Tree algebras over topological vector spaces in rough path theory

Thomas Cass[*]   Martin P. Weidner[†]

Imperial College London

Abstract

We work with non-planar rooted trees which have a label set given by an arbitrary vector space V. By equipping V with a complete locally convex topology, we show how a natural topology is induced on the tree algebra over V. In this context, we introduce the Grossman-Larson and Connes-Kreimer topological Hopf algebras over V, and prove that they form a dual pair in a certain sense. As an application we define the class of branched rough paths over a general Banach space, and propose a new definition of a solution to a rough differential equation (RDE) driven by one of these branched rough paths. We show equivalence of our definition with a Davie-Friz-Victoir-type definition, a version of which is widely used for RDEs with geometric drivers, and we comment on applications to RDEs with manifold-valued solutions.

Keywords: Labelled trees, Hopf algebras, Rough differential equations

Introduction

A cornerstone of classical rough path theory is to understand differential equations of the form

$$\mathrm{d}y_t = V(y_t)\,\mathrm{d}\mathbf{x}_t \text{ for } t \in [0, T],\, \text{started at } y_0 \in F, \tag{1}$$

in which $\mathbf{x}$ is a *geometric p-rough path* on a Banach space E, and V is a continuous linear map from E into the space of vector fields on another Banach space F. A number of different approaches to this definition exist. The first chronologically is due to Lyons [19] for whom a solution is defined as a fixed point, in geometric rough path space, of the rough integral equation corresponding to (1). More recently, Gubinelli introduced an essentially equivalent definition which characterises solutions in terms of *controlled rough paths*, cf.

[14, 15] and [16]. In the finite-dimensional case $E = \mathbf{R}^d$ and $F = \mathbf{R}^n$, Davie [9] (for $2 \leq p < 3$) and then Friz and Victoir [11] (for general $p \geq 2$) proposed a notion of solution motivated by the Euler approximation for ordinary differential equations. This definition requires that y and $\mathbf{x}$ be related[1] by

$$y_t - y_s \simeq \sum_{k=1}^{\lfloor p \rfloor} \sum_{i_1,\ldots,i_k=1}^{d} V_{i_1} \cdots V_{i_k} \mathrm{id}\,(y_s)\, \mathbf{x}_{s,t}^{k;i_1,\ldots,i_k} \text{ for } s,t \in [0,T]\,, \tag{2}$$

where $\mathbf{x}_{s,t}^{k;i_1,\ldots,i_k} := \langle e_{i_1} \otimes \cdots \otimes e_{i_k}, \mathbf{x}_{s,t} \rangle$, and $V_i(y) := V(y)(e_i)$ when $\{e_i\}_{i=1}^{d}$ denotes the standard basis of $\mathbf{R}^d$. Somewhat differently, and inspired by the work of Strichartz [24], Bailleul in [2] interprets a solution as the unique *rough flow* (a diffeomorphism on F) associated to a class of approximate flows. This rough flow can be used to construct a path y which satisfies

$$f(y_t) - f(y_s) \simeq \sum_{k=1}^{\lfloor p \rfloor} \sum_{i_1,\ldots,i_k=1}^{d} V_{i_1}...V_{i_k} f(y_s)\, \mathbf{x}_{s,t}^{k;i_1,\ldots,i_k} \text{ for } s,t \in [0,T]\,, \tag{3}$$

for all smooth enough functions $f : F \to G$ and for any Banach space G. See Proposition 3.5 of [2].

In many situations (3) is a more natural test than (2) to determine if a given path y solves (1). For example, on any smooth manifold (3) still has meaning even if the increment on the left side of (2) does not. This observation has been used in [5] in order to study rough differential equations on submanifolds of $\mathbf{R}^d$. One application of this paper will be to show the equivalence of (2) and (3); this is currently unknown even for the special case sketched above where E and F are finite-dimensional and for geometric $\mathbf{x}$. The basic challenge is algebraic: a general smooth function f can have non-zero derivatives of any order, while the derivatives of the identity function of degree two and higher vanish identically. The summands in (3) therefore typically involve many more terms than those in (2), and to handle these we need to systematically keep track of terms involving derivatives of f of a fixed degree. The algebraic structure needed is exactly the Grossman-Larson Hopf algebra on labelled rooted non-planar trees. These ideas, which orginally go back to Cayley [6], appear in Butcher's seminal study of numerical methods for differential equations [4]. Connes and Kreimer [8] and Grossman and Larson [12] subsequently introduced two different Hopf algebra structures based on rooted trees; these are dual to each another, cf. [21] and [17], allowing one to view the Butcher group either as the character group of the Connes-Kreimer algebra, or as the group-like elements of the Grossman-Larson algebra. Recently the connection between tree algebras, renormalisation theory and rough paths has led to new results such as the work by Bruned, Chevyrev, Friz and Preiß [3].

---

[1]The meaning of $\simeq$ in (2) is made precise later, but it relates two terms whose difference is negligible in a certain sense.

Our analysis admits two important generalisations on the motivating discussion above, namely:

1. It allows for *branched rough path* drivers. These have been introduced by Gubinelli and developed further by Hairer and Kelly. They generalise the notion of a geometric rough path by removing the algebraic constraints which come from the classical integration-by-parts relations. As such, a wider class of driving signals can be accommodated, e.g. stochastic rough paths constructed via iterated Itô integration.

2. It permits E and F to be general Banach spaces. An key advantage of classical rough path theory is its ability to treat models in which the signal and response evolve in infinite dimensional state-spaces. This poses an extra challenge to the algebraic methods, most of which work for trees with a finite – or at most countable – label set. In this article, by constrast, we will work with general locally convex topological vector spaces as label sets.

The eventual outcome of this paper will thus be to close two gaps in the literature: we present the first rigorous treatment of infinite-dimensional branched rough paths and then, in this general setting, we prove that (3) is equivalent to the classical Davie-Friz-Victoir notion of solution to a rough differential equation.

The structure of this work is as follows. In Section 1 we define trees as partially ordered sets and introduce the properties and operations that are relevant in later sections. Section 2 contains the construction of algebras of labelled trees, where the label set is given by an arbitrary vector space V. If V carries a topology, then this topology induces a topology on the tree algebras associated to V. This topology is introduced in Section 3. Once the algebraic and topological algebras of Grossman-Larson and Connes-Kreimer are introduced, we show in Section 4 that they form a dual pair of Hopf algebras. Finally we apply this duality in the context of rough paths in Section 5.

1 The set of non-planar rooted trees

There exist various equivalent definitions for the set of non-planar rooted trees and here we choose to follow [17] and work with equivalence classes of partially ordered sets. Two partially ordered sets are defined to be *equivalent* if there exists an order preserving bijection between them.

Definition 1.1 A *Non-planar rooted tree* $\mathfrak{t}$ is an equivalence class of finite partially ordered sets S such that S has a unique minimal element and for every $v \in S$ the set $\{w \in S : w \prec v\}$ is totally ordered. The set of all (non-planar rooted) trees is denoted by $\mathfrak{Tr}$.

From now on, by a tree $\mathfrak{t} \in \mathfrak{Tr}$ we mean a generic representative of the equivalence class and we will address the problem of well-definedness where we deem it necessary.

The elements of a tree $\mathfrak{t} \in \mathfrak{Tr}$ are called *vertices* and its unique minimal element is called its *root* and it is denoted by $r(\mathfrak{t})$. By $v(\mathfrak{t}) := \mathfrak{t}\backslash r(\mathfrak{t})$ we denote the set of all vertices of $\mathfrak{t}$ without the root. Finally we define $|\mathfrak{t}|$ to be the number of elements of $v(\mathfrak{t})$ and we call this number the *degree* of $\mathfrak{t}$. There exists a unique $\mathfrak{t} \in \mathfrak{Tr}$ with $|\mathfrak{t}| = 0$ and we denote this tree by $\mathbf{1}$.

The *symmetry group* $SG(\mathfrak{t})$ of $\mathfrak{t} \in \mathfrak{Tr}$ is the group of all order-preserving bijections from $v(\mathfrak{t})$ to $v(\mathfrak{t})$.

For $\mathfrak{t} \in \mathfrak{Tr}$ and $v \in \mathfrak{t}$, denote by $c(v)$ the set of *children* of v, i.e. all vertices $w \in \mathfrak{t}$ such that $w \succ v$ and such that there exists no $x \in \mathfrak{t}$ with $w \succ x \succ v$.

Trees can be canonically interpreted as graphs where the vertices of the graph correspond to the vertices of the tree and the edges of the graph correspond to the statement 'is a child of'.

An important operation is that of *grafting* two trees together. Let $\mathfrak{t}, \mathfrak{s} \in \mathfrak{Tr}$ be two trees. A grafting map from $\mathfrak{t}$ to $\mathfrak{s}$ is a map $d : c(r(\mathfrak{t})) \to \mathfrak{s}$, i.e. it assigns a vertex of $\mathfrak{s}$ to every child of the root of $\mathfrak{t}$. Denote by $Gr(\mathfrak{t}, \mathfrak{s})$ the set of all such maps. Given $d \in Gr(\mathfrak{t}, \mathfrak{s})$ we define $\mathfrak{t} \xrightarrow{d} \mathfrak{s} := v(\mathfrak{t}) \sqcup \mathfrak{s}$ with the additional relation $v \succ d(v)$ for every $v \in c(r(\mathfrak{t}))$ (plus all necessary additional relations so that transitivity holds). Here $\sqcup$ denotes the disjoint union.

In the special case where $d(v) = r(\mathfrak{s})$ for all $v \in c(r(\mathfrak{t}))$ we write $\mathfrak{t} \circ \mathfrak{s}$ instead of $\mathfrak{t} \xrightarrow{d} \mathfrak{s}$. Observe that we have $\mathfrak{t} \circ \mathfrak{s} = \mathfrak{s} \circ \mathfrak{t}$, since the operation $\circ$ simply identifies the roots of $\mathfrak{t}$ and $\mathfrak{s}$.

Example 1.2 These are all possible ways of grafting $\mathfrak{t} = $ to $\mathfrak{s} = $:
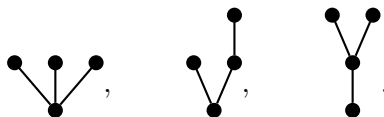
Note however that $|Gr(\mathfrak{t}, \mathfrak{s})| = 4$ since there are two possibilities to obtain the middle one of the above three trees. Furthermore the first tree corresponds to $\mathfrak{t} \circ \mathfrak{s}$.

Let $\mathfrak{t} \in \mathfrak{Tr}$ be a tree. We call a subset $C \subset v(\mathfrak{t})$ of its vertices an *admissible cut* if none of its elements are comparable, i.e. if for all $v, w \in C$ we have neither $v \prec w$ nor $w \prec v$. The set of all admissible cuts of $\mathfrak{t}$ is denoted by $Cut(\mathfrak{t})$. For a given cut $C \in Cut(\mathfrak{t})$ we define the trees

$$R^C(\mathfrak{t}) := \mathfrak{t}\backslash\{v \in v(\mathfrak{t}) : \text{there exists } w \in C \text{ such that } w \preceq v\}$$

and

$$P^C(\mathfrak{t}) := \{r\} \sqcup \{v \in v(\mathfrak{t}) : \text{there exists } w \in C \text{ such that } w \preceq v\}$$

with the additional relations $r \prec w$ for all $w \in C$ (plus all necessary additional relations so that transitivity holds). Intuitively, $R^C(\mathfrak{t})$ contains the tree that remains after cutting

away the elements of C and $P^C(\mathfrak{t})$ is obtained by grafting all the subtrees that have been removed to a new root.

