CrossMark

Invariants of Multidimensional Time Series Based on Their Iterated-Integral Signature

Joscha Diehl[1] · Jeremy Reizenstein[2]

Abstract We introduce a novel class of features for multidimensional time series that are invariant with respect to transformations of the ambient space. The general linear group, the group of rotations and the group of permutations of the axes are considered. The starting point for their construction is Chen's iterated-integral signature.

1 Introduction

The analysis of multidimensional time series is a standard problem in data science. Usually, as a first step, features of a time series must be extracted that are (in some sense) robust and that characterize the time series. In many applications the features should additionally be invariant to a particular group acting on the data. In Human Activity Recognition for example, the orientation of the measuring device is often unknown. This leads to the requirement of rotation invariant features [37]. In EEG analysis, invariants to the general linear group are beneficial [12]. In other applications, the labeling of coordinates is arbitrary, which leads to permutation invariant features.

As any time series in discrete time can, via linear interpolation, be thought of as a multidimensional curve, one is naturally led to the search of invariants of curves. Invariant features of curves have been treated using various approaches, mostly focussing on two-dimensional curves. Among the techniques are Fourier series (of closed curves) [21, 27, 52], wavelets [6], curvature based methods [2, 36] and integral invariants [13, 35].

The usefulness of iterated integrals in data analysis has recently been realized, see for example [20, 26, 32, 51] and the introduction in [5]. Let us demonstrate the appearance of

✉ J. Diehl
joscha.diehl@gmail.com

[1] Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

[2] Centre for Complexity Science, University of Warwick, Coventry, UK

iterated integrals on a very simple example. Let $X : [0, T] \to \mathbb{R}^2$ be a smooth curve. Say we are looking for a feature describing this curve that remains unchanged if one is handed a rotated version of X. Maybe the simplest one that one can come up with is the (squared) total displacement length $|X_T - X_0|^2$. Now,

$$
\begin{aligned}
|X_T - X_0|^2 &= \left(X_T^1 - X_0^1\right)^2 + \left(X_T^2 - X_0^2\right)^2 \\
&= 2 \int_0^T \left(X_r^1 - X_0^1\right) \dot{X}_r^1 dr + 2 \int_0^T \left(X_r^2 - X_0^2\right) \dot{X}_r^2 dr \\
&= 2 \int_0^T \left(\int_0^r \dot{X}_u^1 du\right) \dot{X}_r^1 dr + 2 \int_0^T \left(\int_0^r \dot{X}_u^2 du\right) \dot{X}_r^2 dr \\
&= 2 \int_0^T \int_0^r dX_u^1 dX_r^1 + 2 \int_0^T \int_0^r dX_u^2 dX_r^2,
\end{aligned}
$$

where we have applied the fundamental theorem of calculus twice and then introduced the notation dX_r^i for $\dot{X}_r^i dr$. We see that we have expressed this simple invariant in terms of iterated integrals of X; the collection of which is usually called its *signature*. The aim of this work can be summarized as describing *all* invariants that can be obtained in this way. It turns out, when formulated in the right way, this search for invariants reduces to classical problems in invariant theory. We note that already in the early work of Chen (see for example [4, Chap. 3]) the topic of invariants arose, although a systematic study was missing (see also [23]).

The aim of this work is threefold. Firstly, we adapt classical results in invariant theory regarding non-commuting polynomials (or, equivalently, multilinear maps), to our situation. These results are spread out in the literature and sometimes need a little massaging. Secondly, it lays out the usefulness of the iterated-integral signature in the search for invariants of d-dimensional curves. We show, see Sect. 7, that certain "integral invariants" found in the literature are in fact found in the signature and our approach simplifies their enumeration. Lastly, we present new geometric insights into some entries found in the signature, Sect. 3.3.[1]

The paper is structured as follows. In the next section we introduce the iterated-integral signature of a multidimensional curve, as well as some algebraic language to work with it. Based on this signature, we present in Sect. 3 and Sect. 4 invariants to the general linear group and the special orthogonal group. Both are based on classical results in invariant theory. For completeness, we present in Sect. 5 the invariants to permutations, which have been constructed in [1]. In Sect. 6 we show how to use all these invariants if an additional (time) coordinate is introduced. In Sect. 7 we relate our work to the integral invariants of [13] and demonstrate that the invariants presented there cannot be complete. We formulate the conjecture of completeness for our invariants and point out open algebraic questions.

For readers who want to use these invariants without having to go into the technical results, we propose the following route. The required notation is presented in the next section. The invariants are presented in Proposition 3.11, Proposition 4.4 and Proposition 5.4. Examples are given in Sect. 3.1 (in particular Remark 3.14), Example 4.7 and Example 5.6. All these invariants are also implemented in the software package [9]. For calculating the iterated-integral signature in Python we propose using the package `iisignature`, as described in [40].

---

[1]The signature is notorious for being hard to interpret in geometric terms.

2 The Signature of Iterated Integrals

By a multidimensional **curve** X we will denote a continuous mapping $X : [0, T] \to \mathbb{R}^d$ of bounded variation.[2] The aim of this work is to find features (i.e. complex or real numbers) describing such a curve that are invariant under the general linear group, the group of rotations and the group of permutations. Note that in practical situations one is usually presented with a discrete sequence of data points in $\mathbb{R}^d$, a multidimensional **time series**. Such a time series can be easily transformed into a (piecewise) smooth curve by linear interpolation.

It was proven in [22], which extends the work of [4], that a curve $X = (X^1, .., X^d)$ is almost completely characterized by the collection of its iterated integrals[3]

$$\int_0^T \int_0^{r_n} \ldots \int_0^{r_2} dX_{r_1}^{i_1} \ldots dX_{r_n}^{i_n}, \quad n \geq 1, \quad i_1, \ldots i_n \in \{1, \ldots, d\}.$$

The collection of all these integrals is called the **signature**[4] of X. In a first step, we can hence reduce the goal

Find functions $\Psi : curves \to \mathbb{R}$ that are invariant under the action of a group G,

to the goal

Find functions $\Psi : signature$ of $curves \to \mathbb{R}$ that are invariant under the action of a group G.

By the shuffle identity (Lemma 2.1), any polynomial function on the signature can be rewritten as a linear function on the signature. Assuming that arbitrary functions are well-approximated by polynomial functions, we are led to the final simplification, which is the goal of this paper

Find linear functions $\Psi : signature$ of $curves \to \mathbb{R}$ that are invariant under the action of a group G.

2.1 Algebraic Underpinning

Let us introduce some algebraic notation in order to work with the collection of iterated integrals. Denote by $T((\mathbb{R}^d))$ the space of formal power series in d *non-commuting* variables $x_1, x_2, \ldots, x_d$. We can conveniently store all the iterated integrals of the curve X in $T((\mathbb{R}^d))$, by defining the **signature** of X to be

$$S(X)_{0,T} := \sum x_{i_1} \ldots x_{i_n} \int_0^T \int_0^{r_n} \ldots \int_0^{r_2} dX_{r_1}^{i_1} \ldots dX_{r_n}^{i_n}.$$

---

[2] The reader might prefer to just think of a (piecewise) smooth curve.

[3] Since X is of bounded variation the integrals are well-defined using classical Riemann-Stieltjes integration (see for example Chap. 6 in [43]). This generalizes the notation in the introduction above beyond smooth curves. It can be pushed much further though. In fact the following considerations are purely algebraic and hence hold for any curve for which a sensible integration theory (in particular: obeying integration by parts) exists. A relevant example is Brownian motion which, although being almost surely nowhere differentiable, nonetheless admits a stochastic (Stratonovich) integral.

[4] Also called the "rough path signature".

Here the sum is taken over all $n \geq 0$ and all $i_1, \ldots, i_n \in \{1, 2, .., d\}$. For $n = 0$ the summand is, for algebraic reasons, taken to be the constant 1.

The algebraic dual of $T((\mathbb{R}^d))$ is $T(\mathbb{R}^d)$, the space of polynomials[5] in $x_1, x_2, \ldots, x_d$. The dual pairing, denoted by $\langle \cdot, \cdot \rangle$ is defined by declaring all monomials to be orthonormal, so for example

$$\left\langle x_1 + 15 \cdot x_1 x_2 - 2 \cdot x_1 x_2 x_1, \ x_1 x_2 \right\rangle = 15.$$

Here, we write the element of $T((\mathbb{R}^d))$ on the left and the element of $T(\mathbb{R}^d)$ on the right. We can "pick out" iterated integrals from the signature as follows

$$\left\langle S(X)_{0,T}, \ x_{i_1} \ldots x_{i_n} \right\rangle = \int_0^T \int_0^{r_n} \ldots \int_0^{r_2} dX_{r_1}^{i_1} \ldots dX_{r_n}^{i_n}.$$

The space $T((\mathbb{R}^d))$ becomes an algebra by extending the usual product of monomials, denoted $\cdot$, to the whole space by bilinearity. Note that $\cdot$ is non-commutative.

On $T(\mathbb{R}^d)$ we often use the shuffle product $\shuffle$ which, on monomials, interleaves them in all order-preserving ways, so for example

$$x_1 \shuffle x_2 x_3 = x_1 x_2 x_3 + x_2 x_1 x_3 + x_2 x_3 x_1.$$

Note that $\shuffle$ is commutative.

Monomials, and hence homogeneous polynomials, have the usual concept of **order** or homogeneity. For $n \geq 0$ we denote the projection on polynomials of order n by π_n, so for example

$$\pi_2(x_1 + 15 \cdot x_1 x_2 - 2 \cdot x_1 x_2 x_1) = 15 \cdot x_1 x_2.$$

See [41] for more background on these spaces.

As mentioned above, every polynomial expression in terms of the signature can be re-written as a linear expression in (different) terms of the signature. This is the content of the following lemma, which is proven in [39] (see also [41, Corollary 3.5]).

Lemma 2.1 (Shuffle identity) *Let $X : [0, T] \to \mathbb{R}^d$ be a continuous curve of bounded variation, then for every $a, b \in T(\mathbb{R}^d)$*

$$\left\langle S(X)_{0,T}, a \right\rangle \left\langle S(X)_{0,T}, b \right\rangle = \left\langle S(X)_{0,T}, a \shuffle b \right\rangle.$$

Remark 2.2 We have used this fact already in the introduction, where we confirmed by hand that

$$\left(\left\langle S(X)_{0,T}, x_1 \right\rangle \right)^2 + \left(\left\langle S(X)_{0,T}, x_2 \right\rangle \right)^2 = 2 \left\langle S(X)_{0,T}, x_1 x_1 \right\rangle + 2 \left\langle S(X)_{0,T}, x_2 x_2 \right\rangle$$
$$\left(= \left\langle S(X)_{0,T}, x_1 \shuffle x_1 \right\rangle + \left\langle S(X)_{0,T}, x_2 \shuffle x_2 \right\rangle \right).$$

The concatenation of curves is compatible with the product on $T((\mathbb{R}^d))$ in the following sense (for a proof, see for example [16, Theorem 7.11]).

---

[5] In contrast to a power series, a polynomial only has finitely many terms.

Lemma 2.3 (Chen's relation) *For curves $X : [0, T] \to \mathbb{R}^d$, $Y : [0, T] \to \mathbb{R}^d$ denote their concatenation*

$$X \sqcup Y : [0, 2T] \to \mathbb{R}^d,$$

as $X_.$ on $[0, T]$ and $Y._{-T} - Y_0 + X_T$ on $[T, 2T]$. Then

$$S(X \sqcup Y)_{0,2T} = S(X)_{0,T} \cdot S(Y)_{0,T}.$$

We will use the following fact repeatedly, which also explains the commonly used name **tensor algebra** for $T(\mathbb{R}^d)$.

Lemma 2.4 *The space of all multilinear maps on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$ (n-times) is in a one-to-one correspondence with homogeneous polynomials of order n in the non-commuting variables $x_1, \ldots, x_d$ by the following bijective linear map*

$$\psi \mapsto \mathsf{poly}(\psi) := \sum_{i_1,\ldots,i_n \in \{1,\ldots,d\}} \psi(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) x_{i_1} \cdot x_{i_2} \cdot \ldots \cdot x_{i_n},$$

with e_i being the i-th canonical basis vector of $\mathbb{R}^d$.

For example, with $d = 2$ and $n = 3$, we can consider the multilinear map ψ which takes $\big((a_1, b_1), (a_2, b_2), (a_3, b_3)\big) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$ to the number $a_1 a_2 b_3$. It maps to $\mathsf{poly}(\psi) = x_1 x_1 x_2$.

3 General Linear Group

Let

$$\mathrm{GL}\big(\mathbb{R}^d\big) = \big\{A \in \mathbb{R}^{d \times d} : \det(A) \neq 0\big\},$$

be the general linear group of $\mathbb{R}^d$.

Definition 3.1 For $w \in \mathbb{N}$, we call $\phi \in T(\mathbb{R}^d)$ a GL **invariant of weight** w if

$$\big\langle S(AX)_{0,T}, \phi \big\rangle = (\det A)^w \big\langle S(X)_{0,T}, \phi \big\rangle$$

for all $A \in \mathrm{GL}(\mathbb{R}^d)$ and all curves X.

Definition 3.2 Define a linear action of $\mathrm{GL}(\mathbb{R}^d)$ on $T((\mathbb{R}^d))$ and $T(\mathbb{R}^d)$, by specifying on monomials

$$A x_{i_1}..x_{i_n} := \sum_j (A e_{i_1})_{j_1} x_{j_1}..(A e_{i_n})_{j_n} x_{j_n}$$

$$= \sum_j A_{j_1 i_1}..A_{j_n i_n} x_{j_1}..x_{j_n}.$$

Lemma 3.3 *For all $A \in \mathbb{R}^{d \times d}$ and any curve X,*

$$\big\langle S(AX)_{0,T}, \phi \big\rangle = \big\langle S(X)_{0,T}, A^\top \phi \big\rangle.$$

Proof It is enough to verify this on monomials $\phi = x_{\ell_1}..x_{\ell_m}$. Then, since the ℓ_r-th component of the curve AX is equal to $(AX)^{\ell_r} = \sum_{j_r} A_{\ell_r j_r} X^{j_r}$, we get

$$
\begin{aligned}
\langle S(AX), \phi \rangle &= \int d(AX)^{\ell_1} \ldots d(AX)^{\ell_m} \\
&= \sum_j A_{\ell_1 j_1} \ldots A_{\ell_m j_m} \int dX^{j_1} \ldots dX^{j_m} \\
&= \left\langle S(X), \sum_j A_{\ell_1 j_1} x_{j_1} .. A_{\ell_m j_m} x_{j_m} \right\rangle \\
&= \langle S(X), A^\top \phi \rangle. \qquad \square
\end{aligned}
$$

We can simplify the concept of GL invariants further, using the next lemma. Owing to the shuffle identity, signatures of curves live in a nonlinear subset of the whole tensor algebra $T((\mathbb{R}^d))$, the set of "grouplike elements" (compare [41, Sect. 3.1]). It turns out though that they linearly span all of $T((\mathbb{R}^d))$.

