Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

The rough Veronese variety

by

Francesco Galuppi

Preprint no.: 91 2018

THE ROUGH VERONESE VARIETY

FRANCESCO GALUPPI

Abstract. We study signature tensors of paths from an algebraic geometric viewpoint. The signatures of a given class of paths parametrize a variety inside the space of tensors, and these signature varieties provide both new tools to investigate paths and new challenging questions about their behavior. This paper focuses on signatures of rough paths. Their signature variety shows surprising analogies with the Veronese variety, and our aim is to prove that this so-called Rough Veronese is toric. The same holds for the universal variety. Answering a question of Amendola, Friz and Sturmfels, we show that the ideal of the universal variety does not need to be generated by quadrics.

Introduction

A *path* is a continuous map $X : [0,1] \to \mathbb{R}^d$. Classically, the components $X_1, \dots, X_d$ of X are assumed to be sufficiently smooth. For every positive integer k, it is therefore possible to define an order k tensor $\sigma^{(k)}(X)$, whose $(i_1 \dots i_k)$-th entry is

$$\int_0^1 \int_0^{t_k} \cdots \int_0^{t_3} \int_0^{t_2} \dot{X}_{i_1}(t_1) \cdot \ldots \cdot \dot{X}_{i_k}(t_k) dt_1 \dots dt_k.$$

By convention, we define $\sigma^{(0)}(X) = 1$. The sequence

$$\sigma(X) = (\sigma^{(k)}(X) \mid k \in \mathbb{N})$$

is called the *signature* of X.

Signatures were first defined in [2], and they enjoy many interesting properties. For instance, up to a mild equivalence relation, the signature allows to uniquely recover a piecewise differentiable path (see [3, Theorem 4.1]).

Many physical behaviors and experiments can be modeled by using paths, and signatures are useful tools to encode the information carried by paths into a compact form. In [1], the authors consider signature tensors from an algebraic geometry perspective. If we fix a certain class of paths and the order k of the tensors, then the k-th signature $\sigma^{(k)}$ is a polynomial map into $(\mathbb{R}^d)^{\otimes k}$. Its image variety parametrizes the closure of the set of all k-th signatures of paths of the chosen class. The study of this map, this

1991 *Mathematics Subject Classification.* 14Q15, 14M25, 60H99.

Key words and phrases. Rough Veronese variety, signature tensor, rough path, Lie algebra, Lyndon words, exponential.

The Nonlinear Algebra research group at Max Planck Institute is a very stimulating environment, and I would like to thank Rosa Preiß for introducing me to the topic, Jane Coons for her help with the combinatorics, Joscha Diehl for his many explanations about the tensor algebra, Mateusz Michałek for his patience and his insight, Laura Colmenarejo for providing lots of examples, and Max Pfeffer, Anna Seigal and Bernd Sturmfels for useful suggestions and careful proofreading.

variety and its geometric properties can be interesting for many reasons. For instance, in applied problems one sometimes has a signature, coming from empirical data, and wants to know if there is a path of a certain kind (say, piecewise linear) having that given signature. Knowing whether the map is injective, or at least finitely many to one, tells us if there are chances to solve this inverse problem. Another issue is the study of singularities. What does it mean, for a path, to have a signature which is a singular point in the image variety?

In [1] we find a detailed study of the image varieties of polynomial paths, piecewise linear paths and also random paths arising from Brownian motion. These three classes of paths have a common generalization: the class of rough paths. Rough paths have a number of applications, for instance the study of controlled ODEs and stochastic PDEs (see [6]), as well as sound compression (see [10]). While they are not necessarily piecewise differentiable, it is possible to define their signature, and therefore to study their signature variety. Even at a first glance, such a variety exhibits analogies with the Veronese variety, and it is therefore named the *Rough Veronese variety* in [1, Section 5.4]. The main purpose of this paper is to study its geometry. We will prove that the Rough Veronese variety is a toric variety, and we will characterize the monomials parameterizing it.

Notations and preliminaries

The k-th signature of a path X belongs to $(\mathbb{R}^d)^{\otimes k}$, but we need a space to store the whole signature $\sigma(X)$. In this Section we define such a space, which has a rich algebraic structure. We will recall the features we need, but we do not attempt to any extent to describe all its properties. Every definition and result of this section can be found in [12].

Definition 1. The *tensor algebra* over $\mathbb{R}^d$ is the graded $\mathbb{R}$-vector space

$$T((\mathbb{R}^d)) = \mathbb{R} \times \mathbb{R}^d \times (\mathbb{R}^d)^{\otimes 2} \times \dots$$

of formal power series in the non-commuting variables $x_1, \dots, x_d$. It is an $\mathbb{R}$-algebra with respect to the tensor product, and we denote by $p_k : T((\mathbb{R}^d)) \to (\mathbb{R}^d)^{\otimes k}$ the projection. The algebraic dual of $T((\mathbb{R}^d))$ is the graded $\mathbb{R}$-algebra

$$T(\mathbb{R}^d) = \mathbb{R}\langle x_1, \dots, x_d\rangle$$

of polynomials in the non-commuting variables $x_1, \dots, x_d$. It is the unique free algebra over $x_1, \dots, x_d$.

Notation 2. Given an element $T \in T((\mathbb{R}^d))$, we denote by $T_{i_1 \dots i_k}$ the $(i_1 \dots i_k)$-th entry of the order k element of T. For $y \in \mathbb{R}$, we denote $T_y((\mathbb{R}^d)) = \{T \in T((\mathbb{R}^d)) \mid T_1 = y\}$.

Moreover, it will often be convenient to identify a degree k monomial $x_{i_1} \cdot \dots \cdot x_{i_k}$ with the word $w = \mathbf{i_1} \dots \mathbf{i_k}$ in the alphabet $\{\mathbf{1}, \dots, \mathbf{d}\}$. The number k is called the *length* of w and it is denoted by $|w|$. The degree 0 monomial corresponds to the empty word $\mathbf{e}$. We will write letters in bold in order to distinguish the number 1 from the letter $\mathbf{1} = x_1$. In this way, the product of two words v and w is simply the word obtained by writing

v followed by w, and it is called the *concatenation product*. The natural duality pairing

$$\langle -, - \rangle : T((\mathbb{R}^d)) \times T(\mathbb{R}^d) \to \mathbb{R}$$

is given by $\langle T, \mathbf{i_1} \dots \mathbf{i_k} \rangle = T_{i_1 \dots i_k}$, extended by linearity.

Besides the concatenation of words, there is another product on $T(\mathbb{R}^d)$. It will play a very important role in this paper.

Definition 3. The *shuffle product* of two words v and w is the sum of all order-preserving interleavings of them. It is denoted by $v \sqcup\!\sqcup w$. A more precise, recursive definition can be found in [12, Section 1.4]. Again, the shuffle product can be extended by linearity to $T(\mathbb{R}^d)$. We will sometimes use the notation

$$v^{\sqcup\!\sqcup n} = \underbrace{v \sqcup\!\sqcup \dots \sqcup\!\sqcup v}_{n \text{ times}}.$$

Despite its apparently complicated definition, the shuffle product enjoys good properties. For instance, the space $(T(\mathbb{R}^d), \sqcup\!\sqcup, \mathbf{e})$ is a commutative algebra. Moreover, shuffle behaves nicely with respect to the signatures.