Example 1.3 The result of the cutting operation when C is given by the highlighted vertices.
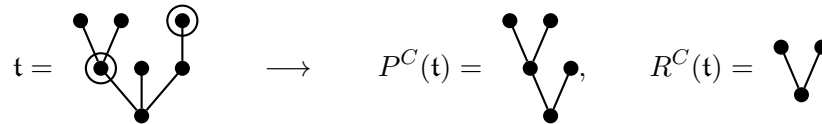
In order to keep track of the vertices of a tree $\mathfrak{t} \in \mathfrak{Tr}$, we would like to consider bijections $\rho : v(\mathfrak{t}) \to \{1, \ldots, |\mathfrak{t}|\}$. However, since we are working with equivalence classes of trees, we have to consider equivalence classes of such bijections. Thus we assume that every representative T of $\mathfrak{t}$ comes with a bijective map $\rho_T : v(T) \to \{1, \ldots, |\mathfrak{t}|\}$ which we call the *enumeration of T*. Furthermore we assume that for any two representatives T and $\tilde{T}$ of $\mathfrak{t}$ the enumerations ρ_T and $\rho_{\tilde{T}}$ are equivalent in the sense that there exists an order preserving bijection $\varphi : T \to \tilde{T}$ with $\rho_T = \rho_{\tilde{T}} \circ \varphi$.

For a tree $\mathfrak{t} \in \mathfrak{Tr}$ we obtain a canonical embedding of $\mathrm{SG}(\mathfrak{t})$ into the symmetric group $\mathfrak{S}_{|\mathfrak{t}|}$ by its enumeration $\rho_{\mathfrak{t}}$, i.e. we define the group

$$\{\rho_{\mathfrak{t}} \pi \rho_{\mathfrak{t}}^{-1} : \pi \in \mathrm{SG}(\mathfrak{t})\} \subseteq \mathfrak{S}_{|\mathfrak{t}|}.$$

We will abuse notation and denote this embedding by $\mathrm{SG}(\mathfrak{t})$ as well.

For $k \in \mathbf{N}$ we denote by $\mathfrak{Tr}_k \subset \mathfrak{Tr}$ the set of all trees whose root has exactly k children and by $\mathfrak{Tr}^k \subset \mathfrak{Tr}$ the set of all trees with degree k. We will also use the notation $\mathfrak{Tr}_n^k := \mathfrak{Tr}^k \cap \mathfrak{Tr}_n$. The set $\mathfrak{Tr}_1$ is especially important as the following trivial yet important lemma shows.

Lemma 1.4 *Let $\mathfrak{t} \in \mathfrak{Tr}_k$. Then there exist unique trees $\mathfrak{t}_1, \ldots, \mathfrak{t}_k \in \mathfrak{Tr}_1$ such that we have the decomposition $\mathfrak{t} = \mathfrak{t}_1 \circ \cdots \circ \mathfrak{t}_k$.*

2 Algebras of labelled trees over a vector space

Let M be a set. The most straightforward way to define a labelling of the vertices of a tree $\mathfrak{t} \in \mathfrak{Tr}$ with labels from M is to consider a map $l : v(\mathfrak{t}) \to M$. More precisely, we have to consider equivalence classes of such maps where two maps l and $\tilde{l}$ are equivalent if there exists $\pi \in \mathrm{SG}(\mathfrak{t})$ such that $l = \tilde{l}\pi$. We can then consider the (free) vector space spanned by all of those equivalence classes of labellings for a fixed tree $\mathfrak{t} \in \mathfrak{Tr}$. Let us denote this vector space by $\mathcal{L}_{\mathfrak{t}}(M)$.

This definition is perfectly fine and especially if M is finite there is not much difference between labelled and unlabelled trees since in that case the space $\mathcal{L}_{\mathfrak{t}}(M)$ is finite dimensional for every $\mathfrak{t} \in \mathfrak{Tr}$. However, if M is infinite, then the vector space $\mathcal{L}_{\mathfrak{t}}(M)$ is infinite dimensional for every $\mathfrak{t} \in \mathfrak{Tr} \backslash \{\mathbf{1}\}$. It is therefore desirable to equip $\mathcal{L}_{\mathfrak{t}}(M)$ with a topology since topological vector spaces are usually nicer to work with than infinite dimensional vector spaces without a topology, for example when it comes to dual spaces.

Hence we would like to establish a different point of view which takes the (free) vector space V spanned by M as the label set rather than the set M itself. In fact, it is usually V that is given in the first place rather than M. Thus our aim is to describe the spaces $\mathcal{L}_t(M)$ in terms of V rather than M so that we can then use a given topology on V to construct a topology on $\mathcal{L}_t(M)$.

In terms of the application to rough paths that we will present in Section 5 the vector space V will take the role of the noise space. The topological tree algebra generated by it – in the way given below – will then be the space in which the rough path lift of the noise takes its values.

2.1 Spaces of labelled trees as tensor products

Let now V be a given vector space. We will perform a purely algebraic construction in this section and then consider a topological vector space in the following one. A permutation $\sigma \in \mathfrak{S}_k$ with $k \in \mathbf{N}$ can be canonically interpreted as a linear map on $V^{\otimes k}$ which is given by $\sigma(v_1 \otimes \cdots \otimes v_k) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}$ for elementary tensors. Given a subgroup $G \subset \mathfrak{S}_k$ we define the projection

$$\text{sym}_G : V^{\otimes k} \to V^{\otimes k} : a \mapsto \frac{1}{|G|} \sum_{\sigma \in G} \sigma(a).$$

Definition 2.1 Let V be a vector space, let $t \in \mathfrak{Tr}$ be a tree and let $\text{SG}(t)$ be its symmetry group. Then the t-*tensor power of V* is defined as the quotient space

$$V^{\otimes t} := V^{\otimes |t|} \Big/ \ker\left(\text{sym}_{\text{SG}(t)}\right)$$

The space $V^{\otimes t}$ is isomorphic to the image of $\text{sym}_{\text{SG}(t)}$, which provides a more direct description of $V^{\otimes t}$. We say that an element $\mathbf{t} \in V^{\otimes t}$ is an *elementary labelled tree* if its equivalence class contains an elementary tensor. The set of all elementary labelled trees in $V^{\otimes t}$ spans $V^{\otimes t}$.

To illustrate an elementary labelled tree we choose one of its elementary representatives as in the following example. E.g., for $t = \mathbf{V}$ we can visualise $\mathbf{t} = [2v_1 \otimes v_2 + v_2 \otimes v_1] \in V^{\otimes t}$ as

$$\mathbf{t} = 3 \overset{v_1 \ v_2}{\mathbf{V}}.$$

Let us now define the following spaces.

$$\mathfrak{Tr}^k(V) := \bigoplus_{t \in \mathfrak{Tr}^k} V^{\otimes t} \text{ for } k \in \mathbf{N}$$

$$\mathfrak{Tr}_k(V) := \bigoplus_{t \in \mathfrak{Tr}_k} V^{\otimes t} \text{ for } k \in \mathbf{N}$$

$$\mathfrak{Tr}_n^k(V) := \bigoplus_{\mathfrak{t} \in \mathfrak{Tr}_n^k} V^{\otimes \mathfrak{t}} = \mathfrak{Tr}^k(V) \cap \mathfrak{Tr}_n(V) \text{ for } k, n \in \mathbf{N}$$

$$\mathfrak{Tr}(V) := \bigoplus_{\mathfrak{t} \in \mathfrak{Tr}} V^{\otimes \mathfrak{t}} = \bigoplus_{k=0}^{\infty} \mathfrak{Tr}^k(V) = \bigoplus_{k=0}^{\infty} \mathfrak{Tr}_k(V) = \bigoplus_{k=0}^{\infty} \bigoplus_{n=0}^{\infty} \mathfrak{Tr}_n^k(V)$$

The definition of $\mathfrak{Tr}(V)$ provides canonical projections onto $\mathfrak{Tr}^k(V)$, $\mathfrak{Tr}_k(V)$ and $\mathfrak{Tr}_n^k(V)$ for $k, n \in \mathbf{N}$ and we will denote those projections by π^k, π_k and π_n^k respectively.

We will also need truncations of the space $\mathfrak{Tr}(V)$. Thus for $k \in \mathbf{N}$ define the subspace

$$I^{(k)} := \bigoplus_{\mathfrak{t} \in \mathfrak{Tr}: |\mathfrak{t}| > k} V^{\hat{\otimes} \mathfrak{t}}$$

and the quotient space

$$\mathfrak{Tr}^{(k)}(V) := \mathfrak{Tr}(V) \Big/ I^{(k)} .$$

The canonical projection $\mathfrak{Tr}(V) \to \mathfrak{Tr}^{(k)}(V)$ will be denoted by $\pi^{(k)}$. We will furthermore need the projection $\pi_n^{(k)} := \pi^{(k)} \circ \pi_n = \pi_n \circ \pi^{(k)}$.

Before we can define the additional algebraic structure on $\mathfrak{Tr}(V)$ to turn it into a Hopf algebra we have to extend two auxiliary constructions from the previous section, namely grafting and cutting, to tensors. We start with grafting.

Definition 2.2 Let $\mathfrak{t}, \mathfrak{s} \in \mathfrak{Tr}$ and let $t \in V^{\otimes |\mathfrak{t}|}$ and $s \in V^{\otimes |\mathfrak{s}|}$ be two elementary tensors. Let furthermore $d \in \mathrm{Gr}(\mathfrak{t}, \mathfrak{s})$ be a grafting map. Then we define

$$t \overset{d}{\curlywedge} s := \left(\rho_{\mathfrak{t},\mathfrak{s}} \circ \rho_{\mathfrak{t} \overset{d}{\curlywedge} \mathfrak{s}}^{-1} \right) (t \otimes s) \in V^{\otimes(|\mathfrak{t}| + |\mathfrak{s}|)}, \tag{4}$$

where

$$\rho_{\mathfrak{t},\mathfrak{s}} : \mathrm{v}(\mathfrak{t}) \sqcup \mathrm{v}(\mathfrak{s}) \to \{1, \ldots, |\mathfrak{t}| + |\mathfrak{s}|\} : v \mapsto \begin{cases} \rho_{\mathfrak{t}}(v) & v \in \mathrm{v}(\mathfrak{t}), \\ \rho_{\mathfrak{s}}(v) + |\mathfrak{t}| & v \in \mathrm{v}(\mathfrak{s}). \end{cases} \tag{5}$$

Remark 2.3 This definition has to be handled with care since the value of $t \overset{d}{\curlywedge} s$ may depend on the choice of representatives of $\mathfrak{t}$ and $\mathfrak{s}$. However, this is unproblematic because it is only an auxiliary construction for Definition 2.7, where the resulting object will be independent of the choice of representatives, which will be shown in Lemma 2.9.

By analogy with to the previous section, if $d \in \mathrm{Gr}(\mathfrak{t}, \mathfrak{s})$ maps all elements of $\mathrm{c}(\mathrm{r}(\mathfrak{t}))$ to $\mathrm{r}(\mathfrak{s})$ we write $t \circ s$ instead of $t \overset{d}{\curlywedge} s$.

Lemma 2.4 Let $\mathfrak{t}, \mathfrak{s} \in \mathfrak{Tr}$ and let $\mathbf{t} \in V^{\otimes \mathfrak{t}}$ and $\mathbf{s} \in V^{\otimes \mathfrak{s}}$ be two labelled trees. Let furthermore $t \in V^{\otimes |\mathfrak{t}|}$ and $s \in V^{\otimes |\mathfrak{s}|}$ be representatives of $\mathbf{t}$ and $\mathbf{s}$ respectively. Then the equivalence class $[t \circ s] \in V^{\otimes (\mathfrak{t} \circ \mathfrak{s})}$ depends neither on the choice of t and s nor on the choice of representatives of $\mathbf{t}$ and $\mathbf{s}$.