Lemma 3.4 *For $n \geq 1$*

$$
\operatorname{span}\{\pi_n S(X)_{0,T} : X \text{ curve}\} = \pi_n T((\mathbb{R}^d)). \tag{1}
$$

Proof It is clear by definition that the left hand side of (1) is included in $\pi_n T((\mathbb{R}^d))$. We show the other direction and use ideas of [3, Proposition 4]. Let $x_{i_n} \cdot \ldots \cdot x_{i_1} \in \pi_n T((\mathbb{R}^d))$ be given. Let X be the piecewise linear path that results from the concatenation of the vectors $t_1 e_{i_1}, t_2 e_{i_2}$ up to $t_n e_{i_n}$, where $e_i, i = 1, .., d$ is the standard basis of $\mathbb{R}^d$. Its signature is given by (see for example [16, Chap. 6])

$$
S(X)_{0,1} = \exp(t_n x_{i_n}) \cdot \ldots \cdot \exp(t_1 x_{i_1}) =: \phi(t_1, \ldots, t_n),
$$

where the exponential function is defined by its power series. Then

$$
\frac{d}{dt_n} \ldots \frac{d}{dt_1} \phi(0, \ldots, 0) = x_{i_n} \cdot \ldots \cdot x_{i_1}.
$$

Combining this with the fact that left hand side of (1) is a closed set we get that

$$
x_{i_n} \cdot \ldots \cdot x_{i_1} \in \operatorname{span}\{\pi_n(S(X)_{0,1}) : X \text{ curve}\}.
$$

These elements span $\pi_n T((\mathbb{R}^d))$, which finishes the proof. $\qquad \square$

Hence ϕ is a GL invariant of weight w in the sense of Definition 3.1 if and only if for all $A \in \operatorname{GL}(\mathbb{R}^d)$

$$
A^\top \phi = (\det A)^w \phi.
$$

Since the action respects homogeneity, we immediately obtain that projections of invariants are invariants (take $B = (\det A)^{-w} A^\top$ in the following lemma):

Lemma 3.5 *If $\phi \in T(\mathbb{R}^d)$ satisfies*

$$B\phi = \phi,$$

for some $B \in \mathrm{GL}(\mathbb{R}^d)$ then

$$B\pi_n\phi = \pi_n\phi,$$

for all $n \geq 1$.

Proof By definition, the action of GL on $T(\mathbb{R}^d)$ commutes with π_n. □

In order to apply classical results in invariant theory, we use the bijection poly between multilinear functions and non-commuting polynomials, given in Lemma 2.4.

Lemma 3.6 *For $\psi : (\mathbb{R}^d)^{\times n} \to \mathbb{R}$ multilinear and $A \in \mathrm{GL}(\mathbb{R}^d)$,*

$$\mathsf{poly}\big[\psi(A\cdot)\big] = A^\top \mathsf{poly}[\psi].$$

Proof

$$\mathsf{poly}\big[\psi(A\cdot)\big] = \sum_i \psi(Ae_{i_1}, ..Ae_{i_n})x_{i_1}..x_{i_n}$$

$$= \sum_{i,j} A_{j_1 i_1}..A_{j_n i_n}\psi(e_{j_1}, ..e_{j_n})x_{i_1}..x_{i_n}$$

$$= \sum_j \psi(e_{j_1}, ..e_{j_n})A^\top x_{j_1}..x_{j_n}$$

$$= A^\top \mathsf{poly}[\psi]. \qquad \square$$

The simplest multilinear function

$$\Psi : \big(\mathbb{R}^d\big)^{\times n} \to \mathbb{R},$$

satisfying $\Psi(Av_1, .., Av_n) = \det(A)\Psi(v_1, .., v_n)$ that one can maybe think of, is the determinant itself. That is, $n = d$ and

$$\Psi(v_1, .., v_n) = \det[v_1 v_2..v_n],$$

where $v_1 v_2..v_n$ is the $d \times d$ matrix with columns v_i. Up to a scalar this is in fact the only one, and it turns out that invariants of higher weight are built only using determinants as a building block.

To state the following classical result, we introduce the notion of Young diagrams, which play an important role in the representation theory of the symmetric group.

Let $\lambda = (\lambda_1, .., \lambda_r)$ be a partition of $n \in \mathbb{N}$, which we assume ordered as $\lambda_1 \geq \lambda_2 \geq .. \geq \lambda_r$. We associate to it a **Young diagram**, which is an arrangement of n boxes into left-justified rows. There are r rows, with λ_i boxes in the i-th row. For example, the partition $(4, 2, 1)$ of 7 gives the Young diagram
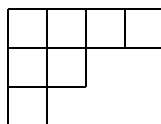
A **Young tableau** is obtained by filling these boxes with the numbers $1, .., n$. Continuing the example, the following is a Young tableau

2	3	7	1
5	4		
6			

A Young tableau is **standard** if the values in every row are increasing (from left to right) and are increasing in every column (from top to bottom). The previous tableau was not standard; the following is.

1	3	5	7
5	4		
6			

The following result is classical, see for example Dieudonné [10, Sect. 2.5], [50] and [18], none of which explicitly give a basis for the invariants though. See [47, Theorem 4.1.12] for a slightly different basis.

Theorem 3.7 *The space of multilinear maps*

$$\psi : \underbrace{\mathbb{R}^d \times \cdots \times \mathbb{R}^d}_{n \text{ times}} \to \mathbb{R}$$

that satisfy

$$\psi(Av_1, Av_2, \ldots, Av_n) = (\det A)^w \psi(v_1, v_2, \ldots, v_n)$$

for all $A \in \mathrm{GL}(\mathbb{R}^d)$ and $v_1, \ldots, v_n \in \mathbb{R}^d$ is non-empty if and only if $n = wd$ for some integer $w \geq 1$.

In that case, a linear basis is given by

$$\left\{ v \mapsto \det[v_{C_1}]..\det[v_{C_w}] \right\}$$

where C_i are the columns of Σ, and Σ ranges over all standard Young tableaux correspond-ing to the partition $\lambda = \overbrace{(w, w, .., w)}^{d \text{ times}}$ of n.

Here, for a sequence $C = (c_1, .., c_d)$, v_C denotes the matrix of column vectors v_{c_i}, i.e.

$$v_C = (v_{c_1}, .., v_{c_d}).$$

Remark 3.8 A consequence of this theorem is the existence of identities between products of determinants. For example, for vectors $v_1, .., v_4 \in \mathbb{R}^2$, one can check by hand

$$\det[v_1 v_4] \det[v_2 v_3] = \det[v_1 v_3] \det[v_2 v_4] - \det[v_1 v_2] \det[v_3 v_4].$$

This is why the product on the left-hand side here is not part of the basis in the previous lemma for $d = 2$, $w = 2$ (compare Sect. 3.1).

Identities of this type are called *Plücker identities*. They have a long history and are a major ingredient in the representation theory of the symmetric group. The procedure of re-ducing certain products of determinants to a basic set of such products is called the *straight-ening algorithm* [44, Sect. 2.6]. See also [30] and [48].

Remark 3.9 The only invariant for $d = 2$, $w = 1$ is

$$x_1 x_2 - x_2 x_1 = [x_1, x_2],$$

a Lie polynomial. One can generally ask for invariant Lie polynomials [41, Sect. 8.6.2]. This seems to be of no relevance to the application of invariant feature extraction for curves though.

Remark 3.10 Let $C_w^{(d)}$ be the number of linear independent invariants of weight w. By Theorem 3.7, this is the number of standard Young tableaux of shape $(w, w, .., w)$. By the Hook formula [44, Theorem 3.10.2]

$$C_w^{(d)} = \frac{\prod_{\ell=1}^{d-1} \ell!}{\prod_{\ell=1}^{d} (w+1)^{d-\ell}} \binom{d \cdot w}{w, w, .., w}$$

$$= \frac{\prod_{\ell=1}^{d-1} \ell! \cdot (d \cdot w)!}{\prod_{\ell=0}^{d-1} (w+\ell)!}.$$

For example for $d = 2$, the number of invariants for weights $w = 0, 1, 2, 3, \ldots$ (and hence for levels $n = 0, 2, 4, 6, \ldots$) are (the Catalan numbers, https://oeis.org/A000108)

$$1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, ..$$

For $d = 3$, the number of invariants for weights $w = 0, 1, 2, 3, \ldots$ (and hence for levels $n = 0, 3, 6, 9, \ldots$) are (the 3-dimensional Catalan numbers, https://oeis.org/A005789)

$$1, 1, 5, 42, 462, 6006, 87516, 1385670, 23371634, 414315330, ..$$

Proof of Theorem 3.7 Write $V = (\mathbb{R}^d)^*$, the dual space of $\mathbb{R}^d$. Every $\phi \in V^{\otimes n}$ that satisfies

$$A\phi = (\det A)^w \phi, \tag{2}$$

clearly spans a one-dimensional irreducible representation of $GL(V)$. Hence we need to investigate all one-dimensional irreducible representation of $GL(V)$ contained in $V^{\otimes n}$ (and it will turn out that all of them satisfy (2)).

The (diagonal) action of $GL(V)$ on $V^{\otimes n}$ is best understood by simultaneously studying the left action of S_n on $V^{\otimes n}$ given by

$$\tau \cdot v_1 \otimes .. \otimes v_n := v_{\tau^{-1}(1)} \otimes .. \otimes v_{\tau^{-1}(n)}.$$

By Schur-Weyl duality, [29, Theorem 6.4.5.2], as $S_n \times GL(V)$ modules,

$$V^{\otimes n} \simeq \bigoplus_{\lambda \vdash n} S^\lambda \otimes V^\lambda, \tag{3}$$

where the sum is over integer partitions λ of n, the S^λ are irreducible representations of S_n, to be detailed below and the V^λ are irreducible representations of $GL(V)$. The exact form of the latter is irrelevant here, we only need to know that V^λ is one-dimensional if and only if $\lambda = (w, .., w)$, d-times, for some integer $w \geq 1$, [10, p. 21]. This gives the condition $n = wd$ in the statement. We assume this to hold from now on.

We are hence left with understanding the unique copy of the "Specht module" S^λ inside of $V^{\otimes n}$. We sketch its classical construction. Let us recall that a **tabloid** is an equivalence class of Young tableaux modulo permutations leaving the set of entries in each row invariant [44, Chap. 2].[6] For t a Young tableau denote $\{t\}$ its tabloid, so for example

$$\left\{\begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & 4 \\ \hline \end{array}\right\} = \left\{\begin{array}{|c|c|} \hline 1 & 3 \\ \hline 4 & 2 \\ \hline \end{array}\right\} = \left\{\begin{array}{|c|c|} \hline 3 & 1 \\ \hline 2 & 4 \\ \hline \end{array}\right\} = \left\{\begin{array}{|c|c|} \hline 3 & 1 \\ \hline 4 & 2 \\ \hline \end{array}\right\}.$$

The symmetric group S_n acts on Young tableaux as

$$(\tau \cdot t)_{ij} := \tau(t_{ij}).$$

For example

$$(234) \cdot \begin{array}{|c|c|} \hline 2 & 4 \\ \hline 1 & 3 \\ \hline \end{array} = \begin{array}{|c|c|} \hline 3 & 2 \\ \hline 1 & 4 \\ \hline \end{array}$$

It then acts on tabloids by $\tau \cdot \{t\} := \{\tau \cdot t\}$. Define for a Young tableau t

$$e_t := \sum_\pi \text{sign}(\pi)\pi \cdot \{t\},$$

where the sum is over all $\pi \in S_n$ that leave the set of values in each *column* invariant. For example with

$$t = \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array}$$

we get

$$e_t = \{t\} - (13) \cdot \{t\} - (24) \cdot \{t\} + (13)(24) \cdot \{t\}.$$

Then

$$\text{Irrep}_{(w,..,w)} := \text{span}\big\{e_t : t \text{ Young tableau of shape } (w, .., w)\big\}$$

is an irreducible representation of S_n and

$$\big\{e_t : t \text{ standard Young tableau of shape } (w, .., w)\big\},$$

forms a basis [44, Theorem 2.5.2]. This concludes the reminder on representation theory for S_n.

Define the map ι from the space of tabloids of shape $(w, .., w)$ into $V^{\otimes n}$ as follows,

$$\iota\big(\{t\}\big) := e_{j_1}^* \otimes .. \otimes e_{j_n}^*,$$

___

[6]One can also think of a tabloid as the following element of the vector space spanned by Young tableaux,

$$\{t\} = \sum_\pi \pi t.$$

Here the sum is over all permutations π that leave the elements of each *row* of t unchanged.

where e_i^* is the canonical basis of V and

$$j_\ell = i \quad \Leftrightarrow \quad \ell \in i\text{-th row of } \{t\}.$$

For example

$$\iota\left(\left\{\begin{array}{|c|c|c|} \hline 1 & 2 & 5 \\ \hline 3 & 4 & 6 \\ \hline \end{array}\right\}\right) = e_1^* \otimes e_1^* \otimes e_2^* \otimes e_2^* \otimes e_1^* \otimes e_2^*.$$

This is a homomorphism of S_n representations. Indeed,

$$\iota\big(\tau \cdot \{t\}\big) = e_{j_1}^* \otimes .. \otimes e_{j_n}^*,$$

with

$$j_\ell = i \quad \Leftrightarrow \quad \ell \in i\text{-th row of } \tau \cdot \{t\}.$$

On the other hand

$$\tau \cdot \iota\big(\{t\}\big) = \tau \cdot e_{r_1}^* \otimes .. \otimes e_{r_n}^*$$
$$= e_{p_1}^* \otimes .. \otimes e_{p_n}^*,$$

with $p_\ell := r_{\tau^{-1}(\ell)}$ and

$$p_\ell = i \quad \Leftrightarrow \quad r_{\tau^{-1}(\ell)} = i$$
$$\Leftrightarrow \quad \tau^{-1}(\ell) \in i\text{-th row of } \{t\}$$
$$\Leftrightarrow \quad \ell \in i\text{-th row of } \tau \cdot \{t\}.$$

So indeed $\iota(\tau \cdot \{t\}) = \tau \cdot \iota(\{t\})$, and ι is a homomorphism of S_n representations. It is a bijection from the space of $(w, .., w)$ tabloids into the space spanned by the vectors

$$e_{i_1}^* \otimes .. \otimes e_{i_n}^* : \#\{\ell : i_\ell = j\} = w, \quad j = 1, .., d.$$

Restricting to $\text{Irrep}_{(w,..,w)}$ then yields an isomorphism of irreducible S_n representations. Hence $\iota(\text{Irrep}_{(w,..,w)})$ is the (unique) realization of S^λ inside of $V^{\otimes n}$ in (3). We finish by describing its image.