Lemma 4 (Shuffle identity)**.** *If $X : [0,1] \to \mathbb{R}^d$ is a piecewise C^1 path, then*

$$\langle \sigma(X), v \rangle \cdot \langle \sigma(X), w \rangle = \langle \sigma(X), v \sqcup\!\sqcup w \rangle$$

for all words $v, w \in T(\mathbb{R}^d)$.

The shuffle identity is proved in [12, Proof of Corollary 3.5], for a more general class of paths.

Up to now we see that signatures do not fill the whole tensor space $T((\mathbb{R}^d))$, but rather they live in the subset of elements with constant term 1 and satisfying the shuffle identity. This is one of the many possible motivations for the next definition.

Definition 5. We will denote

$$\mathcal{G}(\mathbb{R}^d) = \{T \in T_1((\mathbb{R}^d)) \mid \langle T, v \rangle \cdot \langle T, w \rangle = \langle T, v \sqcup\!\sqcup w \rangle \text{ for all words } v, w \in T(\mathbb{R}^d)\}.$$

The object we have just defined is worth a few remarks. It not only contains the signatures of all piecewise C^1 paths, but it is also a group with respect to the tensor product. This is why its elements are sometimes called *group-like* elements in the literature. $\mathcal{G}(\mathbb{R}^d)$ is not linear, but it is closely related to a linear space.

Definition 6. On $T((\mathbb{R}^d))$ there is a bracketing $[T, S] = TS - ST$. Then we can define $\mathrm{Lie}(\mathbb{R}^d)$ to be the free Lie algebra generated by $x_1, \dots, x_d$, that is, the smallest linear subspace of $T((\mathbb{R}^d))$ that contains $x_1, \dots, x_d$ and is closed with respect to the bracketing.

This Lie algebra and $\mathcal{G}(\mathbb{R}^d)$ are linked by two maps.

Definition 7. Define $\exp : T_0((\mathbb{R}^d)) \to T_1((\mathbb{R}^d))$ by the formal power series

$$\exp(T) = \sum_{n=0}^{\infty} \frac{T^{\otimes n}}{n!}.$$

Not surprisingly, exp has a two-sided inverse $\log : T_1((\mathbb{R}^d)) \to T_0((\mathbb{R}^d))$ defined by

$$\log(S) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(S-1)^n.$$

The two maps restrict to a bijection between $\mathrm{Lie}(\mathbb{R}^d)$ and $\mathcal{G}(\mathbb{R}^d)$. For our purposes, we need to point out that all the definitions we recalled have a truncated version. Namely, one can fix $m \in \mathbb{N}$ and consider

$$T^m(\mathbb{R}^d) = \bigoplus_{k=0}^{m} (\mathbb{R}^d)^{\otimes k},$$

where tensors of order greater than n are set to zero. Inside $T^m(\mathbb{R}^d)$ there are $\mathcal{G}^m(\mathbb{R}^d)$ and $\mathrm{Lie}^m(\mathbb{R}^d)$. The maps exp and log are defined in the same way. In order to avoid confusion, we will write $\exp^{(m)}$ to denote the map $T_0^m((\mathbb{R}^d)) \to T_1^m((\mathbb{R}^d))$.

We need a last definition before we move to our Rough Veronese variety.

Definition 8. A non-empty word w is a *Lyndon word* if, whenever we write $w = pq$ as the concatenation of two nonempty words, we have $w < q$ in the lexicographic order. We denote by $W_{d,m}$ the set of Lyndon words of length at most m in the alphabet $\{\mathbf{1}, \dots, \mathbf{d}\}$.

There exists a unique pair (p, q) of nonempty words such that $w = pq$ and q is minimal with respect to lexicographic order. The *bracketing* of w is $[p, q] = pq - qp$.

We care about Lyndon words because $\mathrm{Lie}^m(\mathbb{R}^d)$ has a basis consisting of all bracketings of Lyndon words of length at most m. Recall that the Möbius function $\mu : \mathbb{N} \to \mathbb{N}$ sends a natural number t to

$$\mu(t) = \begin{cases} 0 & \text{if } t \text{ is divisible by the square of a prime,} \\ 1 & \text{if } t \text{ is the product of an even number of distinct primes,} \\ -1 & \text{if } t \text{ is the product of an odd number of distinct primes.} \end{cases}$$

Then the number of length l Lyndon words in the alphabet $\{\mathbf{1}, \dots, \mathbf{d}\}$, denoted by $\mu_{l,d}$, is

$$\mu_{l,d} = \sum_{t|l} \frac{\mu(t)}{l} d^{\frac{l}{t}},$$

and therefore, as a vector space, $\mathrm{Lie}^m(\mathbb{R}^d)$ has dimension

$$\dim \mathrm{Lie}^m(\mathbb{R}^d) = \sum_{l=1}^{m} \sum_{t|l} \frac{\mu(t)}{l} d^{\frac{l}{t}}.$$

Signatures of rough paths

In this section we introduce the main character of our paper. The way we want to think about rough paths is as a generalization of piecewise differentiable paths. The main reference for rough paths is [7].

Consider a piecewise differentiable path X and let $t \in [0, 1]$. In the definition of k-th signature we can replace indefinite integrals with definite ones. This is the same as restricting X to the sub-interval $[0,t]$, hence we will denote this as $\sigma^{(k)}(X_{|[0,t]})$. As an example,

$$\sigma^{(1)}(X_{|[0,t]})_i = \int_0^t \dot{X}_i(\lambda)d\lambda = X_i(t) - X_i(0).$$

For every k, we notice that $\sigma^{(k)}(X_{|[0,t]})$, as a function of t, is a path $[0,1] \to (\mathbb{R}^d)^{\otimes k}$. If we look at the full signature $\sigma(X_{|[0,t]})$, we get a path $[0,1] \to \mathcal{G}(\mathbb{R}^d)$. Moreover, this $\mathcal{G}(\mathbb{R}^d)$-valued path satisfies a Hölder-like inequality. We will use the symbol $f(t) \lesssim g(t)$ to indicate that there is a constant c such that $f(t) \leq c \cdot g(t)$ for every t.

Lemma 9. *Let $X : [0,1] \to \mathbb{R}^d$ be a piecewise differentiable path and let $k \in \mathbb{N}$. If $s, t \in [0,1]$, then*

$$\left|\sigma^{(k)}(X_{|[0,t]}) - \sigma^{(k)}(X_{|[0,s]})\right| \lesssim |t-s|^k. \tag{1}$$

Proof. Let $S = \sigma^{(k)}(X_{|[0,t]}) - \sigma^{(k)}(X_{|[0,s]})$. Since k is fixed, in order to conclude it is enough to bound every entry of S. By definition

$$\begin{aligned}
|S_{i_1\dots i_k}| &= \left|\int_s^t \int_s^{t_k} \cdots \int_s^{t_3} \int_s^{t_2} \dot{X}_{i_1}(t_1)\cdot \ldots \cdot \dot{X}_{i_k}(t_k) dt_1 \dots dt_k\right| \\
&\leq \int_s^t \int_s^{t_k} \cdots \int_s^{t_3} \int_s^{t_2} \left|\dot{X}_{i_1}(t_1)\cdot \ldots \cdot \dot{X}_{i_k}(t_k)\right| dt_1 \dots dt_k \\
&\leq \sup_{t_1\in[0,1]} |\dot{X}_{i_1}(t_1)| \cdot \ldots \cdot \sup_{t_k\in[0,1]} |\dot{X}_{i_k}(t_k)| \int_s^t \int_s^{t_k} \cdots \int_s^{t_3} \int_s^{t_2} dt_1 \dots dt_k \\
&= \sup_{t_1\in[0,1]} |\dot{X}_{i_1}(t_1)| \cdot \ldots \cdot \sup_{t_k\in[0,1]} |\dot{X}_{i_k}(t_k)| \cdot \frac{|t-s|^k}{k!}.
\end{aligned}$$