7

Proof. Choosing a different representative of $\mathfrak{t}$ and/or a different choice of t amount to replacing t by $\tilde{t}$ with the property that

$$\text{sym}_{\text{SG}(\mathfrak{t})}(t - \tilde{t}) = 0.$$

From (4) we get $t \circ s - \tilde{t} \circ s = (t - \tilde{t}) \circ s$. Thus if $\text{SG}(\mathfrak{t})$ embeds as a subgroup into $\text{SG}(\mathfrak{t} \circ \mathfrak{s})$ we are done because then we have

$$[t \circ s] - [\tilde{t} \circ s] = [(t - \tilde{t}) \circ s] = 0.$$

But this is clearly the case because the operation $\circ$ identifies the roots of $\mathfrak{t}$ and $\mathfrak{s}$ and hence both $\mathfrak{t}$ is a subtree of $\mathfrak{t} \circ \mathfrak{s}$.

The same argument shows that choosing a different representative of $\mathfrak{s}$ and/or a different choice of s does not change $[t \circ s]$. ∎

The lemma allows us to define $\mathbf{t} \circ \mathbf{s} := [t \circ s]$ for representatives t and s of $\mathbf{t}$ and $\mathbf{s}$ respectively. If $\mathbf{t} \in V^{\otimes \mathfrak{t}}$ is an elementary labelled tree as defined above, then the unique decomposition $\mathfrak{t} = \mathfrak{t}_1 \circ \cdots \circ \mathfrak{t}_k$ from Lemma 1.4 induces a decomposition

$$\mathbf{t} = \mathbf{t}_1 \circ \cdots \circ \mathbf{t}_k \text{ with } \mathbf{t}_1 \in V^{\otimes \mathfrak{t}_1}, \ldots, \mathbf{t}_k \in V^{\otimes \mathfrak{t}_k}. \tag{6}$$

Let us now turn to cutting. Let $\mathfrak{t} \in \mathfrak{Tr}$ and $C \in \text{Cut}(\mathfrak{t})$. Then the vertices of $R^C(\mathfrak{t})$ and $P^C(\mathfrak{t})$ are defined as subsets of the vertices of $\mathfrak{t}$ and hence there are canonical injections

$$\phi_{C,R} : \text{v}\left(R^C(\mathfrak{t})\right) \to \text{v}(\mathfrak{t}) \qquad \text{and} \qquad \phi_{C,P} : \text{v}\left(P^C(\mathfrak{t})\right) \to \text{v}(\mathfrak{t}).$$

Definition 2.5 Let $\mathfrak{t} \in \mathfrak{Tr}$ be a tree and let $C \in \text{Cut}(\mathfrak{t})$ be an admissible cut. Then we define

$$\mathbf{R}^C : V^{\otimes |\mathfrak{t}|} \to V^{\otimes |R^C(\mathfrak{t})|} : v_1 \otimes \cdots \otimes v_{|\mathfrak{t}|} \mapsto v_{\varphi_{C,R}(1)} \otimes \cdots \otimes v_{\varphi_{C,R}(|R^C(\mathfrak{t})|)}$$

and

$$\mathbf{P}^C : V^{\otimes |\mathfrak{t}|} \to V^{\otimes |P^C(\mathfrak{t})|} : v_1 \otimes \cdots \otimes v_{|\mathfrak{t}|} \mapsto v_{\varphi_{C,P}(1)} \otimes \cdots \otimes v_{\varphi_{C,P}(|P^C(\mathfrak{t})|)},$$

where $\varphi_{C,R} := \rho_\mathfrak{t} \circ \phi_{C,R} \circ \rho_{R^C(\mathfrak{t})}^{-1}$ and $\varphi_{C,P} := \rho_\mathfrak{t} \circ \phi_{C,P} \circ \rho_{P^C(\mathfrak{t})}^{-1}$.

Remark 2.6 As in Remark 2.3 we should emphasize that $\mathbf{R}^C$ and $\mathbf{P}^C$ depend on the choice of the representative of $\mathbf{t}$. Nevertheless, the following definition makes sense as Lemma 2.9 shows.

Definition 2.7 Let V be a vector space, let $\mathfrak{t}, \mathfrak{s} \in \mathfrak{Tr}$ and let $\mathbf{t} \in V^{\otimes \mathfrak{t}}$ and $\mathbf{s} \in V^{\otimes \mathfrak{s}}$ be elementary labelled trees.

(i) For elementary representatives t and s of $\mathbf{t}$ and $\mathbf{s}$ we define

$$\mathbf{t} \star \mathbf{s} := \sum_{d \in \text{Gr}(\mathfrak{t}, \mathfrak{s})} \left[t \overset{d}{\curlywedge} s\right] \in \mathfrak{Tr}(V).$$

(ii) If $\mathbf{t}$ can be decomposed into $\mathbf{t} = \mathbf{t}_1 \circ \cdots \circ \mathbf{t}_k$ as in (6) we define

$$\Delta_{\mathrm{GL}}(\mathbf{t}) := \sum_{I \subset \{1,\ldots,k\}} \bigcirc_{i \in I} \mathbf{t}_i \otimes \bigcirc_{i \in I^c} \mathbf{t}_i \in \mathfrak{Tr}(V) \otimes \mathfrak{Tr}(V),$$

where the sum ranges over all subsets of $\{1, \ldots, k\}$ and we set $\bigcirc_{i \in \emptyset} \mathbf{t}_i = \mathbf{1}$.

(iii) For an elementary representative t of $\mathbf{t}$ we define

$$\Delta_{\mathrm{CK}}(\mathbf{t}) := \sum_{C \in \mathrm{Cut}(\mathbf{t})} \left[\mathbf{P}^C(t) \right] \otimes \left[\mathbf{R}^C(t) \right] \in \mathfrak{Tr}(V) \otimes \mathfrak{Tr}(V).$$

(iv) For $\lambda \in \mathbf{R}$ we define

$$\eta(\lambda) := \lambda \mathbf{1} \in \mathfrak{Tr}(V)$$

(v) Finally we define

$$\varepsilon(\mathbf{t}) := \begin{cases} \mathbf{t}, & \mathbf{t} = \mathbf{1}, \\ 0, & \text{otherwise} \end{cases} \in \mathbf{R}.$$

All of the above maps are then extended linearly to all $\mathbf{t}, \mathbf{s} \in \mathfrak{Tr}(V)$ and $\lambda \in \mathbf{R}$.

Example 2.8

Lemma 2.9 *With the notation as in the previous definition, neither $\mathbf{t} \star \mathbf{s}$ nor $\Delta_{\mathrm{CK}}(\mathbf{t})$ depends on the choice of t and s or on the choice of representatives of $\mathbf{t}$ and $\mathfrak{s}$.*

Proof. Let $\tilde{t}$ and $\tilde{s}$ be different elementary representatives of $\mathbf{t}$ and $\mathbf{s}$ respectively. The simple yet crucial observation is that there exist $\sigma_1 \in \mathrm{SG}(\mathbf{t})$ and $\sigma_2 \in \mathrm{SG}(\mathfrak{s})$ such that we have

$$\left[\tilde{t} \overset{d}{\to} \tilde{s}\right] = \left[\sigma_1(t) \overset{d}{\to} \sigma_2(s)\right] = \left[t \overset{\sigma_2 \circ d \circ \sigma_1}{\to} s\right],$$

where the second equality is a straightforward consequence of Definition 2.2 and the fact that the enumerations of different representatives of the same tree are compatible. Here σ_1 and σ_2 may also depend on the choice of representatives of $\mathbf{t}$ and $\mathfrak{s}$. On the other hand, for fixed $\sigma_1 \in \mathrm{SG}(\mathbf{t})$ and $\sigma_2 \in \mathrm{SG}(\mathfrak{s})$ we have

$$\mathrm{Gr}(\mathbf{t}, \mathfrak{s}) = \{\sigma_2 \circ d \circ \sigma_1 : d \in \mathrm{Gr}(\mathbf{t}, \mathfrak{s})\}.$$

Hence $\mathbf{t} \star \mathbf{s}$ is well-defined as long as the same representatives are chosen for each term in the defining sum.

A similar argument shows that $\Delta_{\mathrm{CK}}(\mathbf{t})$ is well-defined since we sum over all admissible cuts. $\blacksquare$

Proposition 2.10 *Each of the tuples*

$$\mathcal{H}_{\mathrm{GL}}(V) := (\mathfrak{Tr}(V), \star, \eta, \Delta_{GL}, \varepsilon) \ and$$
$$\mathcal{H}_{\mathrm{CK}}(V) := (\mathfrak{Tr}(V), \circ, \eta, \Delta_{CK}, \varepsilon)$$

defines a graded bialgebra and therefore each of them can be equipped with an antipode which turns them into Hopf algebras. We call $\mathcal{H}_{GL}(V)$ the Grossman-Larson Hopf algebra *over V and $\mathcal{H}_{CK}(V)$ the* Connes-Kreimer Hopf algebra *over V.*

Proof. Let us show that the product $\star$ is associative. We will essentially follow [12]. Let thus $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3 \in \mathfrak{Tr}$ and $\mathbf{t}_1 \in V^{\otimes \mathbf{t}_1}$, $\mathbf{t}_2 \in V^{\otimes \mathbf{t}_2}$ and $\mathbf{t}_3 \in V^{\otimes \mathbf{t}_3}$ and fix representatives t_1, t_2 and t_2 of $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_3$ respectively. Then we need to show

$$\sum_{d_1 \in \mathrm{Gr}(\mathbf{t}_1, \mathbf{t}_2)} \sum_{d_2 \in \mathrm{Gr}(\mathbf{t}_1 \overset{d_1}{\to} \mathbf{t}_2, \mathbf{t}_3)} \left[\left(t_1 \overset{d_1}{\to} t_2\right) \overset{d_2}{\to} t_3\right] = \sum_{e_2 \in \mathrm{Gr}(\mathbf{t}_2, \mathbf{t}_3)} \sum_{e_1 \in \mathrm{Gr}(\mathbf{t}_1, \mathbf{t}_2 \overset{e_2}{\to} \mathbf{t}_3)} \left[t_1 \overset{e_1}{\to} \left(t_2 \overset{e_2}{\to} t_3\right)\right]. \tag{7}$$

Thus assume $d_1 \in \mathrm{Gr}(\mathbf{t}_1, \mathbf{t}_2)$ and $d_2 \in \mathrm{Gr}(\mathbf{t}_1 \overset{d_1}{\to} \mathbf{t}_2, \mathbf{t}_3)$. Define $e_2 \in \mathrm{Gr}(\mathbf{t}_2, \mathbf{t}_3)$ to be the restriction of d_2 to the children of the root of $\mathbf{t}_2$ and define $e_1 \in \mathrm{Gr}(\mathbf{t}_1, \mathbf{t}_2 \overset{e_2}{\to} \mathbf{t}_3)$ via

$$e_1(v) := \begin{cases} d_1(v) & d_1(v) \neq \mathrm{r}(\mathbf{t}_2) \\ d_2(v) & d_1(v) = \mathrm{r}(\mathbf{t}_2) \end{cases}.$$

It follows from [12, Lem. 2.6/2.7] that this gives a one-to-one correspondence between the terms on both sides of (7) such that

$$\left(\mathbf{t}_1 \overset{d_1}{\to} \mathbf{t}_2\right) \overset{d_2}{\to} \mathbf{t}_3 = \mathbf{t}_1 \overset{e_1}{\to} \left(\mathbf{t}_2 \overset{e_2}{\to} \mathbf{t}_3\right)$$

holds. In light of (5) we define

$$\rho_{\mathfrak{t}_1,\mathfrak{t}_2,\mathfrak{t}_3} : \mathrm{v}(\mathfrak{t}_1) \sqcup \mathrm{v}(\mathfrak{t}_2) \sqcup \mathrm{v}(\mathfrak{t}_3) \to \{1, \ldots, |\mathfrak{t}_1| + |\mathfrak{t}_2| + |\mathfrak{t}_3|\}$$

$$v \mapsto \begin{cases} \rho_{\mathfrak{t}_1}(v) & v \in \mathrm{v}(\mathfrak{t}_1) \\ \rho_{\mathfrak{t}_2}(v) + |\mathfrak{t}_1| & v \in \mathrm{v}(\mathfrak{t}_2) \\ \rho_{\mathfrak{t}_3}(v) + |\mathfrak{t}_1| + |\mathfrak{t}_2| & v \in \mathrm{v}(\mathfrak{t}_3) \end{cases}$$

and it is then easy to verify that

$$\left(t_1 \overset{d_1}{\leftthreetimes} t_2\right) \overset{d_2}{\leftthreetimes} t_3 = t_1 \overset{e_1}{\leftthreetimes} \left(t_2 \overset{e_2}{\leftthreetimes} t_3\right) = \left(\rho_{\mathfrak{t}_1,\mathfrak{t}_2,\mathfrak{t}_2} \circ \rho^{-1}_{\left(t_1 \overset{d_1}{\leftthreetimes} t_2\right) \overset{d_2}{\leftthreetimes} t_3}\right)(t_1 \otimes t_2 \otimes t_3)$$

where d_1, d_2 and e_1, e_2 are related as above. Hence associativity of $\star$ follows.

