Terry Lyons. "Insertion algorithm for inverting the signature of a path". In: *arXiv preprint arXiv:1907.08423* (2019).

[7] Xi Geng. "Reconstruction for the signature of a rough path". In: *Proceedings of the London Mathematical Society* 114.3 (2017), pp. 495–526.