Consider the standard Young tableau t_{first} of shape $(w, w, .., w)$ obtained by filling the columns from left to right, i.e.

$$t_{\text{first}} := \begin{array}{|c|c|c|c|c|} \hline 1 & d+1 & .. & .. & .. \\ \hline 2 & d+2 & .. & .. & .. \\ \hline .. & .. & .. & .. & .. \\ \hline d & 2d & .. & .. & n \\ \hline \end{array}$$

Clearly, for any (standard) Young tableau t there exists a unique $\sigma_t \in S_n$ such that

$$\sigma_t \cdot t_{\text{first}} = t.$$

We claim

$$\iota(e_t) = \left(v \mapsto \det[v_{\sigma_t(1)}..v_{\sigma_t(d)}] \cdot ... \cdot \det[v_{\sigma_t((w-1)d+1)}..v_{\sigma_t(n)}]\right).$$

Indeed, since ι is a homomorphism of S_n representation,

$$\iota(\sigma_t \cdot e_{t_{\text{first}}})(v_1, .., v_n) = \iota(e_{t_{\text{first}}})(v_{\sigma_t(1)}, .., v_{\sigma_t(n)})$$

It remains to check

$$\iota(e_{t_{\text{first}}}) = \det[v_1..v_d] \cdot ... \cdot \det[v_{(w-1)d+1}..v_n].$$

Every $\pi \in S^n$ that is column-preserving for t_{first} can be written as the product $\pi_1 \cdot .. \cdot \pi_w$, with π_j ranging over the permutations of the entries of the j-th column t_{first}. Then

$$\begin{aligned}
\iota(e_{t_{\text{first}}})(v_1, .., v_n) &= \sum_{\pi} \operatorname{sign} \pi \ \iota\big(\pi\{t\}\big)(v_1, .., v_n) \\
&= \sum_{\pi_j} \prod_j \operatorname{sign} \pi_j \ \iota\big(\pi_1..\pi_w\{t\}\big)(v_1, .., v_n) \\
&= \sum_{\pi_j} \prod_j \operatorname{sign} \pi_j \ e^*_{\pi_1^{-1}(1)} \otimes .. \otimes e^*_{\pi_1^{-1}(d)} \otimes e^*_{(\pi_2^{-1}(d+1) \bmod d)+1} \otimes \cdots \\
&\qquad \otimes e^*_{(\pi_w^{-1}(n) \bmod d)+1}(v_1, .., v_n) \\
&= \det[v_1..v_d] \cdot ... \cdot \det[v_{(w-1)d+1}..v_n],
\end{aligned}$$

as desired. $\qquad\square$

Applying Lemma 2.4 to Theorem 3.7 we get the invariants in $T(\mathbb{R}^d)$.

Proposition 3.11 *A linear basis for the space of* GL *invariants of order* $n = wd$ *is given by*

$$\sum_{i_1,...,i_n \in \{1,...,d\}} g_\Sigma(e_{i_1}, e_{i_2}, \ldots, e_{i_n}) x_{i_1} x_{i_2} \ldots x_{i_n},$$

where

$$g_\Sigma(v) = \det[v_{C_1}].. \det[v_{C_w}],$$

where C_i *are the columns of* Σ, Σ *ranges over all standard Young tableaux corresponding to the partition* $\lambda = \underbrace{(w, w, .., w)}_{d \text{ times}}$ *of* n, *and the notation* v_C *is as introduced in Theorem* 3.7.

Remark 3.12 By Lemma 3.5, for any invariant $\phi \in T(\mathbb{R}^d)$ and $n \geq 1$ we have that $\pi_n \phi$ is also invariant. Hence the previous theorem characterizes *all* invariants we are interested in (Definition 3.1), not just homogeneous ones.

Remark 3.13 Note that each of these invariants ϕ consists only of monomials that contain *every* variable $x_1, \ldots, x_d$ at least once. This implies that $\langle S(X)_{0,T}, \phi \rangle$ consists only of iterated integrals that contain every component $X^1, \ldots, X^d$ of the curve at least once. Hence, if at least one of these components is constant, the whole expression will be zero.

Since ϕ is invariant, this implies that $\langle S(X)_{0,T}, \phi \rangle = 0$ as soon as there is some coordinate transformation under which one component is constant, that is whenever the curve X stays in a hyperplane of dimension strictly less then d.

One of the simplest curves in d dimensions that does *not* lie in any hyperplane of lower dimension is the *moment curve*

$$t \mapsto \left(t, t^2, .., t^d \right).$$

We will come back to this example in Lemma 3.29.

3.1 Examples

We will use the following short notation:

$$\mathtt{i}_1 \ldots \mathtt{i}_\mathtt{n} := x_{i_1} x_{i_2} .. x_{i_n},$$

so, for example

$$\mathtt{1121} := x_1 x_1 x_2 x_1.$$

We present the invariants described in Sect. 2 for some special cases of d and w.

The case $d = 2$

Level 2 ($w = 1$)

$$\mathtt{12} - \mathtt{21}$$

Remark 3.14 Let us make clear that from the perspective of data analysis, the "invariant" of interest is really the action of this element in $T(\mathbb{R}^d)$ on the signature of a curve.

In this example, the real number

$$\left\langle S(X)_{0,T}, \mathtt{12} - \mathtt{21} \right\rangle = \int_0^T \int^{r_2} dX^1_{r_1} dX^2_{r_2} - \int_0^T \int^{r_2} dX^1_{r_1} dX^2_{r_2},$$

changes only by the determinant of $A \in \mathrm{GL}(\mathbb{R}^2)$ when calculating it for the transformed curve AX:

$$\left\langle S(AX)_{0,T}, \mathtt{12} - \mathtt{21} \right\rangle = \det(A) \left\langle S(X)_{0,T}, \mathtt{12} - \mathtt{21} \right\rangle.$$

Level 4 ($w = 2$)

$$\mathtt{1212} - \mathtt{1221} - \mathtt{2112} + \mathtt{2121}$$

$$\mathtt{1122} - \mathtt{1221} - \mathtt{2112} + \mathtt{2211}$$

Remark 3.15 This is a *linear* basis of invariants in the fourth level. If one takes *algebraic* dependencies into consideration, the set of invariants becomes smaller. To be specific, assume that one already has knowledge of the invariant of level 2 (i.e. $\langle S(X)_{0,T}, 12 - 21 \rangle$). If, say in a machine learning application, the learning algorithm can deal sufficiently well with nonlinearities, one should not be required to provide additionally the square of this number. In other words $|\langle S(X)_{0,T}, 12 - 21 \rangle|^2$ can also be assumed to be "known". But, by the shuffle identity (Lemma 2.1), this can be written as

$$
\begin{aligned}
\left| \langle S(X)_{0,T}, 12 - 21 \rangle \right|^2 &= \langle S(X)_{0,T}, 12 - 21 \rangle \cdot \langle S(X)_{0,T}, 12 - 21 \rangle \\
&= \langle S(X)_{0,T}, (12 - 21) \shuffle (12 - 21) \rangle \\
&= \langle S(X)_{0,T}, 4 \cdot 1122 - 4 \cdot 1221 - 4 \cdot 2112 + 4 \cdot 2211 \rangle.
\end{aligned}
$$

Now, seeing that $4 \cdot 1122 - 4 \cdot 1221 - 4 \cdot 2112 + 4 \cdot 2211$ is invariant, there is only one "new" independent invariant in the fourth level, namely $1212 - 1221 - 2112 + 2121$.

A similar analysis can also be carried out for the following invariants, but we refrain from doing so, since it can be easily done with a computer algebra system.

Level 6 ($w = 3$)

$121212 - 121221 - 122112 + 122121 - 211212 + 211221 + 212112 - 212121$
$112212 - 112221 - 122112 + 122121 - 211212 + 211221 + 221112 - 221121$
$121122 - 121221 - 122112 + 122211 - 211122 + 211221 + 212112 - 212211$
$112122 - 112221 - 122112 + 122211 - 211122 + 211221 + 221112 - 221211$
$111222 - 112221 - 121212 + 122211 - 211122 + 212121 + 221112 - 222111$

The case $d = 3$

Level $n = 3$ ($w = 1$)

$$123 - 132 - 213 + 231 + 312 - 321$$

Level $n = 6$ ($w = 2$)

$123123 - 312132 + 312312 + 213132 - 213231 - 213123 + 321213 - 312321 - 132231$
$- 132123 - 321231 + 321132 + 132321 + 132213 + 231231 + 321321 + 213321$
$+ 123231 + 231123 - 312213 - 321123 - 231132 + 213213 + 132132 + 312231$
$- 213312 - 231321 - 132312 - 123213 - 321312 + 312123 - 231213 + 231312$
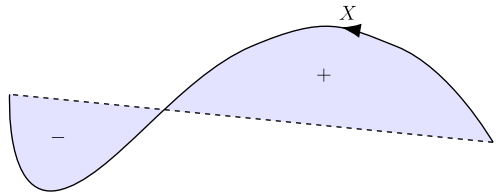$- 123321 + 123312 - 123132$

$+$ 4 more

The case $d = 4$

Level $n = 4$ ($w = 1$)

$$
\begin{aligned}
&1234 - 1243 - 1324 + 1342 + 1423 - 1432 - 2134 + 2143 \\
&+ 2314 - 2341 - 2413 + 2431 + 3124 - 3142 - 3214 + 3241 \\
&+ 3412 - 3421 - 4123 + 4132 + 4213 - 4231 - 4312 + 4321
\end{aligned}
$$

Fig. 1 A curve $X = (X^1, X^2)$ is shown, with shaded area given by
$\frac{1}{2}\langle S(X)_{0,T}, x_1 x_2 - x_2 x_1 \rangle =$
$\frac{1}{2}\int_0^T \int_0^{r_2} dX_{r_1}^1 dX_{r_2}^2 -$
$\frac{1}{2}\int_0^T \int_0^{r_2} dX_{r_1}^2 dX_{r_2}^1$

3.2 The Invariant of Weight One, in Dimension Two

Geometric Interpretation The invariant for $d = 2$, $w = 1$, namely $\phi = x_1 x_2 - x_2 x_1$ has a simple geometric interpretation: it picks out (two times)[7] the area (signed, and with multiplicity) between the curve X and the cord spanned between its starting and endpoint (compare Fig. 1). For (smooth) non-intersecting curves, this follows from Green's theorem [43, Theorem 10.33]. For self-intersecting curves, the mathematically most convenient definition of "signed area" is the integral (in the plane) of its winding number. The claimed relation to the invariant ϕ is for example proven in [34, Proposition 1].

Connection to Correlation Assume that X is a continuous curve, piecewise linear between some time points t_i, $i = 0, \ldots, n$.[8] The area is then explicitly calculated as

$$\int_0^T \int_0^r dX_u^1 dX_r^2 - \int_0^T \int_0^r dX_u^2 dX_r^1$$

$$= \int_0^T (X_r^1 - X_0^1) dX_r^2 - \int_0^T (X_r^2 - X_0^2) dX_r^1$$

$$= \sum_{i=0}^{n-1} \int_0^1 (X_{t_i}^1 + t(X_{t_{i+1}}^1 - X_{t_i}^1) - X_0^1)(X_{t_{i+1}}^2 - X_{t_i}^2) dt$$

$$\quad - \sum_{i=0}^{n-1} \int_0^1 (X_{t_i}^2 + t(X_{t_{i+1}}^2 - X_{t_i}^2) - X_0^2)(X_{t_{i+1}}^1 - X_{t_i}^1) dt$$

$$= \sum_{i=0}^{n-1} X_{t_i}^1 X_{t_{i+1}}^2 - \sum_{i=0}^{n-1} X_{t_i}^2 X_{t_{i+1}}^1 + X_0^2(X_{t_n}^1 - X_{t_0}^1) - X_0^1(X_{t_n}^2 - X_{t_0}^2)$$

$$= \mathrm{Corr}(X^2, X^1)_1 - \mathrm{Corr}(X^1, X^2)_1 + X_0^2(X_{t_n}^1 - X_{t_0}^1) - X_0^1(X_{t_n}^2 - X_{t_0}^2).$$

Here, for two vectors a, b of length n

$$\mathrm{Corr}(a, b)_1 := \sum_{i=0}^{n-1} a_{i+1} b_i,$$

---

[7]The prefactor $1/2$ is irrelevant, so we will speak of ϕ and also of $\frac{1}{2}\phi$ as picking out the area.

[8]The standard example is a time series that is discretely observed at times t_i and linearly interpolated in between.

the lag-one cross-correlation, which is a commonly used feature in signal analysis, see for example [38, Chap. 13.2].[9] In particular, if the curve starts at 0, we have

$$\int_0^T \int_0^r dX_u^1 dX_r^2 - \int_0^T \int_0^r dX_u^2 dX_r^1 = \mathrm{Corr}(X^2, X^1)_1 - \mathrm{Corr}(X^1, X^2)_1,$$

which is an antisymmetrized version of the lag-one cross-correlation.