In a similar way one can verify that the remaining (co)associativity and compatibility conditions hold. ∎

Observe that $I^{(k)}$ as defined above is an ideal for both $\star$ and $\circ$. Hence $\mathfrak{Tr}^{(k)}$ can be interpreted as a quotient algebra of both $(\mathfrak{Tr}(V), \star)$ and $(\mathfrak{Tr}(V), \circ)$. On the other hand $I^{(k)}$ is a coideal for neither Δ_{GK} nor Δ_{CL} and hence the quotient is not a quotient Hopf algebra. Furthermore $\mathfrak{Tr}^{(k)}(V)$ is isomorphic to the closed subspace

$$\bigoplus_{\mathfrak{t} \in \mathfrak{Tr} : |\mathfrak{t}| \leq k} V^{\otimes \mathfrak{t}}$$

of $\mathfrak{Tr}(V)$, which justifies the name *truncated tree algebra*. We will use this isomorphism implicitly without mentioning it. Note however that this vector space isomorphism is not an algebra morphism.

3 Algebras of labelled trees over a topological vector space

Let V be a complete locally convex vector space. Our aim is to understand how the topology of V can induce a topology on $\mathfrak{Tr}(V)$. This question can be split into two problems. First we have to introduce a topology on each of the spaces $V^{\otimes \mathfrak{t}}$ for $\mathfrak{t} \in \mathfrak{Tr}$ and second we have to choose an appropriate topology on their direct sum.

Let $\mathfrak{t} \in \mathfrak{Tr}$ be a tree. The space $V^{\otimes \mathfrak{t}}$ is (isomorphic to) a quotient of a tensor power of V and therefore we can use the well-developed theory of topological tensor products. We recommend [13] for the necessary background. Most importantly, there is no canonical topology on $V^{\otimes k}$ for $k \in \mathbf{N}$. However, if $(\xi_i)_{i \in I}$ is a family of seminorms that generates the topology of V, then for every $k \in \mathbf{N}$ and indices $i_1, \ldots, i_k \in I$ there exists a seminorm $\xi_{i_1} \otimes \cdots \otimes \xi_{i_k}$ on $V^{\otimes k}$ such that the system of seminorms thus obtained satisfies

$$\xi_{i_1} \otimes \cdots \otimes \xi_{i_{m+n}}(a \otimes b) = \xi_{i_1} \otimes \cdots \otimes \xi_{i_m}(a)\xi_{i_{m+1}} \otimes \cdots \otimes \xi_{i_{m+n}}(b)$$

for all $m, n \in \mathbf{N}$, $a \in V^{\otimes m}$, $b \in V^{\otimes n}$ and $i_1, \ldots, i_{m+n} \in I$. A system of seminorms with this property is called a system of *cross (semi)norms*. This system is in general not unique and the most important choices are the *projective seminorms*, given by

$$\xi_{i_1} \otimes \cdots \otimes \xi_{i_k}(a) := \inf \left\{ \sum_{j=1}^{N} \xi_{i_1}(v_1^j) \cdots \xi_{i_k}(v_k^j) : a = \sum_{j=1}^{N} v_1^j \otimes \cdots \otimes v_k^j \right\}$$

and the *injective seminorms*, given by

$$\xi_{i_1} \otimes \cdots \otimes \xi_{i_k}(a) := \sup \left\{ |(v_1' \otimes \cdots \otimes v_k')(a)| : v_1', \ldots, v_k' \in V', \xi_{i_1}(v_1') = \cdots = \xi_{i_k}(v_k') = 1 \right\}$$

for all $a \in V^{\otimes k}$. Both the projective and the injective seminorms are symmetric in the sense of the following definition.

Definition 3.1 Let V be a locally convex vector space and let $(\xi_i)_{i \in I}$ be a family of seminorms that generates the topology of V. A system of cross seminorms on the tensor powers of V is called *symmetric* if for all $k \in \mathbf{N}$, all $i_1, \ldots, i_k \in I$, all $a \in V^{\otimes k}$ and all $\sigma \in \mathfrak{S}_k$ we have

$$\xi_1 \otimes \cdots \otimes \xi_k(a) = \xi_{\sigma(1)} \otimes \cdots \otimes \xi_{\sigma(k)}(\sigma(a)).$$

From now on we will always assume that the tensor powers of V are equipped with a symmetric system of cross seminorms and for all $k \in \mathbf{N}$ we will denote by $V^{\hat{\otimes}k}$ the completion of $V^{\otimes k}$ with respect to the respective family of seminorms. These assumptions imply that for all trees $\mathfrak{t} \in \mathfrak{Tr}$ the projection $\mathrm{sym}_{\mathrm{SG}(\mathfrak{t})}$ defined above is continuous. Therefore it can be uniquely extended to a continuous map on $V^{\hat{\otimes}|\mathfrak{t}|}$ and we define

$$V^{\hat{\otimes}\mathfrak{t}} := V^{\hat{\otimes}|\mathfrak{t}|} \Big/ \ker\left(\mathrm{sym}_{\mathrm{SG}(\mathfrak{t})}\right).$$

Equivalently, $V^{\hat{\otimes}\mathfrak{t}}$ is the completion of $V^{\otimes \mathfrak{t}}$ under the quotient topology.

$$\widehat{\mathfrak{Tr}}(V) := \mathrm{cl}\left(\bigoplus_{\mathfrak{t} \in \mathfrak{Tr}} V^{\hat{\otimes}\mathfrak{t}}\right),$$

where the closure is taken with respect to the product topology. Furthermore we observe that for $\mathfrak{t}, \mathfrak{s} \in \mathfrak{Tr}$ we have an isomorphism

$$V^{\otimes \mathfrak{t}} \otimes V^{\otimes \mathfrak{s}} \simeq \left(V^{\otimes|\mathfrak{t}|} \otimes V^{\otimes|\mathfrak{s}|}\right) \Big/ \left(V^{\otimes|\mathfrak{t}|} \otimes \ker\left(\mathrm{sym}_{\mathrm{SG}(\mathfrak{s})}\right) + \ker\left(\mathrm{sym}_{\mathrm{SG}(\mathfrak{t})}\right) \otimes V^{\otimes|\mathfrak{s}|}\right)$$

and therefore the system of cross-seminorms that we have chosen for the $V^{\otimes k}$ gives a locally convex topology on $V^{\otimes \mathfrak{t}} \otimes V^{\otimes \mathfrak{s}}$. Thus, using distributivity of the tensor product and direct sums, we can define the completion of $\widehat{\mathfrak{Tr}}(V) \otimes \widehat{\mathfrak{Tr}}(V)$ via

$$\widehat{\mathfrak{Tr}}(V) \hat{\otimes} \widehat{\mathfrak{Tr}}(V) = \mathrm{cl}\left(\bigoplus_{\mathfrak{t}, \mathfrak{s} \in \mathfrak{Tr}} \left(V^{\hat{\otimes}\mathfrak{t}} \hat{\otimes} V^{\hat{\otimes}\mathfrak{s}}\right)\right).$$

We can now state a topological version of Proposition 2.10.

12

Proposition 3.2 *Each of the tuples*

$$\widehat{\mathcal{H}}_{\mathrm{GL}}(V) := \left(\widehat{\mathfrak{Tr}(V)}, \star, \eta, \Delta_{GL}, \varepsilon \right) \ \ and$$

$$\widehat{\mathcal{H}}_{\mathrm{CK}}(V) := \left(\widehat{\mathfrak{Tr}(V)}, \circ, \eta, \Delta_{CK}, \varepsilon \right)$$

defines a topological graded bialgebra and therefore each of them can be equipped with an antipode which turns them into topological Hopf algebras. We call $\mathcal{H}_{GL}(V)$ the topological Grossman-Larson Hopf algebra over V *and $\mathcal{H}_{CK}(V)$ the* topological Connes-Kreimer Hopf algebra over V.

Proof. It is immediately clear that all operations of $\widehat{\mathcal{H}}_{\mathrm{GL}}(V)$ and $\widehat{\mathcal{H}}_{\mathrm{CK}}(V)$ are continuous since they only consist of tensor multiplications and permutations. Hence the claim follows directly from Proposition 2.10. ∎

Remark 3.3 Technically, a topological Hopf algebra $\mathcal{H}$ is not a Hopf algebra since its coproduct may take values in $\mathcal{H} \hat{\otimes} \mathcal{H}$ rather than $\mathcal{H} \otimes \mathcal{H}$. Nevertheless, when we write *Hopf algebra* in the sequel, we always mean *Hopf algebra or topological Hopf algebra*.

Remark 3.4 There are many sensible choices for a topology on $\bigoplus_{\mathfrak{t} \in \mathfrak{Tr}} V^{\hat{\otimes} \mathfrak{t}}$. For instance, one could consider a topology akin to the one that is introduced in [7] for the tensor algebra. If V is Banach, considering the l^p direct sum for $p \in [1, \infty]$ might be a useful choice. We have chosen the product topology because it is the coarsest sensible choice (in the sense that it renders the canonical projections continuous) and hence it leads to the largest possible closure (which turns out to be the direct product $\prod_{\mathfrak{t} \in \mathfrak{Tr}} V^{\hat{\otimes} \mathfrak{t}}$). Therefore the separation results in the next section hold automatically for any other reasonable choice of topology.

4 The Grossman-Larson and Connes-Kreimer algebras form a dual pair

In the unlabelled case (which corresponds to $V = \mathbf{R}$ in our notation) it has been shown in [17] that $\mathcal{H}_{\mathrm{CK}}(\mathbf{R})$ and $\mathcal{H}_{\mathrm{GL}}(\mathbf{R})$ are graded duals of each other. The graded dual V^{gr} of a graded vector space V is defined as the direct sum of the duals of its homogeneous subspaces. This is no longer true if $\mathbf{R}$ is replaced by an infinite dimensional vector space V, because in this case the spaces $(\mathfrak{Tr}(V) \otimes \mathfrak{Tr}(V))^{\mathrm{gr}}$ and $\mathfrak{Tr}(V)^{\mathrm{gr}} \otimes \mathfrak{Tr}(V)^{\mathrm{gr}}$ are no longer canonically isomorphic. Nevertheless, the Grosmman-Larson and Connes-Kreimer algebras still form a dual pair in a sense which will be made precise in this section.

