

Cartan's Path Development, the Logarithmic Signature and a Conjecture of Lyons-Sidorova

Horatio Boedihardjo*, Xi Geng[†] and Sheng Wang[‡]

Abstract

The signature transform, which is defined in terms of iterated path integrals, provides a faithful representation of the group of tree-reduced geometric rough paths. While the signature coefficients are known to decay factorially fast, the coefficients of the logarithmic signature generically only possess geometric decay. It was conjectured by T. Lyons and N. Sidorova [20] that the only tree-reduced paths with bounded variation (BV) whose logarithmic signature can have infinite radius of convergence (R.O.C.) are straight lines. This conjecture was confirmed in the same work for two special classes of paths and the general BV case remains unsolved.

The aim of the present article is to develop a deeper understanding towards the Lyons-Sidorova conjecture. In the first place, we show that having infinite R.O.C. for the logarithmic signature is a conjugacy class property and as a consequence, the aforementioned conjecture needs to be modified accordingly.

Our main result is that if the logarithmic signature has infinite R.O.C., the signature coefficients must satisfy an infinite system of rigid algebraic identities defined in terms of iterated integrals along complex exponential one-forms. These iterated integral identities impose strong geometric constraints on the underlying path, and in some special situations, confirm the conjecture.

As a non-trivial application of our integral identities, we prove a weak version of the conjecture, which asserts that if the logarithmic signature of a BV path has infinite radius of convergence over all sub-intervals of time, the underlying path must live on a straight line.

Our methodology relies on Cartan's path development onto the complex semisimple Lie algebras $\mathfrak{sl}_m(\mathbb{C})$. The special root patterns of $\mathfrak{sl}_m(\mathbb{C})$ allow one to project the infinite-dimensional Baker-Campbell-Hausdorff (BCH) formula in

*Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom.
Email: horatio.boedihardjo@warwick.ac.uk.

[†]School of Mathematics and Statistics, University of Melbourne, Parkville VIC 3010, Australia.
Email: xi.geng@unimelb.edu.au.

[‡]School of Mathematics and Statistics, University of Melbourne, Parkville VIC 3010, Australia.
Email: shewang4@student.unimelb.edu.au.

a very special finite dimensional manner to yield meaningful quantitative relations between BCH-type singularities and the vanishing of certain iterated path integrals.

Contents

1	Introduction	3
1.1	Motivation: a conjecture of T. Lyons and N. Sidorova	4
1.2	Summary of main results and novelty	5
1.3	Further questions	7
2	The logarithmic signature transform	8
2.1	Definitions and basic properties	8
2.2	The Lyons-Sidorova conjecture	12
3	First order integral identities	14
3.1	Two immediate applications	17
3.2	Proof of Theorem 3.1	19
3.2.1	The Baker-Campbell-Hausdorff formula	19
3.2.2	An adjoint representation of signature	20
3.2.3	Basic notions on Cartan’s path development	21
3.2.4	A two-dimensional development	22
4	A weak version of the LS conjecture	27
4.1	An application of the integral condition to winding number	28
4.2	A zero winding lemma	30
4.3	Proof of Theorem 4.1	33
5	Second order integral identities	39
5.1	A second order extension of the line integral condition	39
5.2	Homogeneous projection formulae	42
5.3	Complex semisimple Lie algebras	45
5.4	Proof of Theorem 5.1 through $\mathfrak{sl}_3(\mathbb{C})$ -development	46
6	Higher order iterated integral identities	49
6.1	A Chen-Strichartz type formula for $D_m\tilde{L}$	50
6.2	$\mathfrak{sl}_{m+1}(\mathbb{C})$ -development of the logarithmic signature	51
6.3	Proof of Theorem 6.5	57
6.4	The analyticity lemma	59
6.4.1	A recursive formula for the Hausdorff series	59
6.4.2	A combinatorial identity	62
6.4.3	Proof of Lemma 6.15	63

1 Introduction

The *signature transform* (or simply the signature) of a multidimensional continuous path $\gamma : [0, T] \rightarrow \mathbb{R}^d$ with bounded variation (BV) is the formal tensor series

$$S(\gamma) \triangleq \left(1, \gamma_T - \gamma_0, \int_{0 < s < t < T} d\gamma_s \otimes d\gamma_t, \dots, \int_{0 < t_1 < \dots < t_n < T} d\gamma_{t_1} \otimes \dots \otimes d\gamma_{t_n}, \dots\right)$$

defined by the global iterated path integrals of all orders. This path transformation was originally introduced by the geometer K.T. Chen [9] in the 1950s to study the topology of loop spaces over manifolds. The signature transform plays a fundamental role in Lyons' rough path theory as well as its broad applications to problems in stochastic analysis and more recently, in machine learning. Its theoretical importance partly comes from the *signature uniqueness theorem*, which asserts that every rough path is uniquely determined by its signature up to tree-like equivalence (cf. [8, 16, 3]). A probabilistic counterpart of this theorem was obtained by Chevyrev-Lyons [10]. The monographs [14, 19] contain excellent exposition on the mathematical theory of the signature transform and some applications to stochastic analysis are discussed e.g. in [1, 18, 21]. Some recent applications of signature-based methods in machine learning are contained e.g. in [11, 24, 27] (and references therein).