Since X is piecewise differentiable, all the suprema are finite. □

So we see that a piecewise differentiable path $X : [0,1] \to \mathbb{R}^d$ induces a path $\sigma(X_{|[0,-]}) : [0,1] \to \mathcal{G}(\mathbb{R}^d)$ satisfying inequality (1). If we want a rough path to be a generalization of a piecewise differentiable path, we can define it in a similar flavor, also allowing different exponents. Recall that $p_k : T((\mathbb{R}^d)) \to (\mathbb{R}^d)^{\otimes k}$ is the projection.

Definition 10. A *rough path* of order m is a path $\mathbf{X} : [0,1] \to \mathcal{G}^m(\mathbb{R}^d)$ such that $|p_k(\mathbf{X}(s)^{-1} \otimes \mathbf{X}(t))| \lesssim |t-s|^{\frac{k}{m}}$ for every $k \in 1, \dots, m$ and every $s, t \in [0,1]$. The inverse in taken in the group $\mathcal{G}^m(\mathbb{R}^d)$.

Following [1, Section 5.4], we will focus on a special subclass of rough paths of order m, indexed by elements of $L \in \mathrm{Lie}^m(\mathbb{R}^d)$.

Definition 11. For $L \in \mathrm{Lie}^m(\mathbb{R}^d)$, consider the path $\mathbf{X}_L : [0,1] \to \mathcal{G}^m(\mathbb{R}^d)$ sending t to $\exp^{(m)}(tL)$. By [7, Exercise 9.17], this is indeed an order m rough path, and we define its *signature* $\sigma(\mathbf{X}_L) = \exp(L) \in \mathcal{G}(\mathbb{R}^d)$.

The relation between the signature defined by iterated integrals and the signature of a rough path is also pointed out in [7, Exercise 9.17].

We want to study the set parameterizing the k-th signatures of $\mathbf{X}_L$, when L ranges over $\mathrm{Lie}^m(\mathbb{R}^d)$. Such set is the image of $p_k \circ \exp$. A priori, this is just a semialgebraic subset of $(\mathbb{R}^d)^{\otimes k}$, that is, it is described by a finite number of polynomial equations and inequalities. Semialgebraic sets are usually hard to handle. In order to make our analysis simpler, we will follow a common approach in applied algebraic geometry and take the Zariski closure of this set, which means that we only look at the equations without considering the inequalities. Furthermore, from a geometric viewpoint it is convenient to work over an algebraically closed field, so we look at the variety that these equations define in $(\mathbb{C}^d)^{\otimes k}$, instead of $(\mathbb{R}^d)^{\otimes k}$. Finally, we want to work up to scalar multiples, so we pass to the projectivization and hence we deal with a projective variety.

Definition 12. The *Rough Veronese variety* $\mathcal{R}_{d,k,m}$ is the closure of the image of the composition

$$f_{d,k,m}: \mathrm{Lie}^m(\mathbb{R}^d) \xrightarrow{\exp} \mathcal{G}(\mathbb{R}^d) \xrightarrow{p_k} (\mathbb{R}^d)^{\otimes k} \to (\mathbb{R}^d)^{\otimes k} \otimes \mathbb{C} = (\mathbb{C}^d)^{\otimes k} \twoheadrightarrow \mathbb{P}^{d^k-1}.$$

First of all observe that, being the image of a morphism, $\mathcal{R}_{d,k,m}$ is irreducible. There are several reasons to compare $\mathcal{R}_{d,k,m}$ to a Veronese variety. Since $\mathrm{Lie}^1(\mathbb{R}^d) = \mathbb{R}^d$, an element $L \in \mathrm{Lie}^1(\mathbb{R}^d)$ is just a vector and therefore, up to a multiplicative constant, $p_k(\exp(L)) = L^{\otimes k}$ can be viewed as the Veronese embedding of $\mathbb{R}^d$ into $\mathrm{Sym}^k\,\mathbb{R}^d \subset (\mathbb{R}^d)^{\otimes k}$. In other words, $\mathcal{R}_{d,k,1} = \mathcal{V}_{d-1,k}$. Moreover, $\mathcal{V}_{d,k}$ is toric, defined by all degree k monomials in $d+1$ variables. We will see that $\mathcal{R}_{d,k,m}$ is toric as well, and it is defined by monomials of weighted degree k, for a suitable choice of weights imposed by the structure of $\mathrm{Lie}^m(\mathbb{R}^d)$. Unlike $\mathcal{V}_{d-1,k}$, however, in general $\mathcal{R}_{d,k,m}$ fails to be smooth.

The inclusions $\mathrm{Lie}^i(\mathbb{R}^d) \subset \mathrm{Lie}^{i+1}(\mathbb{R}^d)$ show that the Rough Veronese varieties are nested. On the other hand, this chain stabilizes. Indeed, when we apply p_k and project onto the order k summand, we do not see anything of order greater than k. So $\mathcal{R}_{d,k,k+i} = \mathcal{R}_{d,k,k}$ for every $i \in \mathbb{N}$. Being the image of $\mathcal{G}^k(\mathbb{R}^d)$ under p_k, $\mathcal{R}_{d,k,k}$ contains all the k-th signatures of piecewise C^1 paths. For this reason, in [1, Section 4.3] it is called *universal variety* and denoted by $\mathcal{U}_{d,k}$. Summarizing, there is a chain of strict inclusions

$$\mathcal{V}_{d-1,k} = \mathcal{R}_{d,k,1} \subset \mathcal{R}_{d,k,2} \subset \ldots \subset \mathcal{R}_{d,k,k} = \mathcal{R}_{d,k,k+1} = \ldots = \mathcal{U}_{d,k}.$$

It is not restrictive to assume $d \geq 2$ and $m \leq k$. The first thing we want to do is to determine $\dim \mathcal{R}_{d,k,m}$.

Proposition 13. *If $m \leq k$, then the map $f_{d,k,m}: \mathrm{Lie}^m(\mathbb{R}^d) \to \mathbb{P}^{d^k-1}$ is generically k to 1. In particular,* $\dim \mathcal{R}_{d,k,m} = \dim(\mathbb{P}(\mathrm{Lie}^m(\mathbb{R}^d)))$.

This was already noted in [1, Remark 6.5]. We now want to understand the geometry of $\mathcal{R}_{d,k,m}$, and we start by looking at the simplest example.

Example 14. Consider $d = k = m = 2$. We want to write down $f_{2,2,2}$. The Lyndon words of length at most 2 in the alphabet $\{\mathbf{1}, \mathbf{2}\}$ are $\mathbf{1}$, $\mathbf{2}$ and $\mathbf{12}$, hence $\dim \mathrm{Lie}^2(\mathbb{R}^2) = 3$. An element of $\mathrm{Lie}^2(\mathbb{R}^2)$ can therefore be written as $x_1\mathbf{1} + x_2\mathbf{2} + a(\mathbf{12} - \mathbf{21})$. If we look at it in tensor terms, we see a vector $x = (x_1, x_2)$ as order 1 summand, and a 2×2 matrix A as order 2 summand.