Definition 4.1 Let V and W be vector spaces and let $\langle \cdot, \cdot \rangle : V \times W \to \mathbf{R}$ be a bilinear map. The triple $(V, W, \langle \cdot, \cdot \rangle)$ is called a *dual pair of vector spaces* if

(i) The set of linear maps $\{ \langle \cdot, w \rangle : w \in W \}$ separates V and

(ii) the set of linear maps $\{\langle v, \cdot \rangle : v \in V\}$ separates W.

If V and W are locally convex vector spaces, we say that $(V, W, \langle \cdot, \cdot \rangle)$ is a *dual pair of locally convex vector spaces* if in addition $\langle \cdot, w \rangle$ and $\langle v, \cdot \rangle$ are continuous for all $v \in V$ and $w \in W$.

Remark 4.2 Any bilinear map $\langle \cdot, \cdot \rangle : V \times W \to R$ induces a bilinear map on $V^{\otimes k} \times W^{\otimes k}$ for $k \in \mathbf{N}$ which is given by

$$(v_1 \otimes \cdots \otimes v_k, w_1 \otimes \cdots \otimes w_k) \mapsto \langle v_1, w_1 \rangle \cdots \langle v_k, w_k \rangle.$$

We will use this fact below and denote the resulting map also by $\langle \cdot, \cdot \rangle$ without further comment.

Lemma 4.3 *Let* $\mathfrak{t} \in \mathfrak{Tr}$ *and let* $(V, W, \langle \cdot, \cdot \rangle)$ *be a dual pair of vector spaces. Define*

$$\langle \cdot, \cdot \rangle_{\mathfrak{t}} : V^{\otimes \mathfrak{t}} \times W^{\otimes \mathfrak{t}} \to \mathbf{R} : (\mathbf{t}_1, \mathbf{t}_2) \mapsto \sum_{\sigma \in \mathrm{SG}(\mathfrak{t})} \langle \sigma(t_1), t_2 \rangle = \sum_{\sigma \in \mathrm{SG}(\mathfrak{t})} \langle t_1, \sigma(t_2) \rangle \qquad (8)$$

where $t_1 \in V^{\otimes |\mathfrak{t}|}$ *and* $t_2 \in W^{\otimes |\mathfrak{t}|}$ *are representatives of* $\mathbf{t}_1$ *and* $\mathbf{t}_2$ *respectively. (Recall that we denote by* $\mathrm{SG}(\mathfrak{t})$ *both the symmetry group of* $\mathfrak{t}$ *and its embedding into* $\mathfrak{S}_{|\mathfrak{t}|}$ *which is induced by the enumeration* $\rho_{\mathfrak{t}}$.*) Then the following hold.*

(i) $(V^{\otimes \mathfrak{t}}, W^{\otimes \mathfrak{t}}, \langle \cdot, \cdot \rangle_{\mathfrak{t}})$ *is a dual pair of vector spaces.*

(ii) *If* $(V, W, \langle \cdot, \cdot \rangle)$ *is a dual pair of locally convex vector spaces, then* $\langle \cdot, \mathbf{t}_2 \rangle_{\mathfrak{t}}$ *can be extended continuously to* $V^{\hat{\otimes} \mathfrak{t}}$ *for all* $\mathbf{t}_2 \in W^{\otimes \mathfrak{t}}$ *and* $\left(V^{\hat{\otimes} \mathfrak{t}}, W^{\otimes \mathfrak{t}}, \langle \cdot, \cdot \rangle_{\mathfrak{t}} \right)$ *is a dual pair of vector spaces.*

Proof. Let us first show that (8) does not depend on the choice of t_1 or t_2. Thus, let $\tilde{t}_1$ be different representatives of $\mathbf{t}_1$. Then we have

$$\sum_{\sigma \in \mathrm{SG}(\mathfrak{t})} \langle \sigma(t_1), t_2 \rangle - \sum_{\sigma \in \mathrm{SG}(\mathfrak{t})} \langle \sigma(\tilde{t}_1), t_2 \rangle = \left\langle \sum_{\sigma \in \mathrm{SG}(\mathfrak{t})} \sigma(t_1 - \tilde{t}_1), t_2 \right\rangle = 0$$

and similarly for t_2.

Let us now prove (i). Fix $\mathbf{t}_2 \in W^{\otimes \mathfrak{t}}$ and assume that we have $\langle \mathbf{t}_1, \mathbf{t}_2 \rangle_{\mathfrak{t}} = 0$ for all $\mathbf{t}_1 \in V^{\otimes \mathfrak{t}}$. Thus if we fix a representative t_2 of $\mathbf{t}_2$ we have

$$0 = \langle \mathbf{t}_1, \mathbf{t}_2 \rangle_{\mathfrak{t}} = \sum_{\sigma \in \mathrm{SG}(\mathfrak{t})} \langle t_1, \sigma(t_2) \rangle = \left\langle t_1, \sum_{\sigma \in \mathrm{SG}(\mathfrak{t})} \sigma(t_2) \right\rangle$$

for all $t_1 \in V^{\otimes |\mathfrak{t}|}$ which implies $\sum_{\sigma \in \mathrm{SG}(\mathfrak{t})} \sigma(t_2) = 0$, i.e. $\mathbf{t}_2 = 0$. This last implication holds because it is shown in [23, Lem. VII/1] that $\left(V^{\otimes |\mathfrak{t}|}, W^{\otimes |\mathfrak{t}|}, \langle \cdot, \cdot \rangle \right)$ is a dual pair. Therefore

14

the maps $\langle \mathbf{t}_1, \cdot \rangle_{\mathsf{t}}$ with $\mathbf{t}_1 \in V^{\otimes \mathsf{t}}$ separate $W^{\otimes \mathsf{t}}$. The converse statement is shown in exactly the same way and hence the proof of (i) is finished.

In order to prove (ii), we equip $V^{\otimes \mathsf{t}}$ with the weak topology induced by $\langle \cdot, \cdot \rangle_{\mathsf{t}}$, i.e. with the weakest topology which makes the maps $\langle \cdot, \mathbf{t}_2 \rangle_{\mathsf{t}}$ continuous for all $\mathbf{t}_2 \in W^{\otimes \mathsf{t}}$. Denote this topology by τ_1. Thus $W^{\otimes \mathsf{t}}$ is the topological dual of $V^{\otimes \mathsf{t}}$ with respect to τ_1 and therefore it is also the topological dual of the completion of $V^{\otimes \mathsf{t}}$ under τ_1. As a consequence $W^{\otimes \mathsf{t}}$ separates this completion. Since we assume that the pairing between V and W is continuous with respect to their locally convex topologies, τ_1 is weaker than any topology on $V^{\otimes \mathsf{t}}$ that is induced by cross-norms as introduced in Section 3. Therefore the space $V^{\hat{\otimes} \mathsf{t}}$ is contained in the τ_1-completion of $V^{\otimes \mathsf{t}}$ and hence $W^{\otimes \mathsf{t}}$ also separates $V^{\hat{\otimes} \mathsf{t}}$. Thus we have shown that $\left(V^{\hat{\otimes} \mathsf{t}}, W^{\otimes \mathsf{t}}, \langle \cdot, \cdot \rangle_{\mathsf{t}} \right)$ is a dual pair. ∎

Example 4.4

$$= \langle v_1, w_1 \rangle \langle v_2, w_2 \rangle \langle v_3, w_3 \rangle + \langle v_1, w_1 \rangle \langle v_3, w_2 \rangle \langle v_2, w_3 \rangle$$

Definition 4.5 Let A and B be two Hopf algebras. We say that they form a *dual pair of Hopf algebras* with the pairing $\langle \cdot, \cdot \rangle : A \times B \to \mathbf{R}$ if $(A, B, \langle \cdot, \cdot \rangle)$ is a dual pair of vector spaces which satisfies the conditions

$$\langle \Delta(a), b_1 \otimes b_2 \rangle = \langle a, b_1 b_2 \rangle \qquad \text{and} \qquad \langle a_1 \otimes a_2, \Delta(b) \rangle = \langle a_1 a_2, b \rangle \tag{9}$$

for all $a, a_1, a_2 \in A$ and $b, b_1, b_2 \in B$.

Remark 4.6 An alternative way to formulate (9) is the following. Consider the linear maps

$$\varphi : A \to B^* : a \mapsto \langle a, \cdot \rangle \qquad \text{and} \qquad \psi : B \to A^* : b \mapsto \langle \cdot, b \rangle,$$

where A^* and B^* denote the (algebraic) duals of A and B respectively. Then (9) holds if and only if both φ and ψ are algebra morphisms.

Lemma 4.7 *Let* $(A, B, \langle \cdot, \cdot \rangle)$ *be a dual pair of Hopf algebras. Let* $g \in A$ *be group-like, i.e. such that* $\Delta(g) = g \otimes g$. *Then for all elements* $b_1, \ldots, b_k \in B$ *we have*

$$\langle g, b_1 \cdots b_k \rangle = \langle g, b_1 \rangle \cdots \langle g, b_k \rangle$$

Proof. Let $k = 2$. Then we have

$$\langle g, b_1 b_2 \rangle = \langle \Delta(g), b_1 \otimes b_2 \rangle = \langle g \otimes g, b_1 \otimes b_2 \rangle = \langle g, b_1 \rangle \langle g, b_2 \rangle.$$

The claim for $k > 2$ follows by induction. ∎

Theorem 4.8 *Let $(V, W, \langle \cdot, \cdot \rangle)$ be a dual pair of vector spaces. Define*

$$\langle \cdot, \cdot \rangle_{\mathfrak{Tr}} : \mathfrak{Tr}(V) \times \mathfrak{Tr}(W) \to \mathbf{R} : (\mathbf{t}_1, \mathbf{t}_2) \mapsto \delta_{\mathsf{t}_1, \mathsf{t}_2} \langle \mathbf{t}_1, \mathbf{t}_2 \rangle_{\mathsf{t}_1}$$

for $\mathbf{t}_1 \in V^{\otimes \mathsf{t}_1}$ and $\mathbf{t}_2 \in W^{\otimes \mathsf{t}_2}$. Then the following hold.

(i) *$(\mathcal{H}_{\mathrm{GL}}(V), \mathcal{H}_{\mathrm{CK}}(W), \langle \cdot, \cdot \rangle_{\mathfrak{Tr}})$ is a dual pair of Hopf algebras.*

(ii) *If $(V, W, \langle \cdot, \cdot \rangle)$ is a dual pair of locally convex vector spaces, then $\langle \cdot, \mathbf{t}_2 \rangle_{\mathfrak{Tr}}$ can be continuously extended to $\widehat{\mathfrak{Tr}}(V)$ for all $\mathbf{t}_2 \in \mathfrak{Tr}(W)$ and both $\left(\widehat{\mathcal{H}}_{\mathrm{GL}}(V), \mathcal{H}_{\mathrm{CK}}(W), \langle \cdot, \cdot \rangle_{\mathfrak{Tr}} \right)$ and $\left(\widehat{\mathcal{H}}_{\mathrm{CK}}(V), \mathcal{H}_{\mathrm{GL}}(W), \langle \cdot, \cdot \rangle_{\mathfrak{Tr}} \right)$ are dual pairs of Hopf algebras.*

Proof. We start with the proof of (i), which means we need to show two things. Firstly we have to make sure that $\mathfrak{Tr}(V)$ and $\mathfrak{Tr}(W)$ separate each other. This follows immediately from Lemma 4.3. Secondly we have to show that the pairing satisfies (9). This has been shown in [17, Proposition 4.4] for the special case $V = \mathbf{R}$, i.e. for unlabelled trees.