Remark 3.16 The antisymmetrized version of the lag τ cross-correlation, for each $\tau \geq 2$, is also a $\mathrm{GL}(\mathbb{R}^2)$ invariant of the curve. In general these invariants cannot be found in the signature, and we thank the anonymous referee for pointing out the following example. Consider the treelike curve which linearly interpolates the following points

$$(0, 0), (1, 0), (1, 1/2), (1, 1), (1, 0), (0, 0).$$

Its signature is trivial, but

$$\mathrm{Corr}(X^1, X^2)_2 - \mathrm{Corr}(X^2, X^1)_2 = 1 - 0.5 = 0.5.$$

3.3 The Invariant of Weight One, in Any Dimension

Whatever the dimension d of the curve's ambient space, the space of invariants of weight 1 has dimension 1 and is spanned by

$$\mathrm{Inv}_d := \mathrm{Inv}_d(x_1, .., x_d) := \sum_{\sigma \in S_d} \mathrm{sign}(\sigma)\, x_{\sigma(1)}..x_{\sigma(d)} = \det \begin{pmatrix} x_1 & .. & x_d \\ .. & .. & .. \\ x_1 & .. & x_d \end{pmatrix}. \qquad (4)$$

Here, for a matrix C of non-commuting variables, (compare [14, Definition 3.1])

$$\det C := \sum_\tau \mathrm{sign}\,\tau \prod_i C_{i\tau(i)}.$$

This invariant is of homogeneity d. The following lemma tells us that we can write Inv_d in terms of expressions on lower homogeneities.

To state it, we first define the operation $\mathsf{InsertAfter}(x_i, r)$ on monomials of order $n \geq r$, as the insertion of the variable x_i after position r, and extend it linearly. For example

$$\mathsf{InsertAfter}(x_1, 1)\, \mathrm{Inv}_2(x_2, x_3) = \mathsf{InsertAfter}(x_1, 1)(x_2 x_3 - x_3 x_2)$$

$$= x_2 x_1 x_3 - x_3 x_1 x_2.$$

Lemma 3.17 *In any dimension d and for any $r = 0, 1, .., d - 1$*

$$\mathrm{Inv}_d(x_1, .., x_d) = (-1)^r \sum_{j=1}^d (-1)^{j+1} \mathsf{InsertAfter}(x_j, r)\, \mathrm{Inv}_{d-1}(x_1, .., \widehat{x_j}.., x_d),$$

where $\widehat{x_j}$ denotes the omission of that argument.

---

[9] Note the nomenclature used in signal analysis. A probabilist or statistician would tend to call this a "covariance" and not a "correlation".

For d odd,

$$\mathrm{Inv}_d(x_1, .., x_d) = \sum_{j=1}^{d} (-1)^{j+1} x_j \shuffle \mathrm{Inv}_{d-1}(x_1, .., \widehat{x_j}.., x_d).$$

Remark 3.18 For completeness, we also note the related *de Bruijn's formula*. For *d* even,

$$\mathrm{Inv}_d(x_1, .., x_d) = \mathrm{Pf}_{\shuffle}[A],$$

where

$$A_{ij} = \mathrm{Inv}_2(x_i, x_j),$$

and the Pfaffian (with respect to the shuffle product), is

$$\mathrm{Pf}_{\shuffle}[A] = \frac{1}{2^{d/2}(d/2)!} \sum_{\sigma \in S_d} \mathrm{sign}(\sigma) A_{\sigma(1),\sigma(2)} \shuffle A_{\sigma(3),\sigma(4)} \shuffle .. \shuffle A_{\sigma(d-1),\sigma(d)}.$$

For a proof see [7] and [33].

Proof The first statement follows from expressing the determinant in (4) in terms of minors with respect to the row $r + 1$ (since the x_i are non-commuting, this does not work with columns!).

Regarding the second statement, since d is odd and then using the first statement

$$\mathrm{Inv}_d = \sum_{r=0}^{d-1} (-1)^r \mathrm{Inv}_d$$

$$= \sum_{r=0}^{d-1} (-1)^r (-1)^r \sum_{j=1}^{d} (-1)^{j+1} \mathsf{InsertAfter}(x_j, r) \mathrm{Inv}_{d-1}(x_1, .., \widehat{x_j}.., x_d)$$

$$= \sum_{j=1}^{d} (-1)^{j+1} \sum_{r=0}^{d-1} \mathsf{InsertAfter}(x_j, r) \mathrm{Inv}_{d-1}(x_1, .., \widehat{x_j}.., x_d)$$

$$= \sum_{j=1}^{d} (-1)^{j+1} x_j \shuffle \mathrm{Inv}_{d-1}(x_1, .., \widehat{x_j}.., x_d),$$

as claimed. $\qquad\square$

An immediate consequence is the following lemma.

Lemma 3.19 *If the ambient dimension d is odd and the curve X is closed (i.e. $X_T = X_0$) then*

$$\langle S(X)_{0,T}, \mathrm{Inv}_d \rangle = 0.$$

Proof By Lemma 3.17 and then by the shuffle identity (Lemma 2.1)

$$\langle S(X)_{0,T}, \mathrm{Inv}_d \rangle = \sum_{j=1}^{d} \langle S(X)_{0,T}, (-1)^{j+1} x_j \shuffle \mathrm{Inv}_{d-1}(x_1, .., \widehat{x_j}.., x_d) \rangle$$

$$= \sum_{j=1}^{d} (-1)^{j+1} \langle S(X)_{0,T}, x_j \rangle \langle S(X)_{0,T}, \mathrm{Inv}_{d-1}(x_1, .., \widehat{x_j} .., x_d) \rangle$$

$$= 0,$$

since the increment $\langle S(X)_{0,T}, x_j \rangle = X_T^j - X_0^j$ is zero for all j by assumption.   □

In even dimension we have the phenomenon that closing a curve does not change the value of the invariant.

Lemma 3.20 *If the ambient dimension d is even, then for any curve X*

$$\langle S(X)_{0,T}, \mathrm{Inv}_d \rangle = \langle S(\bar{X})_{0,T}, \mathrm{Inv}_d \rangle,$$

where $\bar{X}$ is X concatenated with the straight line connecting X_T to X_0.

Proof Let $\bar{X}$ be parametrized on $[0, 2T]$ as follows: $\bar{X} = X$ on $[0, T]$ and it is the linear path connecting X_T to X_0 on $[T, 2T]$. By translation invariance we can assume $X_0 = 0$ and by $\mathrm{GL}(\mathbb{R}^d)$-invariance that X_T lies on the x_1 axis. Then the only component of $\bar{X}$ that is non-constant on $[T, 2T]$ is the first one, $\bar{X}^1$.

By Lemma 3.17

$$\mathrm{Inv}_d = - \sum_{j=1}^{d} (-1)^{j+1} \mathrm{Inv}_{d-1}(x_1, .., \hat{x}_j, ..x_d) x_j.$$

Letting the summands act on $S(\bar{X})_{0,t}$ we get ± 1 times

$$\int_0^t \langle S(\bar{X})_{0,r}, \mathrm{Inv}_{d-1}(x_1, .., \hat{x}_j, ..x_d) \rangle d\bar{X}_r^j.$$

For $j \neq 1$ these expressions are constant on $[T, 2T]$, since we arranged things so that those $\bar{X}^j$ do not move on $[T, 2T]$. But also for $j = 1$ this expression is constant on $[T, 2T]$. Indeed, the integrand

$$\langle S(\bar{X})_{0,r}, \mathrm{Inv}_{d-1}(x_2, x_3, .., x_d) \rangle,$$

is zero on $[T, 2T]$, since X, projected on the $x_2 - .. - x_d$ hyperplane, is a closed curve, and so Lemma 3.19 applies.   □

Lemma 3.21 *Let X be the piecewise linear curve through $p_0, .., p_d \in \mathbb{R}^d$. Then*

$$\langle S(X)_{0,T}, \mathrm{Inv}_d \rangle = \det \begin{bmatrix} 1 & 1 & .. & 1 \\ p_0 & p_1 & .. & p_d \end{bmatrix}$$

Proof First, for any $v \in \mathbb{R}^d$,

$$\det \begin{bmatrix} 1 & 1 & .. & 1 \\ p_0 + v & p_1 + v & .. & p_d + v \end{bmatrix} = \det \begin{bmatrix} 1 & 1 & .. & 1 \\ p_0 & p_1 & .. & p_d \end{bmatrix}.$$

Since the signature is also invariant to translation, we can assume $p_0 = 0$. Now both sides of the statement transform the same way under the action of $\mathrm{GL}(\mathbb{R}^d)$ on the points $p_1, .., p_d$. It is then enough to prove this for

$$p_0 = 0$$

$$p_1 = e_1$$

$$p_2 = e_1 + e_2$$

$$..$$

$$p_d = e_1 + .. + e_d.$$

Now, for this particular choice of points the right hand side is clearly equal to 1. For the left hand side, the only non-zero term is

$$\langle S(X)_{0,T}, 12..d \rangle = \int dX^1 .. dX^d$$

$$= 1.$$ $\qquad\square$

The modulus of the determinant

$$\det \begin{bmatrix} 1 & 1 & .. & 1 \\ 0 & p_1 & .. & p_d \end{bmatrix} = \det \begin{bmatrix} p_1 & .. & p_d \end{bmatrix}$$

gives the Lebesgue measure of the parallelepiped spanned by the vectors $p_1 - p_0, .., p_d - p_0$. The polytope spanned by the points $p_0, p_1, .., p_d$ fits $d!$ times into that parallelepiped. We hence have the relation to classical volume as follows.

Lemma 3.22 *Let $p_0, .., p_d \in \mathbb{R}^d$, then*

$$\left| \text{Convex-Hull}(p_0, .., p_d) \right| = \frac{1}{d!} \left| \det \begin{bmatrix} 1 & 1 & .. & 1 \\ p_0 & p_1 & .. & p_d \end{bmatrix} \right|$$

We now proceed to piecewise linear curves with more than d vertices.

Lemma 3.23 *Let X be the piecewise linear curve through, $p_0, .., p_n \in \mathbb{R}^d$, with $n \geq d$. Then,*

$$\langle S(X)_{0,T}, \mathrm{Inv}_d \rangle = \sum_i \det \begin{bmatrix} 1 & 1 & .. & 1 \\ p_{i_0} & p_{i_1} & .. & p_{i_d} \end{bmatrix}. \tag{5}$$

For d even, the subsequences i are chosen as follows:

$$i_0 = 0$$

and $i_1, .., i_d$ ranges over all possible increasing subsequences of $1, 2, .., n$ such that for ℓ odd: $i_\ell + 1 = i_{\ell+1}$.

For d odd, they are chosen as follows:

$$i_0 = 0$$

$$i_d = n,$$

and $i_1, .., i_{d-1}$ ranges over all possible increasing subsequences of $1, 2, .., n-1$ such that for ℓ odd: $i_\ell + 1 = i_{\ell+1}$

Remark 3.24 The number of indices is easily calculated. In the even case, we have $B := d/2$ "groups of two" to place, $A := n - d$ "fillers" in between. This gives

$$\binom{A+B}{B} = \binom{n-d+d/2}{d/2} = \binom{\lfloor \frac{d}{2} \rfloor + n - d}{\lfloor \frac{d}{2} \rfloor},$$

where $\lfloor r \rfloor$ is the largest integer less than or equal to r.

In the odd case, we have $B := (d-1)/2$ "groups of two" to place, with $A := n - 1 - (d-1)$ "fillers" in between. This gives

$$\binom{A+B}{B} = \binom{n-1-\frac{d-1}{2}}{\frac{d-1}{2}} = \binom{\lfloor \frac{d}{2} \rfloor + n - d}{\lfloor \frac{d}{2} \rfloor}.$$

Remark 3.25 Consider the case $d = 2$, and a curve X through the points $p_0, p_1, .., p_n \in \mathbb{R}^d$, with $p_0 = 0$. Then

$$\langle S(X)_T, \mathrm{Inv}_2 \rangle = \langle S(X)_T, 12 - 21 \rangle$$

$$= \sum_{i=1}^{n-1} \det \begin{bmatrix} 1 & 1 & 1 \\ p_0 & p_i & p_{i+1} \end{bmatrix}$$

$$= \sum_{i=1}^{n-1} \det \begin{bmatrix} p_i & p_{i+1} \end{bmatrix}$$

$$=: \sum_{i=1}^{n-1} P_{i,i+1}.$$

We can express Inv_d as a linear combination of the 2×2 minors $P_{i,j}$ of the $2 \times n$ matrix $(p_1, p_2, .., p_n)$. Generally, it is well-known that *all* invariants to $\mathrm{GL}(\mathbb{R}^2)$ of a tuple of points are expressible in terms of these minors [47, Sect. 3.2]. So, for a piecewise linear curve through 0, $p_1, .., p_n$, *all* our integral invariants are—a fortiori—expressible in terms of them. In the simple case shown here, this expression is just a linear combination. Experimentally, for higher order invariants, polynomial combinations appear with a lot of structure. This poses the question on whether one can set up some kind of "GL invariant integration", where, instead of the classical Riemann integration that uses increments, one "integrates" using only these $P_{i,j}$.

Example 3.26 For $d = 2$, $n = 5$ we get the subsequences

$$[0, 1, 2]$$

$$[0, 2, 3]$$

$$[0, 3, 4]$$

For $d = 4$, $n = 7$ we get the subsequences

$$[0, 1, 2, 3, 4]$$
$$[0, 1, 2, 4, 5]$$
$$[0, 1, 2, 5, 6]$$
$$[0, 2, 3, 4, 5]$$
$$[0, 2, 3, 5, 6]$$
$$[0, 3, 4, 5, 6]$$

For $d = 5$, $n = 8$ we get the subsequences

$$[0, 1, 2, 3, 4, 7]$$
$$[0, 1, 2, 4, 5, 7]$$
$$[0, 1, 2, 5, 6, 7]$$
$$[0, 2, 3, 4, 5, 7]$$
$$[0, 2, 3, 5, 6, 7]$$
$$[0, 3, 4, 5, 6, 7]$$

Proof of Lemma 3.23 **The case $d = 2$**
Let X be the curve through the points p_0, p_1, .., p_n. We can write it as concatenation of the curves $X^{(i)}$, where $X^{(i)}$ is the curve through the points p_0, p_i, p_{i+1}, p_0. The time-interval of definition for these curves (and all curves in this proof) do not matter, so we omit the subscript of $S(.)$. Then, by Chen's lemma (Lemma 2.3)

$$\langle S(X), 12 - 21 \rangle = \langle S(X^{(n-1)}) \cdot ... \cdot S(X^{(1)}), 12 - 21 \rangle$$
$$= \sum_{i=1}^{n-1} \langle S(X^{(i)}), 12 - 21 \rangle.$$

For the last equality we used that

$$\langle gh, 12 - 21 \rangle = \langle g, 12 - 21 \rangle + \langle h, 12 - 21 \rangle + \langle g, 1 \rangle \langle h, 2 \rangle - \langle g, 2 \rangle \langle h, 1 \rangle,$$

and that the increments of all curves $X^{(i)}$ are zero. Now by Lemma 3.20 we can omit the last straight line in every $X^{(i)}$ and hence by Lemma 3.21

$$\langle S(X^{(i)}), 12 - 21 \rangle = \det \begin{bmatrix} 1 & 1 & 1 \\ p_0 & p_i & p_{i+1} \end{bmatrix},$$

which finishes the proof for $d = 2$.