Being a multiple of $(\mathbf{12}-\mathbf{21})$, A is skew-symmetric. Then

$$\begin{aligned} p_2(\exp(x+A)) &= p_2\left(1+x+A+\frac{(x+A)^2}{2}+\dots\right) \\ &= p_2\left(1+x+A+\frac{x^2+xA+Ax+A^2}{2}+\dots\right) \\ &= A+\frac{x^2}{2} = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x_1^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{pmatrix}. \end{aligned}$$

In coordinates, the map $f_{2,2,2} : \mathbb{R}^3 \to \mathbb{P}^3$ is

$$(x_1,x_2,a) \mapsto \begin{bmatrix} \frac{x_1^2}{2} & \frac{x_1x_2}{2}+a \\ \frac{x_1x_2}{2}-a & \frac{x_2^2}{2} \end{bmatrix}.$$

Up to a linear change of coordinates, it becomes

$$(x_1,x_2,a) \mapsto \begin{bmatrix} x_1^2 & x_1x_2 \\ a & x_2^2 \end{bmatrix}.$$

We can make several important remarks. First of all, the map is now defined by monomials, and so $\mathcal{R}_{2,2,2}$ is a toric variety. It is a cone over the Veronese variety $\mathcal{V}_{1,2}$ and it spans the whole $\mathbb{P}^3$. Its ideal is generated by a quadric polynomial. Finally, it can be seen as the embedding

$$\mathbb{P}(1,1,2) \hookrightarrow \mathbb{P}^3$$

of a weighted projective plane, defined by all monomials of (weighted) degree 2.

Our main goal is to generalize these remarks to all values of d,k,m.

$\mathcal{R}_{d,k,m}$ AS A TORIC VARIETY

Roughly speaking, a variety is toric if it is the image of a monomial map. A toric variety not only has nice properties - for instance, it is irreducible, rational and its ideal is generated by binomials - but it can be associated to a polytope that completely encodes its geometry. This makes toric varieties accessible from a theoretical, combinatorial and computational viewpoint. A good reference on toric varieties is [4]. This section is devoted to prove that the Rough Veronese is indeed toric and to provide an explicit way to make computations on it.

If we write an element $L \in \mathrm{Lie}^m(\mathbb{R}^d)$ as the sum $L = L_1 + \ldots + L_m$ of terms of order $1,\ldots,m$, then $p_k(\exp(L)) \in (\mathbb{R}^d)^{\otimes k}$ is a linear combination of all possible ways to get an order k tensor by multiplying $L_1,\ldots,L_m$. If we consider coordinates $T_{i_1\ldots i_k}$ on $\mathbb{P}^{d^k-1}$, we can rephrase this observation by saying that every coordinate $T_{i_1\ldots i_k}$ of $f_{d,k,m}$ is a linear combination of weighted degree k monomials. The weight of a variable corresponding to a length i Lyndon word is i. We can define

$$g_{d,k,m} : \mathrm{Lie}^m(\mathbb{R}^d) \to \mathbb{P}^{d^k-1}$$

by using all such weighted monomials. By our observation, there is a linear change of coordinates sending the image of $g_{d,k,m}$ to the image of $f_{d,k,m}$. We want to prove that such change of coordinates is invertible, that is, every weighted monomial can be obtained as a linear combination of the entries of

$f_{d,k,m}$, and we also want to do that as explicitly as possible. This section is based on [5, Section 3] and [8, Section IV]. Let us start with two definitions.

Definition 15. Let v be a word. If $v = \mathbf{e}$ is the empty word, set $S_{\mathbf{e}} = \mathbf{e}$. Otherwise, we define S_v in the following recursive way.

(1) If v is Lyndon, write $v = l \cdot w$, where l is a letter, and define $S_v = l \cdot S_w$.
(2) Otherwise, write $v = w_1^{i_1} \cdot \ldots \cdot w_k^{i_k}$ as concatenation of decreasing Lyndon words. This can be done uniquely by [12, Section 7.4]. Define
$$S_v = \frac{S_{w_1}^{\sqcup\!\sqcup i_1} \sqcup\!\sqcup \ldots \sqcup\!\sqcup S_{w_k}^{\sqcup\!\sqcup i_k}}{i_1! \ldots i_k!}.$$

The next ingredient we need is the following.

Definition 16. Define a linear map $\psi : T(\mathbb{R}^d) \to T(\mathbb{R}^d)$ that acts on a word v by
$$v \mapsto \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sum_{\substack{u_1,\ldots,u_n \\ \text{nonempty} \\ \text{words}}} \langle v, u_1 \cdot \ldots \cdot u_n \rangle u_1 \sqcup\!\sqcup \ldots \sqcup\!\sqcup u_n.$$

In [8], the map ψ is called π_1', while it appears in [5] as $\pi_1^\top$. Observe that only finitely many terms of the sum are non-zero. Indeed, an element in $T(\mathbb{R}^d)$ is the linear combination of finitely many words, and if v is a word then there are only finitely many ways to write it as a concatenation $u_1, \ldots, u_n$. Moreover, if v is a word of length l and $v = u_1, \ldots, u_n$, then $u_1 \sqcup\!\sqcup \ldots \sqcup\!\sqcup u_n$ is a sum of length l words. This means that ψ preserves the grading of $T(\mathbb{R}^d)$.

The following result about the exponential map will be of great help. Recall that $W_{d,m}$ is the set of Lyndon words of length at most m in the alphabet with d letters.

Lemma 17. *Let $\{P_w \mid w \in W_{d,m}\}$ be the basis of* $\mathrm{Lie}^m(\mathbb{R}^d)$ *indexed by the Lyndon words and let $w_1, \ldots, w_r \in W_{d,m}$. Then*
$$\left\langle \exp\left(\sum_{w \in W_{d,m}} \alpha_w P_w\right), \psi(S_{w_1}) \sqcup\!\sqcup \ldots \sqcup\!\sqcup \psi(S_{w_r}) \right\rangle = \alpha_{w_1} \cdot \ldots \cdot \alpha_{w_r}.$$

Proof. By [5, Section 3], for every $i \in \{1, \ldots, r\}$ we have
$$\left\langle \exp\left(\sum_{w \in W_{d,m}} \alpha_w P_w\right), \psi(S_{w_i}) \right\rangle = \alpha_{w_i}.$$
See also [8, Theorem 1]. Now thesis follows by Lemma 4. □

Clearly every entry of $f_{d,k,m}$ is a linear combination of monomials. Lemma 17 shows that the converse holds. Every monomial can be obtained as linear combination of entries of $f_{d,k,m}$. In other words, we can use the map ψ to build the linear forms we needed to pass from $f_{d,k,m}$ to $g_{d,k,m}$, allowing us to identify them. Let us summarize the conclusion.