We will extend their arguments to the labelled case.

Let $\mathsf{t}_1, \mathsf{t}_2, \mathsf{s} \in \mathfrak{Tr}$ be trees and let $\mathbf{t}_1 \in V^{\otimes \mathsf{t}}$, $\mathbf{t}_2 \in V^{\otimes \mathsf{t}_2}$ and $\mathbf{s} \in W^{\otimes \mathsf{s}}$ be elementary trees with elementary representatives t_1, t_2 and s respectively.

It is an easy exercise to show that both $\mathcal{H}_{\mathrm{GL}}(V)$ and $\mathcal{H}_{\mathrm{CK}}(V)$ are generated as an algebra by the elements of $\mathfrak{Tr}_1(V)$. In light of Remark 4.6 it is therefore sufficient to show

$$\langle \mathbf{t}_1 \star \mathbf{t}_2, \mathbf{s} \rangle_{\mathfrak{Tr}} = \langle \mathbf{t}_1 \otimes \mathbf{t}_2, \Delta_{\mathrm{CK}}(\mathbf{s}) \rangle_{\mathfrak{Tr}} \quad \text{and} \quad \langle \mathbf{t}_1 \circ \mathbf{t}_2, \mathbf{s} \rangle_{\mathfrak{Tr}} = \langle \mathbf{t}_1 \otimes \mathbf{t}_2, \Delta_{\mathrm{GL}}(\mathbf{s}) \rangle_{\mathfrak{Tr}}$$

for $\mathsf{t}_1 \in \mathfrak{Tr}_1$. By definition the first equality is equivalent to

$$
\begin{aligned}
&\sum_{d \in \mathrm{Gr}(\mathsf{t}_1, \mathsf{t}_2; \mathsf{s})} \sum_{\sigma \in \mathrm{SG}(\mathsf{s})} \langle t_1 \overset{d}{\curvearrowright} t_2, \sigma(s) \rangle \\
&= \sum_{C \in \mathrm{Cut}(\mathsf{s}; \mathsf{t}_1, \mathsf{t}_2)} \sum_{\sigma_1 \in \mathrm{SG}(\mathsf{t}_1)} \sum_{\sigma_2 \in \mathrm{SG}(\mathsf{t}_2)} \langle \sigma_1(t_1) \otimes \sigma_2(t_2), \mathbf{P}^C(s) \otimes \mathbf{R}^C(s) \rangle,
\end{aligned}
\tag{10}
$$

where we denote by $\mathrm{Gr}(\mathsf{t}_1, \mathsf{t}_2; \mathsf{s})$ the set of all grafting maps d such that $\mathsf{t}_1 \overset{d}{\curvearrowright} \mathsf{t}_2 = \mathsf{s}$ and by $\mathrm{Cut}(\mathsf{s}; \mathsf{t}_1, \mathsf{t}_2)$ the set of all cuts C such that $P^C(\mathsf{s}) = \mathsf{t}_1$ and $R^C(\mathsf{s}) = \mathsf{t}_2$.

Denote by c the unique child of the root of $\mathsf{t}_1 \in \mathfrak{Tr}_1$. Any $d \in \mathrm{Gr}(\mathsf{t}_1, \mathsf{t}_2)$ is uniquely determined by the vertex of t_2 to which c is attached. Let us now rewrite the left-hand side of (10). Instead of summing over all $d \in \mathrm{Gr}(\mathsf{t}_1, \mathsf{t}_2; \mathsf{s})$ we can alternatively fix an arbitrary $\bar{d} \in \mathrm{Gr}(\mathsf{t}_1, \mathsf{t}_2; \mathsf{s})$, denote the associated vertex of t_2 by v and rewrite the expression as

$$\frac{1}{|\mathrm{Fix}(v, \mathsf{t}_2)|} \sum_{\pi \in \mathrm{SG}(\mathsf{t}_2)} \sum_{\sigma \in \mathrm{SG}(\mathsf{s})} \langle t_1 \overset{\bar{d}}{\curvearrowright} \pi(t_2), \sigma(s) \rangle, \tag{11}$$

where $\mathrm{Fix}(v, \mathsf{t}_2) := \{ \sigma \in \mathrm{SG}(\mathsf{t}_2) : \sigma(v) = v \}$. In order to verify that (11) is indeed equal to the left hand side of (10) we observe that

$$\sum_{\sigma \in \mathrm{SG}(\mathsf{s})} \langle t_1 \overset{\bar{d}}{\curvearrowright} \pi(t_2), \sigma(s) \rangle = \sum_{\sigma \in \mathrm{SG}(\mathsf{s})} \langle t_1 \overset{\bar{d}}{\curvearrowright} \tilde{\pi}(t_2), \sigma(s) \rangle$$

holds for $\pi, \tilde{\pi} \in \mathrm{SG}(\mathfrak{t}_2)$ whenever $\pi^{-1}(v) = \tilde{\pi}^{-1}(v)$. For fixed $\pi \in \mathrm{SG}(\mathfrak{t}_2)$ there are $|\mathrm{Fix}(v, \mathfrak{t}_2)|$ ways to choose $\tilde{\pi}$ in such a way.

On the other hand, any cut $C \in \mathrm{Cut}(\mathfrak{s}; \mathfrak{t}_1, \mathfrak{t}_2)$ contains exactly one vertex and hence we can choose $\overline{C} \in \mathrm{Cut}(\mathfrak{s}; \mathfrak{t}_1, \mathfrak{t}_2)$ such that for the canonical injection $\phi_{\overline{C}, P} : \mathrm{v}(\mathfrak{t}_1) \to \mathrm{v}(\mathfrak{s})$ we have that $w := \phi_{\overline{C}, R}(c)$ is the unique vertex which defines $\overline{C}$. Thus we can rewrite the right-hand side of (10) as

$$\frac{1}{|\mathrm{Fix}(w, \mathfrak{s})|} \sum_{\tau \in \mathrm{SG}(\mathfrak{s})} \sum_{\sigma_1 \in \mathrm{SG}(\mathfrak{t}_1)} \sum_{\sigma_2 \in \mathrm{SG}(\mathfrak{t}_2)} \left\langle \sigma_1(t_1) \otimes \sigma_2(t_2), \mathbf{P}^{\overline{C}}(\tau(s)) \otimes \mathbf{R}^{\overline{C}}(\tau(s)) \right\rangle, \qquad (12)$$

where $\mathrm{Fix}(w, \mathfrak{s}) := \{\sigma \in \mathrm{SG}(\mathfrak{s}) : \sigma(w) = w\}$. This can verified as in the previous paragraph. Since $\phi_{\overline{C}, P}$ induces a canonical embedding of $\mathrm{SG}(\mathfrak{t}_1)$ into $\mathrm{SG}(\mathfrak{s})$ as a subgroup we can rewrite (12) as

$$\frac{|\mathrm{SG}(\mathfrak{t}_1)|}{|\mathrm{Fix}(w, \mathfrak{s})|} \sum_{\tau \in \mathrm{SG}(\mathfrak{s})} \sum_{\sigma_2 \in \mathrm{SG}(\mathfrak{t}_2)} \left\langle t_1 \otimes \sigma_2(t_2), \mathbf{P}^{\overline{C}}(\tau(s)) \otimes \mathbf{R}^{\overline{C}}(\tau(s)) \right\rangle$$

$$= \frac{|\mathrm{SG}(\mathfrak{t}_1)|}{|\mathrm{Fix}(w, \mathfrak{s})|} \sum_{\tau \in \mathrm{SG}(\mathfrak{s})} \sum_{\sigma_2 \in \mathrm{SG}(\mathfrak{t}_2)} \left\langle t_1 \overset{\overline{d}}{\to} \sigma_2(t_2), \tau(s) \right\rangle$$

One easily verifies that $|\mathrm{Fix}(w, \mathfrak{s})| = |\mathrm{Fix}(v, \mathfrak{t}_2)||\mathrm{SG}(\mathfrak{t}_1)|$ holds, which shows (10).

The second condition in (9), i.e. the relationship between $\circ$ and Δ_{GL}, is shown analogously.

Finally, (ii) is an easy consequence of (i). ∎

Corollary 4.9 *Let V be a locally convex vector space and let $\mathbf{t} \in \widehat{\mathcal{H}}_{\mathrm{GL}}(V)$ be group-like i.e. such that it satisfies $\Delta_{\mathrm{GL}}(\mathbf{t}) = \mathbf{t} \otimes \mathbf{t}$. Then we have*

$$\mathbf{t} = \exp_\circ(\pi_1(\mathbf{t})) := \sum_{k=0}^{\infty} \frac{(\pi_1(\mathbf{t}))^{\circ k}}{k!}.$$

Proof. Let V' be the continuous dual of V, let $\mathfrak{s} \in \mathfrak{Tr}$ and let $\mathbf{s} \in (V')^{\otimes \mathfrak{s}}$ be an elementary labelled tree with decomposition $\mathbf{s} = \mathbf{s}_1 \circ \cdots \circ \mathbf{s}_k$ as in (6). Then we have

$$\langle \mathbf{t}, \mathbf{s} \rangle_{\mathfrak{Tr}} = \langle \mathbf{t}, \mathbf{s}_1 \circ \cdots \circ \mathbf{s}_k \rangle_{\mathfrak{Tr}}$$
$$= \langle \mathbf{t}, \mathbf{s}_1 \rangle_{\mathfrak{Tr}} \cdots \langle \mathbf{t}, \mathbf{s}_k \rangle_{\mathfrak{Tr}}$$
$$= \langle \pi_1(\mathbf{t}), \mathbf{s}_1 \rangle_{\mathfrak{Tr}} \cdots \langle \pi_1(\mathbf{t}), \mathbf{s}_k \rangle_{\mathfrak{Tr}},$$

where we have used first Lemma 4.7 in combination with Theorem 4.8 and then the fact that $\mathbf{s}_i \in \mathfrak{Tr}_1(V)$ for all $i \in \{1, \ldots, k\}$. On the other hand we have

$$\left\langle (\pi_1(\mathbf{t}))^{\circ k}, \mathbf{s} \right\rangle_{\mathfrak{Tr}} = \left\langle (\pi_1(\mathbf{t}))^{\circ k}, \mathbf{s}_1 \circ \cdots \circ \mathbf{s}_k \right\rangle_{\mathfrak{Tr}}$$

$$=k!\langle \pi_1(\mathbf{t}), \mathbf{s}_1\rangle_{\mathfrak{Tr}} \cdots \langle \pi_1(\mathbf{t}), \mathbf{s}_k\rangle_{\mathfrak{Tr}},$$

which follows from (8) and the multinomial formula.

Since we have chosen an arbitrary $\mathbf{s}$ and since $\mathfrak{Tr}(V')$ separates the points of $\widehat{\mathfrak{Tr}}(V)$ this shows that we have $\pi_k(\mathbf{t}) = \frac{1}{k!}(\pi_1(\mathbf{t}))^{\circ k}$ for all $k \in \mathbf{N}_0$. Hence the claim follows. ∎

Corollary 4.10 *Let V be a locally convex vector space and let $\mathbf{t} \in \widehat{\mathcal{H}}_{\mathrm{GL}}(V)$ be group-like i.e. such that it satisfies $\Delta_{\mathrm{GL}}(\mathbf{t}) = \mathbf{t} \otimes \mathbf{t}$. Then we have*

$$\pi_k^{(n)}(\mathbf{t}) = \frac{\left(\pi_1^{(n)}(\mathbf{t})\right)^{\circ k}}{k!} + \mathfrak{r},$$

where $\mathfrak{r}$ is such that $\pi^{(n)}(\mathfrak{r}) = 0$.