Now assume the statement is true for all dimensions strictly smaller than some d. We show it is true for d.

d is odd

As before we can assume $p_0 = 0$ and that p_n lies on the x_1 axis. Every sequence summed over on the right-hand side of (5) is of the form $i = (0, \ldots, n)$. For each of those, we calculate

$$
\det \begin{bmatrix} 1 & 1 & .. & 1 & 1 \\ p_{i_0} & p_{i_1} & .. & p_{i_{d-1}} & p_{i_d} \end{bmatrix} = \det \begin{bmatrix} 1 & 1 & .. & 1 & 1 \\ 0 & p_{i_1} & .. & p_{i_{d-1}} & \Delta \cdot e_1 \end{bmatrix}
$$

$$
= \Delta \cdot \det \begin{bmatrix} 1 & 1 & .. & 1 \\ 0 & \bar{p}_{i_1} & .. & \bar{p}_{i_{d-1}} \end{bmatrix}.
$$

Here $\bar{p}_j \in \mathbb{R}^{d-1}$ is obtained by deleting the first coordinate of p_j, e_1 is the first canonical coordinate vector in $\mathbb{R}^d$ and $\Delta := (p_0 - p_n)_1 = \langle S(X), x_1 \rangle$ is the total increment of X in the x_1 direction. Here we used that d is odd (otherwise we would get a prefactor -1).

The last determinant is the expression for the summands of the right-hand side of (5), but with dimension $d-1$ and points $0 = \bar{p}_0, \bar{p}_1, .., \bar{p}_{n-1}$. By assumption, summing up all these determinants gives

$$
\Delta \cdot \langle S(\bar{X}), \mathrm{Inv}_{d-1} \rangle = \langle S(X), x_1 \rangle \langle S(\bar{X}), \mathrm{Inv}_{d-1} \rangle,
$$

where $\bar{X}$ is the curve in $\mathbb{R}^{d-1}$ through the points $\bar{p}_0, .., \bar{p}_{n-1}$. Since $\bar{p}_n = \bar{p}_0 = 0$, we can attach the additional point $\bar{p}_n$ to $\bar{X}$ without changing the value here (Lemma 3.20). Hence the sum of determinants is equal to

$$
\langle S(X), x_1 \rangle \langle S(X), \mathrm{Inv}_{d-1}(x_2, .., x_d) \rangle.
$$

Since we arranged matters such that $\langle S(X), x_i \rangle = 0$ for $i \neq 1$, this is equal to

$$
\sum_{i=1}^{d} \langle S(X), x_i \rangle \langle S(X), \mathrm{Inv}_{d-1}(x_1, x_2, .., \hat{x}_i, .., x_d) \rangle
$$

$$
= \left\langle S(X), \sum_{i=1}^{d} x_i \shuffle \mathrm{Inv}_{d-1}(x_1, x_2, .., \hat{x}_i, .., x_d) \right\rangle,
$$

where we used the shuffle identity, Lemma 2.1. By the second part of Lemma 3.17 this is equal to $\langle S(X), \mathrm{Inv}_d \rangle$, which finishes the proof for odd d.

d is even

We proceed by induction on n. For $n = d$ the statement follows from Lemma 3.21.

Let it be true for some n, we show it for a piecewise linear curve through some points $p_0, .., p_{n+1}$. Write $X = X' \sqcup X''$ where X' is the linear interpolation of $p_0, .., p_n$, X'' is the linear path from p_n to p_{n+1} and we recall concatenation $\sqcup$ of paths from Lemma 2.3. By assumption, (5) is true for the curve X'. Adding an additional point p_{n+1}, the sum on the right hand side of (5) gets additional indices of the form

$$
(p_{j_0}, .., p_{j_{d-1}}, p_{n+1}),
$$

where

$$j_0 = 0$$

$$j_{d-1} = n,$$

and where $j_1, .., j_{d-2}$ ranges over all possible increasing subsequences of $1, 2, .., n-1$ such that for ℓ odd $j_\ell + 1 = j_{\ell+1}$.

Assume $p_{n+1} - p_n = \Delta \cdot e_1$ lies on the x_1-axis. Then, summing over those j,

$$\sum_j \det \begin{bmatrix} 1 & 1 & .. & 1 & 1 & 1 \\ 0 & p_{j_1} & .. & p_{j_{d-2}} & p_n & p_{n+1} \end{bmatrix}$$

$$= \sum_j \det \begin{bmatrix} 1 & 1 & .. & 1 & 1 & 1 \\ -p_n & p_{j_1} - p_n & .. & p_{j_{d-2}} - p_n & 0 & p_{n+1} - p_n \end{bmatrix}$$

$$= \sum_j \det \begin{bmatrix} 1 & 1 & .. & 1 & 1 & 1 \\ -p_n & p_{j_1} - p_n & .. & p_{j_{d-2}} - p_n & 0 & \Delta \cdot e_1 \end{bmatrix}$$

$$= -\Delta \cdot \sum_j \det \begin{bmatrix} 1 & 1 & .. & 1 & 1 \\ -\bar{p}_n & \bar{p}_{j_1} - \bar{p}_n & .. & \bar{p}_{j_{d-2}} - \bar{p}_n & 0 \end{bmatrix}$$

$$= -\Delta \cdot \sum_j \det \begin{bmatrix} 1 & 1 & .. & 1 & 1 \\ 0 & \bar{p}_{j_1} & .. & \bar{p}_{j_{d-2}} & \bar{p}_n \end{bmatrix}$$

$$= -\Delta \cdot \langle S(\bar{X}'), \mathrm{Inv}_{d-1} \rangle$$

$$= -\Delta \cdot \langle S(X'), \mathrm{Inv}_{d-1}(x_2, .., x_d) \rangle$$

Here $\bar{X}'$ is the curve in $\mathbb{R}^{d-1}$ through the points $\bar{p}_0, .., \bar{p}_n$, and we used the fact that the indices j here range over the ones used for (5) in dimension $d - 1$ on the points $\bar{p}_0, .., \bar{p}_n$. On the other hand,

$$\langle S(X), \mathrm{Inv}_d \rangle = \langle S(X')S(X''), \mathrm{Inv}_d \rangle$$

$$= \langle S(X'), \mathrm{Inv}_d \rangle - \langle S(X'), \mathrm{Inv}_{d-1}(x_2, .., x_d) \rangle \langle S(X''), x_1 \rangle$$

Here we used that $S(X'') = \exp(\Delta \cdot x_1) = 1 + \Delta \cdot x_1 + O(x_1^2)$ [16, Example 7.21], the fact that each monomial in Inv_d has exactly one occurrence of x_1 and Lemma 3.17. This finishes the proof. $\qquad\square$

Definition 3.27 Let $X : [0, T] \to \mathbb{R}^d$ be any curve. Define its **signed volume** to be the following limit, if it exists,

$$\text{Signed-Volume}(X) := \frac{1}{d!} \lim_{|\pi| \to 0} \sum_i \det \begin{bmatrix} 1 & 1 & .. & 1 \\ X_{t_{i_0}^\pi} & X_{t_{i_1}^\pi} & .. & X_{t_{i_d}^\pi} \end{bmatrix}.$$

Here $\pi = (0 = t_0^\pi, .., t_{n^\pi}^\pi = T)$ is a partition of the interval $[0, T]$ and $|\pi|$ denotes its mesh size. The indices i are chosen as in Lemma 3.23.

Theorem 3.28 Let $X : [0, T] \to \mathbb{R}^d$ a continuous curve of bounded variation. Then its signed volume exists and

$$\text{Signed-Volume}(X) = \frac{1}{d!} \langle S(X)_{0,T}, \mathrm{Inv}_d \rangle$$

Proof Fix some sequence $\{\pi^n\}_{n\in\mathbb{N}}$, of partitions of $[0, T]$ with $|\pi^n| \to 0$ and interpolate X linearly along each π^n to obtain a sequence of linearly interpolated curves X^n. Then by Lemma 3.23

$$\text{Signed-Volume}(X^n) = \frac{1}{d!}\langle S(X^n)_{0,T}, \text{Inv}_d\rangle$$

By stability of the signature in the class of continuous curves of bounded variation [16, Proposition 1.28, Proposition 2.7], we get convergence

$$\langle S(X^n)_{0,T}, \text{Inv}_d\rangle \to \langle S(X)_{0,T}, \text{Inv}_d\rangle$$

and this is independent of the particular sequence π^n chosen. $\square$

The previous theorem is almost a tautology, but there are relations to classical objects in geometry. For $d = 2$, as we have seen in Sect. 3.2,

$$\frac{1}{2}\langle S(X)_{0,T}, \text{Inv}_2\rangle,$$

is equal to the signed area of the curve X. In general dimension, the value of the invariant is related to some kind of classical "volume" if the curve satisfies some kind of monotonicity. This is in particular satisfied for the "moment curve".

Lemma 3.29 *Let X be the moment curve*

$$X_t = (t, t^2, \ldots, t^d) \in \mathbb{R}^d.$$

Then for any $T > 0$

$$\frac{1}{d!}\langle S(X)_{0,T}, \text{Inv}_d\rangle = |\text{Convex-Hull}(X_{[0,T]})|$$

Remark 3.30 It is easily verified that for integers $n_1..n_d$ one has

$$\frac{1}{n_1 \cdot \ldots \cdot n_d}\int_0^T dt_1^{n_1}..dt_d^{n_d} = \frac{1}{n_1}\frac{1}{n_1+n_2}..\frac{1}{n_1+..+n_d}T^{n_1+..+n_d}.$$

We deduce that

$$|\text{Convex-Hull}(X_{[0,T]})| = T^{1+2+..+d}\sum_{\sigma\in S_d}\text{sign}\,\sigma\,\frac{1}{\sigma(1)}\frac{1}{\sigma(1)+\sigma(2)}..\frac{1}{\sigma(1)+..+\sigma(d)}.$$

In [24, Sect. 15], the value of this volume is determined, for $T = 1$, as

$$\prod_{\ell=1}^d \frac{(\ell-1)!(\ell-1)!}{(2\ell-1)!}.$$

We hence get the combinatorial identity

$$\prod_{\ell=1}^d \frac{(\ell-1)!(\ell-1)!}{(2\ell-1)!} = \sum_{\sigma\in S_d}\text{sign}\,\sigma\,\frac{1}{\sigma(1)}\frac{1}{\sigma(1)+\sigma(2)}..\frac{1}{\sigma(1)+..+\sigma(d)}.$$

Proof For $n \geq d$ let $0 = t_0 < .. < t_n \leq T$ be time-points, let $p_i := X_{t_i}$ be the corresponding points on the moment curve and denote by X^n the piecewise linear curve through those points. We will show

$$\frac{1}{d!}\langle S(X^n)_{0,T}, \text{Inv}_d \rangle = \left| \text{Convex-Hull}\left(X^n_{[0,T]}\right) \right|.$$

First note that for any $1 \leq i_0 < i_1 < .. \leq i_d \leq n$,

$$\det \begin{bmatrix} 1 & 1 & .. & 1 \\ p_{i_0} & p_{i_1} & .. & p_{i_d} \end{bmatrix} = \prod_{0 \leq \ell < k \leq n} (t_{i_k} - t_{i_\ell}) > 0, \tag{6}$$

since it is a Vandermonde determinant.

We will decompose $P := \{p_0, .., p_n\}$ into (overlapping) sets S_ℓ with cardinality $d + 1$ and such that[10]

$$\left| \text{Convex-Hull}(p_0, .., p_n) \right| = \sum_\ell \left| \text{Convex-Hull}(S_\ell) \right|.$$

A *face* of P is a subset $F \subset P$ such that its convex hull Convex-Hull(F) equals the intersection of Convex-Hull(P) with some affine hyperspace. A face is a *facet*, if its affine span has dimension $d - 1$. The following is a fact that is true for any polytope spanned by some points P: up to a set of measure zero, for every point x in Convex-Hull(P), the line connecting p_0 to x exits Convex-Hull($p_0, .., p_n$) through a unique facet of Convex-Hull($p_0, .., p_n$) contained in $\{p_1, .., p_n\}$. Hence

$$\left| \text{Convex-Hull}(p_0, .., p_n) \right| = \sum_F \left| \text{Convex-Hull}(p_0 \cup F) \right|,$$

where the sum is over all such facets.

Our points p_i lie on the moment curve. Then, by (6), any collection of points $p_{i_0}, p_{i_1}, .., p_{i_d}$ is in general position. This means that every facet of P must have exactly d points (and not more). Facets of Convex-Hull(P) with d points are characterized by Gale's criterion ([17, Theorem 3], [53, Theorem 0.7]):

the points $p_{i_1}, .., p_{i_d}$, with distinct $i_j \in \{0, .., n\}$ form a facet of P if and only if any two elements of $\{0, .., n\} \setminus \{i_1, .., i_d\}$ are separated by an even number of elements in $\{i_1, .., i_d\}$.[11]

d odd

We are looking for such $\{i_j\}$ such that $i_1 \geq 1$. Those are exactly the indices with

- $i_{\ell+1} = i_\ell + 1$ for ℓ odd
- $i_d = n$.

Together with $i_0 := 0$ these form the indices of Lemma 3.23.

d even

We are looking for such $\{i_j\}$ such that $i_1 \geq 1$. Those are exactly the indices with

---

[10]The following can be formulated in terms of *pulling triangulations*, compare [49, Chap. 16], [31]. For a proof that the pulling triangulation is in fact a triangulation, see [46, Proposition 8.6].

[11]For example, with $n = 4$ and dimension $d = 3$, the indices $\{0, 1, 2\}$, $\{0, 2, 3\}$, $\{0, 3, 4\}$, $\{0, 1, 4\}$, $\{1, 2, 4\}$, $\{2, 3, 4\}$ lead to the facets, which in this dimension are triangles.

- $i_{\ell+1} = i_\ell + 1$ for ℓ odd.

Together with $i_0 := 0$ these form the indices of Lemma 3.23.