Proposition 18. *For every $w \in W_{d,m}$, define a variable x_w and assign it the weight $|w|$. Then, up to a linear change of coordinates, $f_{d,k,m}$ is defined by all monomials of weighed degree k in the variables $\{x_w \mid w \in W_{d,m}\}$. More explicitly, if we set*

$$J = \{(w_1, \ldots, w_r) \mid r \in \mathbb{N}, w_i \in W_{d,m} \text{ and } |w_1| + \ldots + |w_r| = k\},$$

then $\mathcal{R}_{d,k,m}$ is isomorphic to the image of

$$(x_w)_{w \in W_{d,m}} \mapsto [x_{w_1} \cdot \ldots \cdot x_{w_r}]_{(w_1,\ldots,w_r)\in J}.$$

In particular, it is a toric variety.

Notice that the map defined in Proposition 18 is a nondegenerate embedding in a possibly proper linear subspace of $\mathbb{P}^{d^k-1}$. We will discuss the linear span of $\mathcal{R}_{d,k,m}$ in Lemma 22.

For the purpose of practical applications, we want to explicitly describe a linear change of coordinates in $\mathbb{P}^{d^k-1}$ that makes $\mathcal{R}_{d,k,m}$ toric. If $m < k$, then every element of $L \in \mathrm{Lie}^{m+1}(\mathbb{R}^d)$ can be written uniquely as $L = (L_1, L_2)$, where $L_1 \in \mathrm{Lie}^m(\mathbb{R}^d)$. In this case, $f_{d,k,m}(L_1) = f_{d,k,m+1}(L_1, 0)$. This means that every change of coordinates that makes $f_{d,k,m+1}$ a monomial map also makes $f_{d,k,m}$ monomial. Therefore we can assume $m = k$, because the change of coordinates that will make $\mathcal{R}_{d,k,k}$ toric will also work on $\mathcal{R}_{d,k,m}$ for every $m \leq k$.

In the notation of Proposition 18, the change of coordinates in $\mathbb{P}^{d^k-1}$ is

$$T \mapsto [\langle T, \psi(S_{w_1}) \sqcup\!\sqcup \ldots \sqcup\!\sqcup \psi(S_{w_r})\rangle]_{(w_1,\ldots,w_r)\in J}\,. \tag{2}$$

This is indeed well defined because, as we will show in Proposition 23, J has exactly d^k elements when $m = k$.

Example 19. Let us consider $\mathcal{U}_{2,4} = \mathcal{R}_{2,4,4}$. The Lyndon words are

$$W_{2,4} = \{\mathbf{1}, \mathbf{2}, \mathbf{12}, \mathbf{112}, \mathbf{122}, \mathbf{1112}, \mathbf{1122}, \mathbf{1222}\}.$$

By using Definition 15, we get $S_w = w$ for every $w \in W_{2,4}$. The first 5 entries of the change of variables (2) are then

$$\begin{aligned}\langle T, \psi(S_{\mathbf{i}}) \sqcup\!\sqcup \psi(S_{\mathbf{j}}) \sqcup\!\sqcup \psi(S_{\mathbf{k}}) \sqcup\!\sqcup \psi(S_{\mathbf{l}})\rangle &= \langle T, \psi(\mathbf{i}) \sqcup\!\sqcup \psi(\mathbf{j}) \sqcup\!\sqcup \psi(\mathbf{k}) \sqcup\!\sqcup \psi(\mathbf{l})\rangle \\ &= \langle T, \mathbf{i} \sqcup\!\sqcup \mathbf{j} \sqcup\!\sqcup \mathbf{k} \sqcup\!\sqcup \mathbf{l}\rangle \\ &= \sum_{\sigma} T_{\sigma(i)\sigma(j)\sigma(k)\sigma(l)},\end{aligned}$$

where $1 \leq i \leq j \leq k \leq l \leq 2$ and σ ranges among the permutations of $\{i, j, k, l\}$. The next 3 entries are

$$\begin{aligned}\langle T, \psi(S_{\mathbf{12}}) \sqcup\!\sqcup \psi(S_{\mathbf{1}}) \sqcup\!\sqcup \psi(S_{\mathbf{1}})\rangle &= \langle T, \psi(\mathbf{12}) \sqcup\!\sqcup \psi(\mathbf{1}) \sqcup\!\sqcup \psi(\mathbf{1})\rangle \\ &= \langle T, \frac{1}{2}(\mathbf{12} - \mathbf{21}) \sqcup\!\sqcup \mathbf{1} \sqcup\!\sqcup \mathbf{1}\rangle \\ &= \langle T, 3 \cdot \mathbf{1112} + \mathbf{1121} - \mathbf{1211} - 3 \cdot \mathbf{2111}\rangle \\ &= 3T_{1112} + T_{1121} - T_{1211} - 3T_{2111}\end{aligned}$$

and, in a similar way,

$$\begin{aligned}\langle T, \psi(S_{\mathbf{12}}) \sqcup\!\sqcup \psi(S_{\mathbf{1}}) \sqcup\!\sqcup \psi(S_{\mathbf{2}})\rangle &= 2T_{1122} + T_{1212} - T_{2121} - 2T_{2211}, \\ \langle T, \psi(S_{\mathbf{12}}) \sqcup\!\sqcup \psi(S_{\mathbf{2}}) \sqcup\!\sqcup \psi(S_{\mathbf{2}})\rangle &= 3T_{1222} + T_{2122} - T_{2212} - 3T_{2221}.\end{aligned}$$

We go on with

$$\begin{aligned}\langle T, \psi(S_{\mathbf{112}}) \sqcup\!\sqcup \psi(S_{\mathbf{1}})\rangle &= \langle T, \psi(\mathbf{112}) \sqcup\!\sqcup \psi(\mathbf{1})\rangle \\ &= \langle T, \frac{1}{6}(\mathbf{112} - 2\cdot\mathbf{121} + \mathbf{211}) \sqcup\!\sqcup \mathbf{1}\rangle \\ &= \langle T, \frac{1}{2}(\mathbf{1112} - \mathbf{1121} - \mathbf{1211} + \mathbf{2111})\rangle \\ &= \frac{1}{2}(T_{1112} - T_{1121} - T_{1211} + T_{2111})\end{aligned}$$

and, in a similar way,

$$\begin{aligned}\langle T, \psi(S_{\mathbf{112}}) \sqcup\!\sqcup \psi(S_{\mathbf{2}})\rangle &= \frac{1}{3}(T_{1122} + T_{2112} + T_{2211}) - \frac{1}{6}(T_{1212} + T_{2121} + 4T_{2211}), \\ \langle T, \psi(S_{\mathbf{122}}) \sqcup\!\sqcup \psi(S_{\mathbf{1}})\rangle &= \frac{1}{3}(T_{1122} + T_{1221} + T_{2211}) - \frac{1}{6}(T_{1212} + T_{2121} + 4T_{2112}), \\ \langle T, \psi(S_{\mathbf{122}}) \sqcup\!\sqcup \psi(S_{\mathbf{2}})\rangle &= \frac{1}{2}(T_{1222} - T_{2122} - T_{2212} + T_{2221}).\end{aligned}$$

The next entry is

$$\begin{aligned}\langle T, \psi(S_{\mathbf{12}}) \sqcup\!\sqcup \psi(S_{\mathbf{12}})\rangle &= \langle T, \psi(\mathbf{12}) \sqcup\!\sqcup \psi(\mathbf{12})\rangle \\ &= \langle T, \frac{1}{2}(\mathbf{12} - \mathbf{21}) \sqcup\!\sqcup \frac{1}{2}(\mathbf{12} - \mathbf{21})\rangle \\ &= \langle T, \mathbf{1122} + \mathbf{2211} - \mathbf{2112} - \mathbf{1221} + \frac{1}{2}(\mathbf{1212} - \mathbf{2121})\rangle \\ &= T_{1122} - T_{1221} - T_{2112} + T_{2211}.\end{aligned}$$