Proof. This follows immediately from the previous corollary. ∎

5 Change of variables for differential equations driven by branched rough paths in arbitrary dimension

For a general introduction to the theory of rough paths we refer to [10, 11, 20]. In this section we assume that E is a Banach space, which implies that $\widehat{\mathfrak{Tr}}^{(k)}(E)$ is also Banach for all $k \in \mathbf{N}$. For $T > 0$ define $\Delta_T := \{(s,t) \in [0,T] : s < t\}$. By a *control function* or simply *control* we mean a continuous function $\omega : \Delta_T \to \mathbf{R}$ which satisfies $\omega(s,s) = 0$ and $\omega(s,t) + \omega(t,u) \leq \omega(s,u)$ for all $0 \leq s \leq t \leq u \leq T$.

Definition 5.1 Let $T > 0$, let $p \geq 1$ and let ω be a control function. A continuous map $\mathbf{x} : \Delta_T \to \widehat{\mathfrak{Tr}}^{([p])}(E)$ is called a *branched p-rough path controlled by ω* if it satisfies

(i) $\mathbf{x}_{s,u} = \mathbf{x}_{s,t} \star \mathbf{x}_{t,u}$

(ii) $\mathbf{x}_{s,t} = \pi^{([p])}(\mathbf{g})$ for some $\mathbf{g} \in \widehat{\mathfrak{Tr}}(E)$ with $\Delta_{\mathrm{GL}}(\mathbf{g}) = \mathbf{g} \otimes \mathbf{g}$

(iii) $\left\|\pi^k(\mathbf{x}_{s,t})\right\| \leq \omega(s,t)^{\frac{k}{p}}$

for all $0 \leq s \leq t \leq u \leq T$ and $k \in \{1, \ldots, [p]\}$.

Let F be another Banach space and let $V : E \times F \to F$ be a continuous map which is linear in the first argument and smooth (in the Fréchet sense) in the second argument. There are two ways to view such a map. The one that is chosen in [20] is to interpret V as a map from F to the space of linear maps from E to F. However, we prefer to view V as a continuous linear map from E into the space of smooth vector fields on F. To emphasise this point of view we write V_e instead of $V(e, \cdot)$ for $e \in E$. Note that the the vector fields V_e for $e \in E$ are *kinetic vector fields* in the language of [18].

Let furthermore $f : F \to G$ be a smooth map into some Banach space G. Then for $y \in F$ the linear map

$$\Psi_{V,f,y} : \widehat{\mathfrak{Tr}}(E) \to G$$

is constructed in the following way. It maps a labelled tree $\mathbf{t}$ to a product of derivatives of f and V, evaluated at y. The root of $\mathbf{t}$ corresponds to f and a vertex labelled by $e \in E$ corresponds to the vector field V_e. The number of children of a vertex indicates how often the corresponding term has to be differentiated. The terms corresponding to the children are then plugged into the derivative.

Example 5.2

$$\Psi_{V,f,y}\left(\begin{array}{c}\begin{tikzpicture}\end{tikzpicture}\end{array}\right) = f''(y)[V_{e_1}'''(y)[V_{e_3}(y), V_{e_4}(y), V_{e_5}(y)], V_{e_2}(y)]$$

There is a particular choice of f which will become important. For any positive integer k we define the map

$$\mathrm{id}^k : F \to F^{\otimes k} : x \mapsto x^{\otimes k}.$$

The maps $\Psi_{V,\mathrm{id}^k,y}$ have two important properties. First they satisfy

$$\Psi_{V,f,y}(\mathbf{t}) = \frac{1}{k!} f^{(k)}(y)[\Psi_{V,\mathrm{id}^k,y}(\mathbf{t})] \tag{13}$$

for all $\mathbf{t} \in \widehat{\mathfrak{Tr}}(E)$. Second they are compatible with the $\circ$-product of labelled trees in the sense that for two trees $\mathbf{t}_1 \in \widehat{\mathfrak{Tr}}_k(E)$ and $\mathbf{t}_2 \in \widehat{\mathfrak{Tr}}_l(E)$ we have

$$\Psi_{V,\mathrm{id}^{k+l},y}(\mathbf{t}_1 \circ \mathbf{t}_2) = \frac{(k+l)!}{k!l!} \mathrm{sym}_{\mathfrak{S}_{k+l}}\left(\Psi_{V,\mathrm{id}^k,y}(\mathbf{t}_1) \otimes \Psi_{V,\mathrm{id}^l,y}(\mathbf{t}_2)\right). \tag{14}$$

Theorem 5.3 *Let $V : E \times F \to F$ be as above. Let furthermore $\mathbf{x}$ be a branched p-rough path controlled by ω and let $y : [0, T] \to F$ be a continuous path. Then the following are equivalent.*

(i) There exists a function $\theta : \mathbf{R} \to \mathbf{R}$ with $\theta(h)/h \to 0$ for $h \to 0$ such that

$$\|y_t - \Psi_{V,\mathrm{id},y_s}(\mathbf{x}_{s,t})\| \leq \theta(\omega(s,t)) \tag{15}$$

holds for all $(s,t) \in \Delta_T$.

(ii) For every Banach space G and every smooth $f : F \to G$ there exists a function $\theta_f : \mathbf{R} \to \mathbf{R}$ with $\theta_f(h)/h \to 0$ for $h \to 0$ such that

$$\|f(y_t) - \Psi_{V,f,y_s}(\mathbf{x}_{s,t})\| \leq \theta_f(\omega(s,t)) \tag{16}$$

holds for all $(s,t) \in \Delta_T$.

Proof. The implication $(ii) \Rightarrow (i)$ is trivial. For the other direction, the idea is to split the term $\Psi_{V,f,y_s}(\mathbf{x}_{s,t})$ into terms that we can control. We have (see explanations below)

$$\|f(y_t) - \Psi_{V,f,y_s}(\mathbf{x}_{s,t})\|$$

$$\leq \left\| \sum_{k=1}^{\lfloor p \rfloor} \frac{f^{(k)}(y_s)}{k!} \left[y_{s,t}^{\otimes k} - \Psi_{V,\mathrm{id}^k,y_s}(\pi_k(\mathbf{x}_{s,t})) \right] \right\| + C_1 \|y_{s,t}\|^{\lfloor p \rfloor + 1} \tag{17a}$$

$$\leq \left\| \sum_{k=1}^{\lfloor p \rfloor} \frac{f^{(k)}(y_s)}{k!} \left[y_{s,t}^{\otimes k} - \frac{1}{k!} \Psi_{V,\mathrm{id}^k,y_s}\left((\pi_1(\mathbf{x}_{s,t}))^{\circ k} \right) \right] \right\| + C_1 \|y_{s,t}\|^{\lfloor p \rfloor + 1} + C_2 \omega(s,t)^{\frac{\lfloor p \rfloor + 1}{p}} \tag{17b}$$

$$= \left\| \sum_{k=1}^{\lfloor p \rfloor} \frac{f^{(k)}(y_s)}{k!} \left[y_{s,t}^{\otimes k} - \left(\Psi_{V,\mathrm{id},y_s}(\pi_1(\mathbf{x}_{s,t})) \right)^{\otimes k} \right] \right\| + C_1 \|y_{s,t}\|^{\lfloor p \rfloor + 1} + C_2 \omega(s,t)^{\frac{\lfloor p \rfloor + 1}{p}} \tag{17c}$$

$$= \left\| \sum_{k=1}^{\lfloor p \rfloor} \frac{f^{(k)}(y_s)}{k!} \left[(y_{s,t} - \Psi_{V,\mathrm{id},y_s}(\pi_1(\mathbf{x}_{s,t}))) \otimes \sum_{i=0}^{k-1} y_{s,t}^{\otimes(k-1-i)} \Psi_{V,\mathrm{id},y_s}(\pi_1(\mathbf{x}_{s,t}))^{\otimes i} \right] \right\|$$
$$+ C_1 \|y_{s,t}\|^{\lfloor p \rfloor + 1} + C_2 \omega(s,t)^{\frac{\lfloor p \rfloor + 1}{p}} \tag{17d}$$

$$= \left\| \sum_{k=1}^{\lfloor p \rfloor} \frac{f^{(k)}(y_s)}{k!} \left[(y_t - \Psi_{V,\mathrm{id},y_s}(\mathbf{x}_{s,t})) \otimes \sum_{i=0}^{k-1} y_{s,t}^{\otimes(k-1-i)} \Psi_{V,\mathrm{id},y_s}(\pi_1(\mathbf{x}_{s,t}))^{\otimes i} \right] \right\|$$
$$+ C_1 \|y_{s,t}\|^{\lfloor p \rfloor + 1} + C_2 \omega(s,t)^{\frac{\lfloor p \rfloor + 1}{p}}. \tag{17e}$$

In this computation we have used the following arguments.

(17a) Replace f by its Taylor approximation up to degree $\lfloor p \rfloor$ and make use of (13).

(17b) Apply Corollary 4.10, where the remainder $\mathfrak{r}$ leads to the term that involves ω.

(17c) Use (14), where we can ignore the symmetrisation because $f^{(k)}(y_s)$ is a symmetric multilinear form anyway.

(17d) Use the elemetary identity $a^k - b^k = (a - b)(a^{k-1} + a^{k-2}b + \cdots + ab^{k-2} + b^{k-1})$, which holds in principle only for commuting variables a and b. It can be applied nevertheless, again because $f^{(k)}(y_s)$ is symmetric.

(17e) Observe that Ψ_{V,id,y_s} maps any tree which is neither in $\widehat{\mathfrak{Tr}}_1(E)$ nor in $E^{\otimes \mathbf{1}}$ to zero.

Now define the constant C_3 by

$$C_3 := \sup_{0 \leq s < t \leq T} \sup_{a \in F: \|a\| = 1} \left\| \sum_{k=1}^{\lfloor p \rfloor} \frac{f^{(k)}(y_s)}{k!} \left[a \otimes \sum_{i=0}^{k-1} y_{s,t}^{\otimes(k-1-i)} \Psi_{V,\mathrm{id},y_s}(\pi_1(\mathbf{x}_{s,t}))^{\otimes i} \right] \right\|.$$

We have $C_3 < \infty$ because $f^{(k)}(y)$ is a continuous multilinear map for all $k \in \mathbf{N}$ and all $y \in F$. Thus (17e) can be estimated by

$$C_3 \|y_t - \Psi_{V,\mathrm{id},y_s}(\mathbf{x}_{s,t})\| + C_1 \|y_{s,t}\|^{\lfloor p \rfloor + 1} + C_2 \omega(s,t)^{\frac{\lfloor p \rfloor + 1}{p}}$$
$$\leq C_3 \theta(\omega(s,t)) + C_4 \omega(s,t)^{\frac{\lfloor p \rfloor + 1}{p}},$$

for some $C_4 > 0$ where we have also used that y has finite p-variation controlled by (a scalar multiple of) ω. Defining

$$\theta_f(h) := C_3 \theta(h) + C_4 h^{\frac{\lfloor p \rfloor + 1}{p}}$$

finishes the proof. ∎

Definition 5.4 With the notation as in Theorem 5.3 we say that y solves the rough differential equation

$$\mathrm{d}y_t = V(y_t)\mathrm{d}\mathbf{x}_t \tag{18}$$

with initial value y_0 if (i) (and thus also (ii)) of Theorem 5.3 holds.