Hence

$$\left| \text{Convex-Hull}\left(X^n_{[0,T]} \right) \right| = \sum_i \left| \text{Convex-Hull}(p_{i_0}, .., p_{i_d}) \right|.$$

Now by Lemma 3.22

$$\left| \text{Convex-Hull}(p_{i_0}, .., p_{i_d}) \right| = \frac{1}{d!} \left| \det \begin{bmatrix} 1 & 1 & .. & 1 \\ p_{i_0} & p_{i_1} & .. & p_{i_d} \end{bmatrix} \right|.$$

The determinant is in fact positive here, by (6). We can hence omit the modulus and get

$$\left| \text{Convex-Hull}\left(X^n_{[0,T]} \right) \right| = \sum_i \left| \text{Convex-Hull}(p_{i_0}, .., p_{i_d}) \right|$$

$$= \sum_i \frac{1}{d!} \det \begin{bmatrix} 1 & 1 & .. & 1 \\ p_{i_0} & p_{i_1} & .. & p_{i_d} \end{bmatrix}$$

$$= \frac{1}{d!} \langle S(X^n)_{0,T}, \text{Inv}_d \rangle,$$

by Lemma 3.23.

The statement of the lemma now follows by piecewise linear approximation of X using continuity of the convex hull, which follows from [11, Lemma 3.2], and of iterated integrals [16, Proposition 1.28, Proposition 2.7]. □

4 Rotations

Let

$$\text{SO}\left(\mathbb{R}^d \right) = \left\{ A \in \text{GL}\left(\mathbb{R}^d \right) : AA^\top = \text{id}, \det(A) = 1 \right\},$$

be the group of rotations of $\mathbb{R}^d$.

Definition 4.1 We call $\phi \in T(\mathbb{R}^d)$ an SO **invariant** if

$$\langle S(X)_{0,T}, \phi \rangle = \langle S(AX)_{0,T}, \phi \rangle$$

for all $A \in \text{SO}(\mathbb{R}^d)$ and all curves X. Alternatively, as explained in Sect. 3,

$$A^\top \phi = \phi,$$

for all $A \in \text{SO}(\mathbb{R}^d)$, where the action on $T(\mathbb{R}^d)$ was given in Definition 3.2.

Since $\det(A) = 1$, any GL invariant of weight $w \geq 1$ (Sect. 3) is automatically an SO invariant. But there are SO invariants that are not GL invariants (of any weight), for example, for $d = 2$, $\phi := x_1 x_1 + x_2 x_2$.

Switching to the perspective of multilinear maps, this is the map $(v_1, v_2) \mapsto \langle v_1, v_2 \rangle$. It is shown, see for example [50, Theorem 2.9.A], that all invariants are built from the inner product and the determinant.

Recently, a linear basis for these invariants has been constructed. To formulate the result, we need to introduce some notation from [28]. Define

$$I(r, n) := \{ (i_1, .., i_r) : 1 \leq i_1 < .. < i_r \leq n \}.$$

Use the following partial order on these sequences: for $a \in I(r, n)$, $a' \in I(r', n)$

$$a \geq a'$$

if $r \leq r'$ and $a_j \geq a'_j$ for $j \leq r$.

For $c \in I(d, n)$ and $v_1, .., v_n \in \mathbb{R}^d$, define

$u(c)(v_1, .., v_n) := d$-minor of the $d \times n$ matrix $(v_1, .., v_n)$, with columns given by c.

For $a, b \in I(r, n) \times I(r, n)$ with $r \leq d$ and $v_1, .., v_n \in \mathbb{R}^d$, define

$p(a, b)(v_1, .., v_n) := r$-minor of the matrix $\langle v_i, v_j \rangle$, rows given by a, columns given by b

Theorem 4.2 ([28, Theorem 12.5.0.8]) *Let V be a d-dimensional vector space with inner product $\langle \cdot, \cdot \rangle$. A basis for the space of multilinear maps*

$$\psi : \underbrace{V \times \cdots \times V}_{n \text{ times}} \to \mathbb{R}$$

that satisfy

$$\psi(Av_1, Av_2, \ldots, Av_n) = \psi(v_1, v_2, \ldots, v_n)$$

for all $A \in SO(V)$ and $v_1, \ldots, v_n \in V$ is given by the maps

$$F(v_1, .., v_n) = p\big(a^{(1)}, b^{(1)}\big)(v_1, .., v_n) \cdot .. \cdot p\big(a^{(r)}, b^{(r)}\big)(v_1, .., v_n)$$

$$\cdot u\big(c^{(1)}\big)(v_1, .., v_n) \cdot .. \cdot u\big(c^{(s)}\big)(v_1, .., v_n),$$

satisfying

- $c^{(j)} \in I(d, n)$ *for each $j = 1, .., s$*
- $a^{(j)}, b^{(j)} \in I(t_j, n)$ *for some $1 \leq t_j \leq d - 1$ for each $j = 1, .., r$*
- $a^{(1)} \geq b^{(1)} \geq a^{(2)} \geq .. \geq b^{(r)} \geq c^{(1)} \geq .. \geq c^{(s)}$
- *every number $1, .., n$ appears in exactly one of the sequences $a^{(1)}, .., a^{(r)}, b^{(1)}, .., b^{(r)}$, $c^{(1)}, .., c^{(s)}$; (in particular $n = 2 \cdot C_1 + d \cdot C_2$ for some C_1, C_2 non-negative integers)*

Example 4.3 We give examples of these sequences for $d = 2$.
$n = 1$: There is no such set of sequences, since non-negative integers C_1, C_2 with $2 \cdot C_1 + 2 \cdot C_2 = 1$ cannot be found.
$n = 2$: Allowed sets of sequences are

- $c^{(1)} = (1, 2)$; meaning that $F(v_1, v_2) = \langle v_1, v_2 \rangle$
- $a^{(1)} = (2)$, $b^{(1)} = (1)$; meaning that $F(v_1, v_2) = \det[v_1 v_2]$

$n = 3$: There is no such set of sequences.
$n = 4$: Allowed sets of sequences are

- $a^{(1)} = (4)$, $b^{(1)} = (3)$, $a^{(2)} = (2)$, $b^{(2)} = (1)$; meaning that $F(v_1, v_2, v_3, v_4) = \langle v_4, v_3 \rangle \langle v_2, v_1 \rangle$.
- $a^{(1)} = (4)$, $b^{(1)} = (3)$, $c^{(1)} = (1, 2)$; meaning that $F(v_1, v_2, v_3, v_4) = \langle v_4, v_3 \rangle \det[v_1 v_2]$.
- $a^{(1)} = (4)$, $b^{(1)} = (2)$, $c^{(1)} = (1, 3)$
- $a^{(1)} = (3)$, $b^{(1)} = (2)$, $c^{(1)} = (1, 4)$
- $c^{(1)} = (3, 4)$, $c^{(2)} = (1, 2)$
- $c^{(1)} = (2, 4)$, $c^{(2)} = (1, 3)$

In the setting of $T(\mathbb{R}^d)$ we have

Proposition 4.4 *The* SO *invariants of homogeneity n are spanned by*

$$\mathsf{poly}(\Psi),$$

where Ψ ranges over the invariants of the previous theorem and poly *is given in Lemma 2.4.*

In the case $d = 2$, there is another way to arrive at a basis for the invariants. Taking inspiration from [15], which concerns rotation invariants of images, we work in the complex vector space $T(\mathbb{C}^2)$.[12]

Theorem 4.5 *Define*

$$z_1 = x_1 + i x_2$$
$$z_2 = x_1 - i x_2.$$

The space of SO *invariants on level n in $T(\mathbb{C}^2)$ is spanned freely by*

$$z = z_{j_1} \cdot \ldots \cdot z_{j_n} \quad \text{with } \#\{r : j_r = 1\} = \#\{r : j_r = 2\}.$$

The space of SO *invariants on level n in $T(\mathbb{R}^2)$ is spanned freely by*

$$\mathrm{Re}[z], \mathrm{Im}[z] \quad \text{with } \#\{r : j_r = 1\} = \#\{r : j_r = 2\} \text{ and } z_1 = 1.$$

Remark 4.6 In particular for $d = 2$ and n even, the dimension of rotation invariants on level n in $T(\mathbb{R}^2)$ is equal to $\binom{n}{n/2}$.

Proof 1. **The elements z are invariant**
Let

$$A_\theta := \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

---

[12]One may think of this as the space of noncommuting polynomials in x_1, x_2 with *complex* coefficients, or, equivalently, as the complexification [42, Chap. 1] of the vector space $T(\mathbb{R}^2)$. An element $A \in \mathrm{GL}(\mathbb{R}^2)$ then acts on $T(\mathbb{C}^2)$ by the prescription in Definition 3.2. More abstractly: this is the diagonal action of the complexification of A on $T(\mathbb{C}^2)$.

Then (recall Definition 3.2)

$$
\begin{aligned}
A_\theta^\top z_1 &= A_\theta^\top (x_1 + i x_2) \\
&= \cos(\theta) x_1 + \sin(\theta) x_2 + i\big(-\sin(\theta) x_1 + \cos(\theta) x_2\big) \\
&= e^{-i\theta} z_1 \\
A_\theta^\top z_2 &= e^{i\theta} z_2.
\end{aligned}
$$

Hence

$$
A_\theta^\top z_{j_1} \cdot \ldots \cdot z_{j_n} = z_{j_1} \cdot \ldots \cdot z_{j_n} \, \forall \theta \quad \text{if and only} \quad \#\{r : j_r = 1\} = \#\{r : j_r = 2\}.
$$

2. **The elements z form a basis**

Now $x_{j_1}..x_{j_n} : j_\ell \in \{1, 2\}$ is a basis of $\pi_n T(\mathbb{C}^2)$ with respect to $\mathbb{C}$. Hence $z_{j_1}..z_{j_n}$ is (the map $(x_1, x_2) \mapsto (z_1, z_2)$ is invertible). By Step 1 we have hence exhibited a basis (with respect to $\mathbb{C}$) for all invariants in $\pi_n T(\mathbb{C}^2)$.

3. **Real invariants**

The space of SO invariants on level n in $T(\mathbb{C}^2)$ is spanned freely by the set of

$$
z_{j_1} \cdot \ldots \cdot z_{j_n} \quad \text{with} \ \#\{r : j_r = 1\} = \#\{r : j_r = 2\}.
$$

Adding and subtracting the elements with $j_1 = 2$ from the elements with $j_1 = 1$, we get that the space of SO invariants on level n in $T(\mathbb{C}^2)$ is spanned freely by the set of

$$
(z_{j_1} \cdot \ldots \cdot z_{j_n} + z_{3-j_1} \cdot \ldots \cdot z_{3-j_n}) \quad \text{and} \quad (z_{j_1} \cdot \ldots \cdot z_{j_n} - z_{3-j_1} \cdot \ldots \cdot z_{3-j_n})
$$
$$
\text{with} \ \#\{r : j_r = 1\} = \#\{r : j_r = 2\} \ \text{and} \ j_1 = 1.
$$

Because $z_{3-j_1} \cdot \ldots \cdot z_{3-j_n}$ is the complex conjugate of $z_{j_1} \cdot \ldots \cdot z_{j_n}$, this means that the space of SO invariants on level n in $T(\mathbb{C}^2)$ is spanned freely by the set of

$$
\mathrm{Re}(z_{j_1} \cdot \ldots \cdot z_{j_n}) \quad \text{and} \quad \mathrm{Im}(z_{j_1} \cdot \ldots \cdot z_{j_n})
$$
$$
\text{with} \ \#\{r : j_r = 1\} = \#\{r : j_r = 2\} \ \text{and} \ j_1 = 1.
$$

This is an expression for a basis of the SO invariants in terms of real combinations of basis elements of the tensor space. They thus form a basis for the SO invariants for the free *real* vector space on the same set, namely $\pi_n T(\mathbb{R}^2)$. □

Example 4.7 Consider $d = 2$, level $n = 2$

$$
11 + 22
$$
$$
-12 + 21
$$

Level $n = 4$

$$
1111 - 1122 + 1212 + 1221 + 2112 + 2121 - 2211 + 2222
$$
$$
-1112 - 1121 + 1211 - 1222 + 2111 - 2122 + 2212 + 2221
$$
$$
1111 + 1122 - 1212 + 1221 + 2112 - 2121 + 2211 + 2222
$$

$$-1112 + 1121 - 1211 - 1222 + 2111 + 2122 - 2212 + 2221$$
$$1111 + 1122 + 1212 - 1221 - 2112 + 2121 + 2211 + 2222$$
$$1112 - 1121 - 1211 - 1222 + 2111 + 2122 + 2212 - 2221$$

Consider $d = 3$, level $n = 2$

$$11 + 22 + 33$$

Level $n = 3$.

$$123 - 132 + 312 - 321 + 231 - 213$$

Consider $d = 4$, level $n = 2$

$$11 + 22 + 33 + 44.$$

Level $n = 4$

$$1144 + 4422 + 4444 + 3333 + 1122 + 4433 + 1133 + 4411 + 2211 + 3344$$
$$+ 1111 + 2244 + 2222 + 3322 + 2233 + 3311$$
$$+ 4 \text{ more}$$

5 Permutations

Denote by S_d the group of permutations of $[d] := \{1, .., d\}$.

Lemma 5.1 *For $\sigma \in S_d$, define $M(\sigma) \in \mathrm{GL}(\mathbb{R}^d)$ as*

$$M(\sigma)_{ij} = 1 \quad if\ i = \sigma(j).$$

Then $M : S_d \to \mathrm{GL}(\mathbb{R}^d)$ is a group homomorphism and moreover $M(\sigma^{-1}) = M(\sigma)^\top$.[13]

Proof Regarding the first point, for $i = \{1, .., d\}$,

$$M(\sigma)M(\tau)e_i = M(\sigma)e_{\tau(i)} = e_{\sigma(\tau(i))} = M(\sigma\tau)e_i.$$

Regarding the last point, note the following sequence of equivalences.

$$M_{ij}(\sigma^{-1}) = 1 \quad \Leftrightarrow \quad i = \sigma^{-1}(j) \quad \Leftrightarrow \quad j = \sigma(i) \quad \Leftrightarrow \quad M_{ji}(\sigma) = 1.$$

This proves the claim. $\qquad\square$

S_d then acts on $T((\mathbb{R}^d))$ and $T(\mathbb{R}^d)$ via Definition 3.2. Explicitly,

$$\sigma \cdot x_{i_1}..x_{i_n} = x_{\sigma(i_1)}..x_{\sigma(i_n)}.$$

---

[13] M is sometimes called the *defining representation* of S_d.