Finally we consider length 4 Lyndon words.

$$\langle T, \psi(S_{\mathbf{1112}})\rangle = \langle T, \psi(\mathbf{1112})\rangle = \langle T, \frac{1}{6}(\mathbf{1211} - \mathbf{1121})\rangle = \frac{1}{6}(T_{1211} - T_{1121})$$

and, in the same way,

$$\begin{aligned}\langle T, \psi(S_{\mathbf{1122}})\rangle &= \frac{1}{6}(T_{1122} - T_{1212} + T_{2121} - T_{2211}), \\ \langle T, \psi(S_{\mathbf{1222}})\rangle &= \frac{1}{6}(T_{2212} - T_{2122}).\end{aligned}$$

Compare our computations to the ones in [8, Section IV]. Our change of coordinates in $\mathbb{P}^{15}$ sends T to

$$\begin{bmatrix} 24T_{1111} \\ 6(T_{1112}+T_{1121}+T_{1211}+T_{2111}) \\ 4(T_{1122}+T_{1212}+T_{1221}+T_{2211}+T_{2112}+T_{2121}) \\ 6(T_{1222}+T_{2122}+T_{2212}+T_{2221}) \\ 24T_{2222} \\ 3T_{1112}+T_{1121}-T_{1211}-3T_{2111} \\ 2T_{1122}+T_{1212}-T_{2121}-2T_{2211} \\ 3T_{1222}+T_{2122}-T_{2212}-3T_{2221} \\ \frac{1}{2}(T_{1112}-T_{1121}-T_{1211}+T_{2111}) \\ \frac{1}{3}(T_{1122}+T_{2112}+T_{2211})-\frac{1}{6}(T_{1212}+T_{2121}+4T_{1221}) \\ \frac{1}{3}(T_{1122}+T_{1221}+T_{2211})-\frac{1}{6}(T_{1212}+T_{2121}+4T_{2112}) \\ \frac{1}{2}(T_{1222}-T_{2122}-T_{2212}+T_{2221}) \\ T_{1122}-T_{1221}-T_{2112}+T_{2211} \\ \frac{1}{6}(T_{1211}-T_{1121}) \\ \frac{1}{6}(T_{1122}-T_{1212}+T_{2121}-T_{2211}) \\ \frac{1}{6}(T_{2212}-T_{2122}) \end{bmatrix}.$$

Example 20. Let us consider $\mathcal{U}_{3,3}=\mathcal{R}_{3,3,3}$. Then

$$W_{3,3}=\{\mathbf{1},\mathbf{2},\mathbf{3},\mathbf{12},\mathbf{13},\mathbf{23},\mathbf{112},\mathbf{113},\mathbf{122},\mathbf{133},\mathbf{223},\mathbf{233},\mathbf{123},\mathbf{132}\}.$$

By using Definition 15, we get $S_w=w$ for every $w\in W_{3,3}\setminus\{\mathbf{132}\}$, while $S_{\mathbf{132}}=\mathbf{123}+\mathbf{132}$. The first 10 entries of the change of variables (2) are then

$$\begin{aligned}\langle T,\psi(S_{\mathbf{i}})\sqcup\psi(S_{\mathbf{j}})\sqcup\psi(S_{\mathbf{k}})\rangle &= \langle T,\psi(\mathbf{i})\sqcup\psi(\mathbf{j})\sqcup\psi(\mathbf{k})\rangle\\ &= \langle T,\mathbf{i}\sqcup\mathbf{j}\sqcup\mathbf{k}\rangle\\ &= \langle T,\mathbf{kij}+\mathbf{ikj}+\mathbf{ijk}+\mathbf{kji}+\mathbf{jki}+\mathbf{jik}\rangle\\ &= T_{kij}+T_{ikj}+T_{ijk}+T_{kji}+T_{jki}+T_{jik},\end{aligned}$$

for $1\le i\le j\le k\le 3$. Then we have 9 more entries of the form

$$\begin{aligned}\langle T,\psi(S_{\mathbf{i}})\sqcup\psi(S_{\mathbf{jk}})\rangle &= \langle T,\psi(\mathbf{i})\sqcup\psi(\mathbf{jk})\rangle\\ &= \langle T,\mathbf{i}\sqcup\frac{1}{2}(\mathbf{jk}-\mathbf{kj})\rangle\\ &= \langle T,\frac{1}{2}(\mathbf{ijk}+\mathbf{jik}+\mathbf{jki}-\mathbf{ikj}-\mathbf{kij}-\mathbf{kji})\rangle\\ &= \frac{1}{2}(T_{ijk}+T_{jik}+T_{jki}-T_{ikj}-T_{kij}-T_{kji}),\end{aligned}$$

for $i\in\{1,2,3\}$ and $1\le j<k\le 3$. Finally, the length 3 Lyndon words give the last 8 entries as

$$\begin{aligned}\langle T,\psi(S_{\mathbf{132}})\rangle &= \langle T,\psi(\mathbf{123}+\mathbf{132})\rangle\\ &= \langle T,\psi(\mathbf{123})+\psi(\mathbf{132})\rangle\\ &= \langle T,\frac{1}{6}(\mathbf{123}+\mathbf{132}-2\cdot\mathbf{213}+\mathbf{231}-2\cdot\mathbf{312}+\mathbf{321})\rangle\\ &= \frac{1}{6}(T_{123}+T_{132}+T_{231}+T_{321})-\frac{1}{3}(T_{213}+T_{312})\end{aligned}$$

and

$$\begin{aligned}\langle T, \psi(S_{\mathbf{ijk}})\rangle &= \langle T, \psi(\mathbf{ijk})\rangle \\ &= \langle T, \frac{1}{6}(2\cdot \mathbf{ijk} + 2\cdot \mathbf{kji} - \mathbf{kij} - \mathbf{ikj} - \mathbf{jki} - \mathbf{jik})\rangle \\ &= \frac{1}{3}(T_{ijk} + T_{kji}) - \frac{1}{6}(T_{kij} + T_{ikj} + T_{jki} + T_{jik}),\end{aligned}$$

for $\mathbf{ijk} \in \{\mathbf{112}, \mathbf{113}, \mathbf{122}, \mathbf{133}, \mathbf{223}, \mathbf{233}, \mathbf{123}\}$. It is interesting to point out that some of these linear forms are used in [11, Section 5] in order to recover the path of a given signature.

Remark 21. The universal variety contains many interesting subvarieties besides $\mathcal{R}_{d,k,m}$. Examples include the signature variety $\mathcal{L}_{d,k,m}$ of piecewise linear paths with m steps and its subvariety $\mathcal{A}_{\nu,k}$ of axis-parallel paths, both studied in [1]. The change of coordinates given by Lemma 17 proves that $\mathcal{R}_{d,k,m}$ and therefore $\mathcal{U}_{d,k}$ are toric, but it does not necessarily work as nicely with other subvarieties. For instance, a computation with the software Macaulay2 ([9]) shows that, after our change of coordinates, the ideals of both $\mathcal{L}_{2,3,2}$ and $\mathcal{A}_{1212,3}$ are not generated by binomials and therefore they are not toric.