The idea to define solutions of rough differential equations via Euler expansions is originally due to Davie [9] and was later developed further in [11]. Note that, in contrast to Lyons' original work [19, 20] and Gubinelli's theory of controlled rough paths [14], this approach defines a solution as an ordinary F-valued path without any higher order terms.

5.1 Geometric rough paths and differential geometry

Even though we have formulated Theorem 5.3 only for branched rough paths it is also valid for (weakly) geometric rough path since the latter can be interpreted as branched rough paths in a canonical way. The relationship between branched and geometric rough paths has been discussed in detail in [16], at least for the finite-dimensional case. The crucial difference between branched and geometric rough paths is that geometric rough paths take their values in the tensor algebra over E rather than the Grossman-Larson algebra.

Let us define the tensor algebra over E and see how it embeds into the Grossman-Larson algebra. For the algebraic details we refer to [20, 22]. Recall that we have fixed a symmetric system of cross-norms on the tensor powers of E in order to define $\widehat{\mathfrak{Tr}}(E)$. Thus we can equip

$$T(E) := \bigoplus_{k=0}^{\infty} V^{\hat{\otimes}k}$$

with the product topology and denote by $\widehat{T}(E)$ its completion. The space $\widehat{T}(E)$ carries a natural product, namely the tensor product, which turns it into an algebra. Note that the

sub-algebra generated by E is dense in $\widehat{T}(E)$. Thus we can define the coproduct $\Delta_{\sqcup\!\sqcup}$ on E by

$$\Delta_{\sqcup\!\sqcup}(e) := 1 \otimes e + e \otimes 1$$

and then extend it to all of $\widehat{T}(E)$ as a continuous algebra morphism.

Now define $\mathfrak{t} = \mathbf{\bullet\!\!\;}$ and note that $E^{\otimes \mathfrak{t}} \simeq E$. Hence we can define the continuous algebra morphism

$$\Phi : \widehat{T}(E) \to \widehat{\mathcal{H}}_{\mathrm{GL}}(E)$$

which is given by $\Phi(e) = e \in E^{\otimes \mathfrak{t}}$ for $e \in E$. One easily checks that this is in fact a Hopf algebra morphism. Thus Φ lets us embed the space of geometric rough paths into the space of branched rough paths.

In [16] it has in fact been shown that conversely, one can interpret every branched rough path as a geometric one, albeit this geometric rough path will then take values in the tensor algebra over an *augmented* version of the original vector space.

Now assume that V is a linear map from E into the space of vector fields on some smooth manifold $\mathcal{M}$. Let furthermore $y \in \mathcal{M}$ and let $f : \mathcal{M} \to G$ a smooth map into some Banach space G. In this case the 'non-geometricity' of general branched rough paths becomes apparent in the definition of the map $\Psi_{V,f,y}$ that we have defined above. Let $e_1, \ldots, e_k \in E$. Then we have

$$\Psi_{V,f,y}\left(\Phi(e_1 \otimes \cdots \otimes e_k)\right) = \Psi_{V,f,y}\left(\,{}^{e_1}\!\mathbf{\bullet\!\!\;} \star \cdots \star {}^{e_k}\!\mathbf{\bullet\!\!\;}\,\right) = V_{e_1} \cdots V_{e_k} f(y),$$

because the $\star$-product simply encodes the product rule. Thus when we say that a rough path $\mathbf{x}$ is geometric, this is because $\Psi_{V,f,y}(\Phi(\mathbf{x}_{s,t}))$ is a canonically defined object for all V, f, y, s, t, even in a non-linear setting. In contrast, if $\mathbf{x}$ is a branched rough path, then $\Psi_{V,f,y}(\mathbf{x}_{s,t})$ is in general only well-defined once we choose an affine connection on $\mathcal{M}$ because it involves covariant derivatives of vector fields and higher order derivatives of f. Hence it is the independence of the choice of a connection which makes the geometric rough paths special among all branched rough paths. Differential equations driven by geometric rough paths can be interpreted as generalised Stratonovich equations whereas differential equations driven by branched rough paths should be seen as a generalisation of Ito-corrected Stratonovich equations.

In light of the previous paragraph we will write $\Psi_{V,f,y,\nabla}$ in order to make clear that the values of the function on arbitrary labelled trees are to be computed with respect to the connection ∇. To align this with the notation used above, we agree that ∇ is the canonical connection of a vector space if it is not mentioned explicitly. Thus we generalise Definition 5.4 in the following way.

Definition 5.5 With the notation as in Theorem 5.3 we say that y solves the rough differential equation

$$\mathrm{d}y_t = V(y_t)\mathrm{d}\mathbf{x}_t \tag{19}$$

with respect to the connection ∇ on F and with initial value y_0 if there exists a function $\theta : \mathbf{R} \to \mathbf{R}$ with $\theta(h)/h \to 0$ for $h \to 0$ such that

$$\|y_t - \Psi_{V,\mathrm{id},y_s,\nabla}(\mathbf{x}_{s,t})\| \leq \theta(\omega(s,t)) \tag{20}$$

holds for all $(s,t) \in \Delta_T$.

To conclude, let us mention two consequences of Theorem 5.3 which underline its significance in the context of differential geometry. First, we have the invariance of rough differential equations under changes of coordinates as stated in the following corollary.

Corollary 5.6 *Let the notation be as in Theorem 5.3 and assume that $y_t \in U$ for some open set $U \subset F$ and all $t \in [0,T]$. Let $W \subset F$ be another open set and let $\varphi : U \to W$ be a diffeomorphism. For each $e \in E$ define the vector field $\tilde{V}_e$ on W by*

$$\tilde{V}_e(p) = \varphi'(\varphi^{-1}(p))[V_e(\varphi^{-1}(p))].$$

Then y solves (19) (with respect to the canonical connection on F) if and only if $\tilde{y}$ given by $\tilde{y}_t := \varphi(y_t)$ solves

$$\mathrm{d}\tilde{y}_t = \tilde{V}(\tilde{y}_t)\mathrm{d}\mathbf{x}_t,$$

with respect to the pushforward connection $\tilde{\nabla}$ given by

$$\tilde{\nabla}_{\tilde{Y}}\tilde{X}(\tilde{p}) := \tilde{X}'(\tilde{p})[\tilde{Y}(\tilde{p})] - \varphi''(\varphi^{-1}(\tilde{p}))[(\varphi'(\varphi^{-1}(\tilde{p})))^{-1}[\tilde{X}(\tilde{p})], (\varphi'(\varphi^{-1}(\tilde{p})))^{-1}[\tilde{Y}(\tilde{p})]]$$

for vector fields $\tilde{X}, \tilde{Y}$ on F and $\tilde{p} \in W$.

Proof. Writing φ itself in the new coordinates on W gives $\tilde{\varphi}(\tilde{p}) = \varphi(\varphi^{-1}(\tilde{p})) = \tilde{p}$, i.e. $\tilde{\varphi} = \mathrm{id}$. Thus we have

$$\tilde{y}_t - \Psi_{\tilde{V},\mathrm{id},\tilde{y}_s,\tilde{\nabla}}(\mathbf{x}_{s,t}) = \varphi(y_t) - \Psi_{\tilde{V},\tilde{\varphi},\tilde{y}_s,\tilde{\nabla}}(\mathbf{x}_{s,t}) = \varphi(y_t) - \Psi_{V,\varphi,y_s}(\mathbf{x}_{s,t}),$$

where the last equality is obtained by writing everything in the original coordinates. The claim follows now immediately from Theorem 5.3. ∎

A careful reading of [5] shows that the authors have used arguments similar to those that went into the proof of Theorem 5.3 in order to obtain well-definedness of RDEs on submanifolds of $\mathbf{R}^d$, albeit in the less general case $p < 3$.

Second, Theorem 5.3 yields a way to *define* what we mean by the solution of an RDE on a manifold by testing (16) for a suitable set of real-valued test functions. If the manifold is finite-dimensional, then smooth compactly supported test functions are an obvious choice. Even though this definition has been suggested before – see e.g. [1] and the references therein – we believe that Theorem 5.3 is so far the most general way to establish that such a definition is compatible with Definitions 5.4 and 5.5.

References

[1] Jürgen Angst, Ismaël Bailleul, and Camille Tardif, *Kinetic Brownian motion on Riemannian manifolds*, Electron. J. Probab. **20** (2015).

[2] Ismaël Bailleul, *Flows driven by rough paths*, Rev. Mat. Iberoam. **31** (2015), no. 3, 901–934.

[3] Yvain Bruned, Ilya Chevyrev, Peter K. Friz, and Rosa Preiss, *A Rough Path Perspective on Renormalization*, arXiv:1701.01152 (2017).

[4] John C. Butcher, *An algebraic theory of integration methods*, Math. Comp. **26** (1972), no. 117, 79–106.

[5] Thomas Cass, Bruce K. Driver, and Christian Litterer, *Constrained rough paths*, Proceedings of the London Mathematical Society **111** (2015), no. 6, 1471–1518.

[6] Arthur Cayley, *On the Analytical Forms Called Trees*, Am. J. Math. **4** (1881), no. 1, 266–268.

[7] Ilya Chevyrev and Terry Lyons, *Characteristic functions of measures on geometric rough paths*, Ann. Probab. **44** (2016), no. 6, 4049–4082.

[8] Alain Connes and Dirk Kreimer, *Hopf Algebras, Renormalization and Noncommutative Geometry*, Commun. Math. Phys. **199** (1998), no. 1, 203–242.

[9] Alexander M. Davie, *Differential Equations Driven by Rough Paths: An Approach via Discrete Approximation*, Applied Mathematics Research eXpress **2008** (2008).

[10] Peter K. Friz and Martin Hairer, *A course on rough paths*, Universitext, Springer, Cham, 2014.

[11] Peter K. Friz and Nicolas B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010.

[12] Robert Grossman and Richard G Larson, *Hopf-algebraic structure of families of trees*, J. Algebra **126** (1989), no. 1, 184–210.

[13] Alexander Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Memoirs of the American Mathematical Society, vol. 16, American Mathematical Society, Providence, 1955.

[14] Massimiliano Gubinelli, *Controlling rough paths*, Journal of Functional Analysis **216** (2004), no. 1, 86–140.

[15] _____ , *Ramification of rough paths*, J. Diff. Equations **248** (2010), no. 4, 693–721.

[16] Martin Hairer and David Kelly, *Geometric versus non-geometric rough paths*, Ann. I. H. Poincaré-PR **51** (2015), no. 1, 207–251.

[17] Michael E. Hoffman, *Combinatorics of Rooted Trees and Hopf Algebras*, T. Am. Math. Soc. **355** (2003), no. 9, 3795–3811.

[18] Andreas Kriegl and Peter W. Michor, *The Convenient Setting of Global Analysis*, Mathematical Surveys and Monographs, vol. 53, American Mathematical Society, Providence, 1997.

[19] Terry J. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoam. **14** (1998), no. 2, 215–310.

[20] Terry J. Lyons, Michael J. Caruana, and Thierry Lévy, *Differential Equations Driven by Rough Paths*, Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007.

[21] Florin Panaite, *Relating the Connes–Kreimer and Grossman–Larson Hopf Algebras Built on Rooted Trees*, Lett. Math. Phys. **51** (2000), no. 3, 211–219.

[22] Christophe Reutenauer, *Free Lie algebras*, London Mathematical Society Monographs. New Series, vol. 7, Clarendon Press, Oxford, 1993.

[23] Alexander Robertson and Wendy Robertson, *Topological Vector Spaces*, Cambridge Tracts in Mathematics and Mathematical Physics, vol. 53, Cambridge University Press, Cambridge, 1980.

[24] Robert S Strichartz, *The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations*, J. Funct. Anal. **72** (1987), no. 2, 320–345.