Definition 5.2 We call $\phi \in T(\mathbb{R}^d)$ a **permutation invariant** if

$$\left\langle S\big(M(\sigma)X\big)_{0,T}, \phi \right\rangle = \left\langle S(X)_{0,T}, \phi \right\rangle$$

for all $\sigma \in S_d$ and all curves X. Alternatively, as explained in Sect. 3,

$$M(\sigma)^\top \phi = \phi,$$

for all $\sigma \in S_d$. Equivalently,

$$M(\sigma)\phi = \phi,$$

for all $\sigma \in S_d$.

We follow [1, Sect. 3]. To a monomial

$$x_{i_1} \cdot .. \cdot x_{i_n},$$

we associate the following set partition of $[n] := \{1, .., n\}$

$$\nabla(x_{i_1} \cdot .. \cdot x_{i_n}) := \big\{ \{\ell : i_\ell = p\} : p \in [d] \big\} \setminus \big\{ \{\} \big\}.$$

Example 5.3 Let $d = 3$, then

$$\nabla(x_2 x_3 x_2 x_2 x_1) = \big\{ \{1, 3, 4\}, \{2\}, \{5\} \big\}.$$

Note that for every permutation $\sigma \in S_d$,

$$\nabla(x_{i_1} \cdot .. \cdot x_{i_n}) = \nabla(x_{\sigma(i_1)} \cdot .. \cdot x_{\sigma(i_n)}). \tag{7}$$

Proposition 5.4 ([1, Sect. 3]) *Define*

$$M_A := \sum_{i : \nabla(x_{i_1} .. x_{i_n}) = A} x_{i_1} .. x_{i_n}.$$

Then $\{M_A : A$ is set partition of $[n]$ and $|A| \le d\}$ is a linear basis for the space of permutation invariants of homogeneity n.

Remark 5.5 The generating function for partitions with at most d blocks is given by

$$\frac{\sum_{\ell=1}^{d+1} x^\ell \prod_{m=\ell}^{d}(1 - m\,x)}{\prod_{\ell=1}^{d}(1 - \ell\,x)}.$$

This follows from summing up [45, (1.94c)].

For example $d = 2$,

$$\frac{(1-x)(1-2x) + x(1-2x) + x^2}{(1-x)(1-2x)} = \frac{1-x}{1-2x},$$

which is the generating function of the sequence (https://oeis.org/A011782)

$$1, 2^0, 2^1, 2^2, 2^3, 2^4, ..$$

For $d = 3$ one gets, the generating function

$$\frac{(1-x)(1-2x)(1-3x) + x(1-2x)(1-3x) + x^2(1-3x) + x^3}{(1-x)(1-2x)(1-3x)},$$

which is the generating function of the sequence (https://oeis.org/A124302)

$$1, \left(3^0 + 1\right)/2, \left(3^1 + 1\right)/2, \left(3^2 + 1\right)/2, ..$$

We are not aware of a general explicit formula for the number of partitions (i.e. the coefficients of the generating function).

Proof of Proposition 5.4 By (7), each M_A is permutation invariant. Moreover, since $|A| \leq d$, M_A is nonzero.

For A, A' distinct set partitions of $[n]$, the monomials in M_A and the monomials in $M_{A'}$ do not overlap. Hence the proposed basis is linearly independent.

Now, if ϕ is permutation invariant and if for some i, i', $\nabla(x_{i_1}..x_{i_n}) = \nabla(x_{i'_1}..x_{i'_n})$ then the coefficient of x_i and $x_{i'}$ must coincide. Hence the proposed basis spans invariants of homogeneity n. □

Example 5.6 Consider $d = 3$
Order $n = 1$

$$1 + 2 + 3$$

Order $n = 2$

$$33 + 22 + 11$$

$$32 + 31 + 23 + 21 + 13 + 12$$

Order $n = 3$

$$333 + 222 + 111$$

$$332 + 331 + 223 + 221 + 113 + 112$$

$$323 + 313 + 232 + 212 + 131 + 121$$

$$322 + 311 + 233 + 211 + 133 + 122$$

$$321 + 312 + 231 + 213 + 132 + 123$$

6 An Additional (Time) Coordinate

Assume now that $X = (X^0, X^1, .., X^d) : [0, T] \to \mathbb{R}^{1+d}$. Here X^0 plays a special role, in that we assume that it is not affected by the space transformations under consideration.

Adding an "artificial" 0-th component, usually keeping track of time, $X_t^0 := t$, is a common trick to improve the expressiveness of the signature. In particular, if such an X^0 is monotonically increasing, the enlarged curve $(X^0, X^1, .., X^d)$ never has any "tree-like" components (compare Sect. 7), no matter what the original $(X^1, .., X^d)$ was.

Consider GL invariants for the moment.

Definition 6.1 Let

$$\mathrm{GL}_0\big(\mathbb{R}^d\big) := \big\{A \in \mathrm{GL}\big(\mathbb{R}^{1+d}\big) : Ae_0 = A^{-1}e_0 = e_0\big\},$$

the space of invertible maps of $\mathbb{R}^{1+d}$ leaving the first direction unchanged. We call $\phi \in T(\mathbb{R}^{1+d})$ a GL_0 **invariant of weight** w if

$$A^\top \phi = (\det A)^w \phi,$$

for all $A \in \mathrm{GL}_0(\mathbb{R}^d)$.

Consider the $\mathrm{GL}(\mathbb{R}^2)$ invariant of weight 1

$$x_1 x_2 - x_2 x_1.$$

Since elements of $\mathrm{GL}_0(\mathbb{R}^2)$ leave the variable x_0 unchanged, a straightforward way to produce GL_0 invariants presents itself: insert x_0 at the same position in every monomial. For example

$$x_1 x_0 x_2 - x_2 x_0 x_1$$

is a $\mathrm{GL}_0(\mathbb{R}^2)$ invariant of weight 1. We now formalize this idea and show that we get every GL_0 invariant this way.

Define the linear map Remove of "removing instances of x_0" on monomials, as

$$\mathsf{Remove}\, x_{i_1}..x_{i_m} := \prod_{\ell: i_\ell \neq 0} x_{i_\ell},$$

so for example

$$\mathsf{Remove}\, x_0 x_1 x_1 x_0 x_3 = x_1 x_1 x_3$$

$$\mathsf{Remove}\, x_0 x_0 = 1.$$

Define for $U \subset [m]$ and $i = (i_1, .., i_m)$

$$i|_U = (i_\ell : \ell = 1, .., m; \ell \in U).$$

Define the linear map of restriction to U on polynomials of order m by defining on monomials

$$x_i|_U := x_{i|_U}$$

so for example

$$x_{i_1} x_{i_2} x_{i_3}|_{\{1,3\}} = x_{i_1} x_{i_3}.$$

For $z = (z_1, .., z_{m+1}) \in \mathbb{N}^{m+1}$ denote by Insert$_z$ the linear operator on polynomials of order m by defining it on monomials as follows. For a monomial $x_{i_1}..x_{i_m}$ of order m, Insert$_z$ inserts z_1 occurrences of x_0 before x_{i_1}, z_2 occurrences of x_0 before x_{i_2}, .., z_m occurrences of x_0 before x_{i_m} and z_{m+1} occurrences of x_0 after x_{i_m}. For example

$$\mathsf{Insert}_{(2,1,4)} x_1 x_2 = x_0 x_0 x_1 x_0 x_2 x_0 x_0 x_0 x_0.$$

Theorem 6.2 *A basis for the space of* GL_0 *invariants of weight* w, *homogeneous of degree* m, *is given by the polynomials*

$$\mathsf{Insert}_z \psi,$$

with $0 \le n \le m$, ψ *ranging over the basis for* GL *invariant of weight* w *and homogeneity* n (*Proposition* 3.11) *and* $z \in \mathbb{N}^{n+1}$ *such that* $\sum_\ell z_\ell = m - n$.

Proof Let n, ψ, z be as in the statement. Then, for $A_0 = \mathrm{diag}(1, A) \in \mathrm{GL}_0(\mathbb{R}^d)$, with $A \in \mathrm{GL}(\mathbb{R}^d)$,

$$A_0 \, \mathsf{Insert}_z \psi = \mathsf{Insert}_z A \psi = (\det A)^w \, \mathsf{Insert}_z \psi.$$

Therefore $\mathsf{Insert}_z \psi$ is GL_0 invariant of weight w.

On the other hand, let ϕ of order m be a GL_0 invariant modulo time of weight w. Define for $U \subset [m]$

$$\phi^U := \sum_{i : i_\ell = 0, \ell \in U; i_j \neq 0, j \notin U} \langle \phi, x_i \rangle x_i,$$

which collects all monomials having x_0 exactly at the positions in U. Then

$$\phi = \sum_{U \subset [m]} \phi^U.$$

Now, since ϕ is GL_0 invariant of weight w and since GL_0 leaves

$$\mathrm{span}\{x_i : i_\ell = 0, \ell \in U; i_j \neq 0, j \notin U\}$$

invariant, we get that ϕ^U is GL_0 invariant of weight w. Clearly, there is $0 \le n \le m$ and $z \in \mathbb{N}^{n+1}$ such that

$$\mathsf{Insert}_z \, \mathsf{Remove} \, \phi^U = \phi^U.$$

Lastly, $\mathsf{Remove} \, \phi^U$ is GL invariant, since for $A_0 = \mathrm{diag}(1, A) \in \mathrm{GL}_0(\mathbb{R}^d)$, with $A \in \mathrm{GL}(\mathbb{R}^d)$,

$$A \, \mathsf{Remove} \, \phi^U = \mathsf{Remove} \, A_0 \phi^U = (\det A_0)^w \, \mathsf{Remove} \, \phi^U = (\det A)^w \, \mathsf{Remove} \, \phi^U.$$

Hence every invariant is in the span of the set given in the statement. They are linearly independent, and hence form a basis. □

The corresponding statements for rotations and permutations are completely analogous, so we omit them.

7 Discussion and Open Problems

We have presented a novel way to extract invariant features of d-dimensional curves, based on the iterated-integral signature. We have identified all those features that can be written as a finite linear combination of terms in the signature.

Among the techniques used previously for finding invariants of curves, the method of "integral invariants" [13] is closest to ours (it has been used for example in [19] for character recognition). In that work, for a curve $X : [0, T] \to \mathbb{R}^d$, $d = 2, 3$, the building blocks for invariants are expressions of the form

$$\int_0^T (X_r^1)^{\alpha_1} .. (X_r^d)^{\alpha_d} dX_r^i, \quad i = 1, .., d. \tag{8}$$

Using an algorithmic procedure, some invariants to certain subgroups of $G \subset \mathrm{GL}(\mathbb{R}^d)$ are derived. In particular for $d = 2$ and $G = \mathrm{GL}(\mathbb{R}^d)$ the following invariants are given

$$I_1 = \frac{1}{2} \int_0^T X_{0,r}^1 dX_r^2 - \frac{1}{2} X_{0,t}^1 X_{0,t}^2$$

$$I_2 = \int_0^T X_{0,r}^1 X_{0,r}^2 dX_r^2 \, X_{0,t}^1 - \frac{1}{2} \int_0^t (X_r^1)^2 dX_r^2 \, X_{0,t}^2$$

$$I_3 = \int_0^T X_{0,r}^1 (X_{0,r}^2)^2 dX_r^2 \, X_{0,T}^2 - \int_0^T (X_{0,r}^1)^2 X_{0,r}^2 dX_r^2 \, X_{0,T}^1 X_{0,T}^2$$

$$+ \frac{1}{3} \int_0^T (X_{0,r}^1)^3 dX_r^2 \, X_{0,r}^2 X_{0,r}^2 - \frac{1}{12} (X_{0,t}^1)^3 (X_{0,t}^2)^3.$$

By the shuffle identity (Lemma 2.1), we can write these as $I_i = \langle S(X)_{0,T}, \phi_i \rangle$, with

$$\phi_1 := \frac{1}{2} 12 - \frac{1}{2} 12$$

$$\phi_2 := \frac{1}{3} 1221 + \frac{1}{3} 1212 - \frac{2}{3} 1122 + \frac{1}{3} 2121 + \frac{1}{3} 2112 - \frac{2}{3} 2211$$

$$\phi_3 := -121212 - 211122 + 212121 + 221112 - 121221 + 122211 - 112212$$

$$+ 122112 - 211212 - 211221 - 121122 + 122121 - 3222111 + 3111222$$

$$+ 221121 + 212211 - 112122 + 212112 - 112221 + 221211.$$

One can easily check that these lie in the linear span of the invariants given in Proposition 4.4 (or Theorem 4.5), as expected.

We note that expressions of the form (8) are *not* enough to uniquely characterize a path. Indeed, the following lemma gives a counterexample to the conjecture on p. 906 in [13] that "signatures of non-equivalent curves are different" (here, the "signature" of a curve means the set of expressions of the form (8)).

Lemma 7.1 *Consider the two closed curves* X^+ *and* X^- *in* $\mathbb{R}^2$, *given for t in* $[0, 2\pi]$ *as*

$$X_t^{\pm,1} = \pm \cos t$$

$$X_t^{\pm,2} = \sin 2t.$$

Then all the expressions (8) *coincide on* X^+ *and* X^-.[14]

---

[14]Note that X^+ and X^- are not tree-like equivalent and therefore have different (iterated-integral) signatures. The lowest level on which they differ is level 4.

Fig. 2 The lemniscate of
Gerono. Traversing it once from
each of the two starting points
indicated gives two distinct
closed curves with distinct
iterated-integral signatures, but
which cannot be distinguished
with the "signature" of [13]
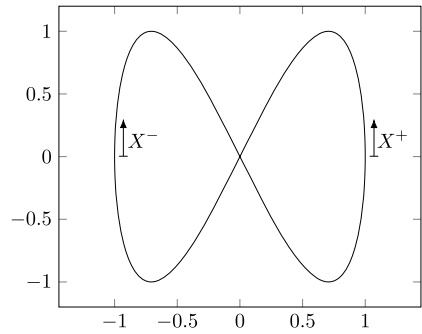
These curves both trace a figure called the *lemniscate of Gerono* which is illustrated in
Fig. 2.

Proof Consider the function $f_n^m(t) := \cos^m t \sin^n t$, where m and n are nonnegative integers.
If n is odd, then $f_n^m(t) = -f_n^m(2\pi - t)$ so $\int_0^{2\pi} f_n^m(t)\,dt$ is zero. If m is odd, then

$$\int_0^{2\pi} f_n^m(t)\,dt = -\int_{\frac{\pi}{2}}^{-\frac{3\pi}{2}} f_n^m\left(\frac{\pi}{2} - t\right)dt = \int_{-\frac{3\pi}{2}}^{\frac{\pi}{2}} f_m^n(t)\,dt = \int_0^{2\pi} f_m^n(t)\,dt = 0.$$

Thus $\int_0^{2\pi} f_n^m(t)\,dt$ can only be nonzero if m and n are both even.