Further geometric properties

In this section we will generalize the remarks we made for Example 14. While $\mathcal{R}_{2,2,2} \subset \mathbb{P}^3$ is nondegenerate, for other values of d, k, m the Rough Veronese may be contained in a smaller linear subspace, as we already observed in Proposition 18.

Proposition 22. *The affine dimension of the linear span of $\mathcal{R}_{d,k,m} \subset \mathbb{P}^{d^k-1}$ is*

$$\sum_{\lambda\vdash k, \lambda_1\le m}\left(\prod_{i=1}^{m}\binom{\mu_{i,d} + \sharp\{j \mid \lambda_j = i\} - 1}{\mu_{i,d}-1}\right).$$

This number equals the coefficient of t^k of the expansion of the generating function

$$\prod_{i=1}^{m}\frac{1}{(1-t^i)^{\mu_{i,d}}}.$$

Proof. As in Proposition 18, we define a variable x_w for every $w \in W_{d,m}$ and we assign it the weight $|w|$. Given a weighted degree k monomial $\varphi \in \mathbb{C}[x_w \mid w \in W_{d,m}]$, we can write it in reverse lexicographic order as a string of possibly repeated variables. Define a partition λ of k by

$$\lambda_j = i \Leftrightarrow \text{ the } j\text{-th entry of the string is } x_w \text{ for some } |w| = i.$$

Now φ is the product of monomials $\varphi = \varphi_1 \cdot \ldots \cdot \varphi_m$, where φ_i is a monomial in $\mathbb{C}[x_w \mid |w| = i]$. The degree of φ_i is the number of times i appears as an entry of λ. Hence, for every φ_i there are $\binom{\mu_{i,d}+\sharp\{j\mid\lambda_j=i\}-1}{\mu_{i,d}-1}$ choices. For the generating function, see [13, Section 1.8]. □

Nonetheless, the universal variety $\mathcal{U}_{d,k} = \mathcal{R}_{d,k,k}$ is indeed nondegenerate for every k.

Proposition 23. *Under our assumption $k \geq m$, the Rough Veronese $\mathcal{R}_{d,k,m} \subset \mathbb{P}^{d^k-1}$ is nondegenerate if and only if $m = k$.*

Proof. Because of the chain of strict inclusions $\mathcal{R}_{d,k,m} \subsetneq \mathcal{R}_{d,k,m+1}$, it is enough to show that $\mathcal{R}_{d,k,k}$ is nondegenerate.

Thanks to Proposition 18, we only have to prove that there are d^k distinct monomials of weighted degree k. Corollary 22 may be difficult to apply, so instead we want to define a bijection between these monomials and the set of length k words in the alphabet with d letters. Since our variables are indexed by Lyndon words, we can think of a monomial as a product of Lyndon words such that the sum of the lengths is k. Observe that the monomial remains the same if we permute the variables, i.e. the Lyndon words. In order to avoid redundancy, we can fix an order. Now we only have to prove that, after fixing an order among Lyndon words, every length k word in the alphabet with d letters can be written uniquely as an ordered product of Lyndon words whose lengths sum to k. This is a well-known fact, and a proof can be found for instance in [12, Section 7.4]. □

The next result shows another feature of rough paths. Not only does the universal variety $\mathcal{U}_{d,k}$ coincide with the last Rough Veronese $\mathcal{R}_{d,k,k}$, but its structure is already determined by the second to last one, $\mathcal{R}_{d,k,k-1}$.

Proposition 24. *The universal variety $\mathcal{U}_{d,k} = \mathcal{R}_{d,k,k}$ is a cone over $\mathcal{R}_{d,k,k-1}$ with vertex $\mathbb{P}^{\mu_{k,d}-1}$. The preimage of the vertex is the vector subspace of $\mathrm{Lie}^k(\mathbb{R}^d)$ defined by the vanishing of the first $k-1$ entries.*

Proof. Let V be the vector subspace of $\mathrm{Lie}^k(\mathbb{R}^d)$ defined by the vanishing of the first $k-1$ entries. Thanks to Proposition 18, we already know that $\mathcal{R}_{d,k,k}$ is defined by the degree k weighted monomials. Exactly $\mu_{k,d}$ of them are $\{x_w \mid w \in W_{d,m} \text{ and } |w| = k\}$. In other words, the map $p_k \circ \exp$ restricts to the identity on V. It follows that

$$p_k(\exp(\mathrm{Lie}^k(\mathbb{R}^d))) = p_k(\exp(\mathrm{Lie}^{k-1}(\mathbb{R}^d))) \times V$$

is a cylinder. When we pass to the projectivization, we get a cone with base $\mathcal{R}_{d,k,k-1}$ and vertex $\mathbb{P}(p_k(\exp(V))) \cong \mathbb{P}(V)$. □

Remark 25. It is interesting to try to classify the paths whose signatures lie in the vertex of the universal variety. Equivalently, we wonder what it means for a rough path to have zeroes in the first $k-1$ entries. It is straightforward to check that for a piecewise differentiable path X, the first signature

$$\sigma^{(1)}(X) = X(1) - X(0)$$

is just the vector joining the endpoint of X to its starting point. The second signature also has a geometric interpretation: if we take the projection of X onto the i, j plane, the signed area of the region bounded by X and the segment between $X(0)$ and $X(1)$ is $\sigma^{(2)}_{ij}(X)$. For instance, if X is a loop then $\sigma^{(1)}(X) = 0$ but $\sigma^{(2)}(X)$ may be nonzero. However, if we allow not only piecewise differentiable paths, but also rough paths, we find more interesting examples. There are rough paths $\mathbf{X}$ such that $p_1(\mathbf{X}(t)) = p_1(\mathbf{X}(0))$ for every t, but with a nonzero second entry. One example of these *pure area paths* is [6, Exercise 2.17], where it is built as a limit of smaller and smaller loops.

In the same fashion, we can think of the vanishing of the first two signatures as a loop with some symmetry that make the signed area zero, such as two circles in $\mathbb{R}^2$ meeting at a point. Then again, there are more exotic examples of rough paths with the same property. However, the geometric meaning of the third and higher order signatures is not yet clear, as it is an open research area.

We now turn our attention to the ideal of $\mathcal{R}_{d,k,m}$. In [11, Section 4], it is proved that the ideal of $\mathcal{R}_{d,3,3}$ is generated by the 2×2 minors of a suitable Henkel matrix, in particular by quadrics. In many other examples this ideal is generated in degree 2. However, this is not true in general, not even if $k = m$. The following counterexample, suggested by Michałek, answers a question posed in [1, Section 4.3].

Proposition 26. *For $14 \leq m \leq k = 20$, the ideal of the Rough Veronese variety $\mathcal{R}_{d,20,m}$ is not generated by quadrics.*

Proof. Let N be the number of weighted degree k monomials. By Proposition 18, up to change of coordinates $\mathcal{R}_{d,20,m}$ is the image of a map $\mathrm{Lie}^m(\mathbb{R}^d) \to \mathbb{P}^{N-1}$. The coordinates T of the target space are indexed by the N weighted monomials. For every $i \in \{1,\dots,m\}$, let $x^{(i)}$ be one of the variables of weight i. Let I be the ideal of $\mathcal{R}_{d,20,m}$ and define

$$\begin{aligned} t_1 &= T_{x^{(1)}x^{(9)}x^{(10)}}, & t_2 &= T_{x^{(5)}x^{(7)}x^{(8)}},\\ t_3 &= T_{x^{(2)}x^{(4)}x^{(14)}}, & t_4 &= T_{x^{(1)}x^{(5)}x^{(14)}},\\ t_5 &= T_{x^{(4)}x^{(7)}x^{(9)}}, & t_6 &= T_{x^{(2)}x^{(8)}x^{(10)}}. \end{aligned}$$

Then $f = t_1t_2t_3 - t_4t_5t_6$ is a degree 3 element of I. Let us show that f is not generated by quadrics. Let $q_1,\dots,q_r$ be the degree two generators of I. Since $\mathcal{R}_{d,20,m}$ is toric, they are binomials and so we can write $q_i = g_i - h_i$, where each g_i and each h_i is a degree 2 monomial. Assume by contradiction that f can be algebraically generated by $q_1,\dots,q_r$. Then there exist linear monomials $l_1,\dots,l_r$ such that, up to order, f is a sum of nonzero polynomials $f = l_1(g_1 - h_1) + \cdots + l_t(g_r - h_r)$. It follows that each term in the sum is a multiple of some g_i. For instance, there is a degree 2 monomial dividing $t_1t_2t_3$. The only three possibilities are t_1t_2, t_1t_3 and t_2t_3. Suppose then that $g_i = t_1t_2$. Since $g_i - h_i \in I$, the product of the two monomials indexing the variables of h_i equals t_1t_2. But it is easy to see that $1+9+10$ and $5+7+8$ are the only ways to obtain 20 by sums of non-repeated elements of $\{1,5,7,8,9,10\}$. Hence $g_i - h_i = 0$, a contradiction. $\square$

Among the features of $\mathcal{R}_{2,2,2}$ we pointed out in Example 14, there is one we still have to check. We saw that $\mathcal{R}_{2,2,2}$ is the embedding of $\mathbb{P}(1,1,2)$ given by its weighted quadrics. If we define the sequence of weights

$$s = (\underbrace{1,\dots,1}_{\mu_{1,d}\text{ times}}, \underbrace{2,\dots,2}_{\mu_{2,d}\text{ times}}, \dots, \underbrace{m,\dots,m}_{\mu_{m,d}\text{ times}}),$$

by Proposition 18 $f_{d,k,m}$ always gives a rational map $\mathbb{P}(s) \dashrightarrow \mathbb{P}^{d^k-1}$, that we will still call in the same way by abuse of notation. However, $f_{d,k,m}$ does not need to be an embedding. Actually, it does not even need to be defined everywhere. For instance, $f_{2,3,2} : \mathbb{P}(1,1,2) \dashrightarrow \mathbb{P}^7$ is defined by

$$[x_1, x_2, a] \mapsto \left[\begin{array}{cc} x_1^3 & x_1^2x_2 \\ x_1x_2^2 & x_2^3 \\ \hline x_1a & x_2a \\ 0 & 0 \end{array}\right],$$

so $[0,0,1]$ is a base point. This is a general behavior.

Proposition 27. *The map $f_{d,k,m} : \mathbb{P}(s) \dashrightarrow \mathbb{P}^{d^k-1}$ is base point free if and only if every entry of w divides k.*

Proof. Assume that k is a multiple of every entry of s. Then for every variable x_w, there is a power of x_w appearing among the monomials of weighted degree k. Therefore the only way for all weighted monomials to vanish is setting all variables to zero. This means there are no base points.

On the other hand, assume that k is not a multiple of one of the entries of s, say i, and consider a variable $x^{(i)}$ of weight i. Since $i \nmid k$, no power of $x^{(i)}$ appears among the monomials defining $f_{d,k,m}$. Then $[0,\dots,0,1,0\dots,0] \in \mathbb{P}(s)$, with a 1 in the entry corresponding to $x^{(i)}$, is a base point for $f_{d,k,m}$. □

Table of invariants

We collect some of the geometric invariants of $\mathcal{R}_{d,k,m}$, obtained with the software Macaulay2. We compute the dimension of the linear span, the dimension of $\mathcal{R}_{d,k,m}$ and its degree. Despite Proposition 26, all the ideals in the examples we present are generated by quadrics. In the last column "gen" we record the number of generators.

d	k	m	span	dim	deg	gen
2	2	2	3	2	2	1
3	2	2	8	5	4	6
4	2	2	15	9	8	20
5	2	2	24	14	16	50
6	2	2	35	20	32	105
2	3	2	5	2	4	6
3	3	2	18	5	24	81
4	3	2	43	9	200	486
5	3	2	84	14	2221	1920
2	4	2	8	2	8	27
3	4	2	38	5	128	528
2	5	2	11	2	12	43
3	5	2	68	5	368	1806
2	6	2	15	2	18	87
2	3	3	7	4	4	6
3	3	3	26	13	24	81
4	3	3	63	29	200	486
2	4	3	12	4	12	33
3	4	3	62	13	672	954
2	4	4	15	7	12	33
2	5	4	25	7	40	150
2	5	5	31	13	40	150
2	6	5	54	13	336	694

References

[1] Améndola C., Friz P., Sturmfels B., *Varieties of signature tensors*, https://arxiv.org/pdf/1804.08325.pdf.

[2] Chen K.-T., *Iterated integrals and exponential homomorphisms*, Proc. London Math. Soc. **4** (1954), 502-512.

[3] Chen K.-T., *Integration of paths - A faithful representation of paths by noncommutative formal power series*, Transactions Amer. Math. Soc. **89** (1958), 395-407.

[4] Cox D., Little J., Schenck H., *Toric varieties*, American Mathematical Society, 2011.

[5] Diehl J., Preiß R., Reizenstein J., *Coordinates of the first kind and "areas of areas"*, in preparation.

[6] Friz P, Hairer M., *A course on rough paths, with an introduction to regularity structures*, Universitext, Springer, Cahm, 2014.

[7] Friz P., Victoir N., *Multidimensional stochastic processes as rough paths* (Cambridge Studies in Advanced Mathematics), Cambridge University Press, Cambridge, 2010.

[8] Gehrig E., Kawski M., *A Hopf-algebraic formula for compositions of noncommuting flows*, Proceedings of the 47th IEEE Conference on Decision and Control (2008), 1569–1574.

[9] Grayson D., Stillman M., *Macaulay2, a software system for research in algebraic geometry*, available at `www.math.uiuc.edu/Macaulay2`.

[10] Lyons T., Sidorova N., *Sound compression: a rough path approach*, Proceedings of the 4th international symposium on Information and communication technologies, 223-228, Trinity College Dublin, 2005.

[11] Pfeffer M., Seigal A., Sturmfels B., *Learning paths from signature tensors*, https://arxiv.org/abs/1809.01588.

[12] Reutenauer C., *Free Lie algebras*, London Mathematical Society Monographs, New Series **7**, Oxford University Press, New York, 1993.

[13] Stanley R., *Enumerative combinatorics* (Cambridge Studies in Advanced Mathematics), Cambridge University Press, Cambridge, 2011.

Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, 04103 Leipzig, Germany

E-mail address: galuppi@mis.mpg.de