Any expression like (8) is either of the form

$$A_{m,n}^{\pm} = \int_0^{2\pi} \left(X_t^{\pm,1}\right)^m \left(X_t^{\pm,2}\right)^n dX_t^{\pm,1}$$

$$= \int_0^{2\pi} (\pm 1)^m \cos^m t \sin^n 2t (\mp \sin t)\,dt$$

$$= \mp 2^n (\pm 1)^m \int_0^{2\pi} \cos^{m+n} t \sin^{n+1} t\,dt$$

$$= \begin{cases} 0 & n \text{ even or } m \text{ even} \\ -2^n \int_0^{2\pi} \cos^{m+n} t \sin^{n+1} t\,dt & \text{otherwise} \end{cases}$$

or of the form

$$B_{m,n}^{\pm} = \int_0^{2\pi} \left(X_t^{\pm,1}\right)^m \left(X_t^{\pm,2}\right)^n dX_t^{\pm,2}$$

$$= \int_0^{2\pi} (\pm 1)^m \cos^m t \sin^n 2t (2\cos 2t)\,dt$$

$$= 2^{n+1} (\pm 1)^m \int_0^{2\pi} \cos^{m+n} t \sin^n t \left(\cos^2 t - \sin^2 t\right)dt$$

$$= \begin{cases} 0 & n \text{ odd or } m \text{ odd} \\ 2^{n+1} \int_0^{2\pi} \cos^{m+n} t \sin^n t (\cos^2 t - \sin^2 t)\,dt & \text{otherwise.} \end{cases}$$

Both these expressions are free from the symbols $\pm$ and $\mp$. Therefore these two curves have the same values on terms of the form (8). □

Moreover, the algorithmic nature of the construction in [13] makes it difficult to proceed to invariants of higher order. In contrast, our method gives an explicit linear basis for the invariants under consideration up to *any* order.

Regarding the question of whether our invariants are complete we propose the following conjecture. As shown in [22], if $S(X)_{0,T} = S(Y)_{0,T}$ for some curves X and Y, then X is "tree-like equivalent" to Y. For the concrete definition of this equivalence we refer to their paper, but let us give one example. Consider in $d = 2$, the constant path $X_t := (0, 0)$, $t \in [0, T]$ and the piecewise linear path Y, between the points $(0, 0)$, $(1, 0)$ and $(0, 0)$. One can check that

$$S(X)_{0,T} = S(Y)_{0,T} = 1.$$

The signature has no chance of picking up these kind of "excursions" in a path; this concept is formalized in "tree-like equivalence". We suspect that the following holds true (with corresponding formulations for the other subgroups of $\mathrm{GL}(\mathbb{R}^d)$).

Conjecture 7.2 *Let $X, Y : [0, T] \to \mathbb{R}^d$ be two curves such that*

$$\langle S(X)_{0,T}, \phi \rangle = \langle S(Y)_{0,T}, \phi \rangle,$$

for all SO *invariants given in Proposition* 4.4. *Then there is a curve* $\bar{X}$, *tree-like equivalent to* X, *and a rotation* $A \in \mathrm{SO}(\mathbb{R}^d)$, *such that*

$$A\bar{X} = Y.$$

In Proposition 3.11, Proposition 4.4 and Proposition 5.4 we have established a linear basis for invariants for every homogeneity. As already mentioned in Remark 3.15, owing to the shuffle identity, there are algebraic relations between elements of different homogeneity. An interesting open problem is then to find a minimal set of generators for the set of invariants, considered as a subalgebra of the shuffle algebra. This applies to all subgroups of $\mathrm{GL}(\mathbb{R}^d)$ and their corresponding invariants.

Lastly, a word on (computational) complexity. We have seen in Remark 3.10 the dimensions of GL invariant elements (which is a lower bound on the dimensions of SO invariant elements).[15] In Remark 5.5 we have seen the dimensions for the permutation invariant elements.

Computing the signature itself up to level n has complexity $\Omega(d^n)$, since $d + .. + d^n$ integrals need to be calculated. So any method that calculates the invariant features of a curve X by first calculating its signature and extracting them (see Remark 3.14) will have computational complexity dominated by the calculation of the signature. Furthermore, the calculation of the invariant elements is a computation that can be done offline (they do not depend on the curve X).

This leaves several directions of future research.

---

[15]In this section we take "invariant element" to mean the elements of the space $T(\mathbb{R}^d)$, incarnations of which can be seen in Sect. 3.1. They do *not* depend on any curve X one might be interested in.

- Is it possible to apply kernelization techniques similar to the ones used for the entire (non-invariant) signature in [25]? These techniques, in the non-invariant setting, allow to use information of the signature up to high levels and dimension for certain learning algorithms.
- We have studied in this paper *linear* expressions on the signature that are invariant to a group action. This was justified by using the shuffle identity (Lemma 2.1), which tells us that any *polynomial* functional on the signature can in fact be linearized. One can also consider a fixed level n of the signature and look for all *nonlinear* expressions that are invariant under the group action. This is the classical problem of invariant theory for polynomial rings [47, Sect. 4]. On the one hand, this makes it possible to "peek ahead" in the signature, since one gets invariant information that would only be seen in linear expressions of higher levels than n. On the other hand, except for special cases, there are no explicit expressions for these nonlinear invariant. One has to proceed algorithmically (for example via Derksen's algorithm, [8]) which only works for low dimension d and low levels n. Since the calculation of those nonlinear invariant elements can also be done offline it would nonetheless be nice to have a tabulation of nonlinear invariants (as far as existing algorithms can reach).
- For GL invariants, in Remark 3.25 we conjecture the existence of a "GL invariant" signature. This could improve computation time, since no non-invariant integrals have to be computed.

References

1. Bergeron, N., Reutenauer, C., Rosas, M., Zabrocki, M.: Invariants and coinvariants of the symmetric group in noncommuting variables. Can. J. Math. **60**(2), 266–296 (2008)
2. Calabi, E., Olver, P.J., Shakiban, C., Tannenbaum, A., Haker, S.: Differential and numerically invariant signature curves applied to object recognition. Int. J. Comput. Vis. **26**(2), 107–135 (1998)
3. Cass, T., Friz, P.: Densities for rough differential equations under Hörmander's condition. Ann. Math. **171**(3), 2115–2141 (2010)
4. Chen, K.-T.: Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. Ann. Math. **65**(1), 163–178 (1957)
5. Chevyrev, I., Kormilitzin, A.: A primer on the signature method in machine learning. arXiv:1603.03788 (2016)
6. Chuang, G.C.H., Kuo, C.-C.J.: Wavelet descriptor of planar curves: theory and applications. IEEE Trans. Image Process. **5**(1), 56–70 (1996)
7. De Bruijn, N.G.: On some multiple integrals involving determinants. J. Indian Math. Soc. **19**, 133–151 (1955)
8. Derksen, H.: Computation of invariants for reductive groups. Adv. Math. **141**(2), 366–384 (1999)
9. Diehl, J.: signature-invariants-py, GitHub repository. https://github.com/diehlj/signature-invariants-py
10. Dieudonné, J.A., Carrell, J.B.: Invariant theory, old and new. Adv. Math. **4**(1), 1–80 (1970)
11. Engström, A., Patrik, N.: Polytopes from subgraph statistics. In: 23rd International Conference on Formal Power Series and Algebraic Combinatorics. Discrete Mathematics & Theoretical Computer Science, pp. 305–316 (2011)
12. Ewald, A., Marzetti, L., Zappasodi, F., Meinecke, F.C., Nolte, G.: Estimating true brain connectivity from EEG/MEG data invariant to linear and static transformations in sensor space. NeuroImage **60**(1), 476–488 (2012)

13. Feng, S., Kogan, I., Krim, H.: Classification of curves in 2D and 3D via affine integral signatures. Acta Appl. Math. **109**(3), 903–937 (2010)
14. Fillmore, J.P., Williamson, S.G.: Permanents and determinants with generic noncommuting entries. Linear Multilinear Algebra **19**(4), 321–334 (1986)
15. Flusser, J.: On the independence of rotation moment invariants. Pattern Recognit. **33**(9), 1405–1410 (2000)
16. Friz, P., Victoir, N.: Multidimensional Stochastic Processes as Rough Paths: Theory and Applications. Cambridge Studies in Advanced Mathematics, vol. 120. Cambridge University Press, Cambridge (2010)
17. Gale, D.: Neighborly and cyclic polytopes. In: Convexity. Sympos. Pure Math., vol. 7, pp. 225–233. AMS, Providence (1963)
18. Gardner, R.B.: The fundamental theorem of vector relative invariants. J. Algebra **36**(2), 314–318 (1975)
19. Golubitsky, O., Vadim, M., Watt, S.M.: Orientation-independent recognition of handwritten characters with integral invariants. In: Proc. Joint Conf. ASCM (2009)
20. Graham, B.: Sparse arrays of signatures for online character recognition. arXiv:1308.0371 (2013)
21. Granlund, G.H.: Fourier preprocessing for hand print character recognition. IEEE Trans. Comput. **100**(2), 195–201 (1972)
22. Hambly, B., Lyons, T.: Uniqueness for the signature of a path of bounded variation and the reduced path group. Ann. Math. **171**(1), 109–167 (2010)
23. Johnson, H.H.: A generalization of KT Chen's invariants for paths under transformation groups. Trans. Am. Math. Soc. **105**(3), 453–461 (1962)
24. Karlin, S., Shapley, L.S.: Geometry of Moment Spaces. Memoirs of the American Mathematical Society, vol. 12. AMS, Providence (1953)
25. Király, F.J., Oberhauser, H.: Kernels for sequentially ordered data. arXiv:1601.08169 (2016)
26. Kormilitzin, A., Saunders, K.E.A., Harrison, P.J., Geddes, J.R., Lyons, T.: Detecting early signs of depressive and manic episodes in patients with bipolar disorder using the signature-based model. arXiv:1708.01206 (2017)
27. Kuhl, F.P., Giardina, C.R.: Elliptic Fourier features of a closed contour. Comput. Graph. Image Process. **18**(3), 236–258 (1982)
28. Lakshmibai, V., Komaranapuram, N.R.: Standard Monomial Theory: Invariant Theoretic Approach. Encyclopaedia of Mathematical Sciences, vol. 137. Springer, Berlin (2007)
29. Landsberg, J.M.: Tensors: geometry and applications. Represent. Theory **381**, 402 (2012)
30. Leclerc, B.: On identities satisfied by minors of a matrix. Adv. Math. **100**(1), 101–132 (1993)
31. Lee, C.: Regular triangulations of convex polytopes. In: Gritzmann, P., Sturmfels, B. (eds.) Applied Geometry and Discrete Mathematics: The Victor Klee Festschrift, pp. 443–456. American Mathematical Society, Providence (1991)
32. Levin, D., Lyons, T., Ni, H.: Learning from the past, predicting the statistics for the future, learning an evolving system. arXiv:1309.0260 (2013)
33. Luque, J.-G., Thibon, J.-Y.: Pfaffian and Hafnian identities in shuffle algebras. Adv. Appl. Math. **29**(4), 620–646 (2002)
34. Lyons, T.J., Phillip, S.Y.: On Gauss-Green theorem and boundaries of a class of Hölder domains. J. Math. Pures Appl. **85**(1), 38–53 (2006)
35. Manay, S., Cremers, D., Hong, B.W., Yezzi, A.J., Soatto, S.: Integral invariants for shape matching. IEEE Trans. Pattern Anal. Mach. Intell. **28**(10), 1602–1618 (2006)
36. Mokhtarian, F., Mackworth, A.: Scale-based description and recognition of planar curves and two-dimensional shapes. IEEE Trans. Pattern Anal. Mach. Intell. **1**, 34–43 (1986)
37. Morales, J., Akopian, D.: Physical activity recognition by smartphones, a survey. Biocybern. Biomed. Eng. **37**(3), 388–400 (2017)
38. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes: The Art of Scientific Computing, 3rd edn. Cambridge University Press, Cambridge (2007)
39. Ree, R.: Lie elements and an algebra associated with shuffles. Ann. Math. **68**(2), 210–220 (1958)
40. Reizenstein, J.: Calculation of iterated-integral signatures and log signatures. arXiv:1712.02757 (2017)
41. Reutenauer, C.: Free Lie Algebras. London Mathematical Society Monographs. New Series, vol. 7. Clarendon, Oxford (1993)
42. Roman, S.: Advanced Linear Algebra, 3rd edn. Graduate Texts in Mathematics, vol. 135. Springer, New York (2005)
43. Rudin, W.: Principles of Mathematical Analysis, 3rd edn. International Series in Pure & Applied Mathematics. McGraw-Hill, New York (1964)
44. Sagan, B.: The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions. Graduate Texts in Mathematics, vol. 203. Springer, Berlin (2013)
45. Stanley, R.P.: Enumerative Combinatorics, Volume 1, 2nd edn. Cambridge Studies in Advanced Mathematics, vol. 49. Cambridge University Press, Cambridge (2011)

46. Sturmfels, B.: Gröbner Bases and Convex Polytopes. University Lecture Series, vol. 8. AMS, Providence (1996)
47. Sturmfels, B.: Algorithms in Invariant Theory. Springer, Berlin (2008)
48. Sturmfels, B., White, N.: Gröbner bases and invariant theory. Adv. Math. **76**(2), 245–259 (1989)
49. Toth, C.D., O'Rourke, J., Goodman, J.E. (eds.): Handbook of Discrete and Computational Geometry. CRC Press, Boca Raton (2004)
50. Weyl, H.: The Classical Groups, Their Invariants and Representations. Princeton University Press, Princeton (1946)
51. Yang, W., Lyons, T., Ni, H., Schmid, C., Jin, L., Chang, J.: Leveraging the path signature for skeleton-based human action recognition. arXiv:1707.03993 (2017)
52. Zahn, C.T., Roskies, R.Z.: Fourier descriptors for plane closed curves. IEEE Trans. Comput. **100**(3), 269–281 (1972)
53. Ziegler, G.: Lectures on Polytopes. Graduate Texts in Mathematics, vol. 152. Springer, Berlin (2012)