While the signature encodes essential geometric information about the underlying path, it contains a lot of algebraic redundancies since different components of the signature are related by some universal algebraic constraints (the shuffle product formula). An effective way of removing such algebraic dependencies is to pass to the so-called *logarithmic signature transform*. The two objects are naturally isomorphic, however, it was a deep theorem of Chen [7] that the logarithmic signature is a free Lie series and thereby does not contain algebraic relations among its components. This leads to dimension-reduction from a practical viewpoint. An important motivation for studying the logarithmic signature is related to the study of controlled differential equations, which will be elaborated further in Section 1.1 below. In recent years, methods based on the logarithmic signature have also been developed to study various problems in machine learning (cf. [12, 29] and references therein).

Analytic properties of the signature and logarithmic signature are very different on the other hand. It is well-known (which is rather trivial in the BV case) that the signature components decay factorially fast with respect to the degree. However, it is highly non-trivial that components of the logarithmic signature generically only possess geometric decay. This is like the infinite dimensional analogue of the elementary fact that the exponential function e^z is entire while the logarithmic function $\log(1+z)$ only has a finite radius of convergence (R.O.C.). Understanding such a property, even just in the BV case, is still an unsolved open problem in rough path theory.

1.1 Motivation: a conjecture of T. Lyons and N. Sidorova

The logarithmic signature arises naturally when one attempts to solve a controlled differential equation (CDE) using the so-called *log-ODE method*. Consider the following CDE

$$dX_t = \sum_{i=1}^d V_i(X_t) d\gamma_t^i, \quad 0 \leq t \leq T \quad (1.1)$$

on a differentiable manifold M , where V_1, \dots, V_d are smooth vector fields on M and γ_t is a smooth \mathbb{R}^d -valued path. At a formal level, the flow of diffeomorphisms on M induced by (1.1) admits a “logarithm” in the Lie algebra $\mathcal{L}(V_1, \dots, V_d)$ of smooth vector fields generated by the V_i ’s. Such a logarithm can be explicitly expressed in terms of the logarithmic signature of γ .

To elaborate this, let W be a smooth vector field on M . We use $\exp(W)$ to denote the diffeomorphism given by the time-one map of the flow induced by W . In other words,

$$\exp(W) : M \rightarrow M, \quad \exp(W)x = y_1$$

where $(y_t)_{0 \leq t \leq 1}$ solves the ODE

$$\dot{y}_t = W(y_t), \quad y_0 = x.$$

The solution at time T to the CDE (1.1) can formally be expressed as

$$X_T = \exp\left(\sum_I \Lambda_I(\gamma) V_{[I]}\right)(x). \quad (1.2)$$

Here the summation is taken over all words $I = (i_1, \dots, i_n)$ where $n \geq 1$ and $i_j \in \{1, \dots, d\}$. The numbers $\Lambda_I(\gamma)$ are the logarithmic signature coefficients of γ ; one has

$$\log S(\gamma) = \sum_{I=(i_1, \dots, i_n)} \Lambda_I(\gamma) [e_{i_1}, [e_{i_2}, \dots, [e_{i_{n-1}}, e_{i_n}] \dots]].$$

The vector fields $V_{[I]}$ are defined by $V_{[I]} \triangleq [V_{i_1}, [V_{i_2}, \dots, [V_{i_{n-1}}, V_{i_n}]]]$. The relation (1.2) is often known as the *Chen-Strichartz formula* (cf. [1, 28]).

However, the formula (1.2) is only formal and it requires additional analytic assumptions on $\log S(\gamma)$ and the vector fields V_i to make it precise. It is accurate if γ is a straight line or if the Lie algebra $\mathcal{L}(V_1, \dots, V_d)$ is nilpotent (in both cases, the series (1.2) becomes a finite sum). In general, one expects that the series (1.2) should converge if the “size” of the vector fields is within the R.O.C. of $\log S(\gamma)$. This makes the R.O.C. for the logarithmic signature a natural object of study. However, it is still unclear how one can make such intuition mathematically precise. In practice, one usually applies a truncated version of (1.2) over a partition of $[0, T]$ to obtain an approximating solution (numerical scheme) and tries to prove convergence when the mesh size is sent to zero (cf. [1, 13]).

Cartan's path development

There is a simple yet useful situation where the analysis can be made very precise. Let G be a matrix Lie group with Lie algebra \mathfrak{g} . Let $V = (V_1, \dots, V_d)$ be d left invariant vector fields on G . The Lie algebra $\mathcal{L}(V_1, \dots, V_d)$ is now a sub-algebra of \mathfrak{g} . One can prove that the series (1.2) is convergent if $\|V\| < \text{R.O.C. of } \log S(\gamma)$, where certain matrix norms has to be taken and fixed. It is known from [10] that the R.O.C. is strictly positive and in particular, the series (1.2) is always convergent (in the current Lie group setting) as long as the vector fields V_i are "small enough". The resulting solution X_t to the CDE (1.1) defines a path in the Lie group G which is known as *Cartan's path development* of γ in G . If γ is a straight line, say $\gamma_t = tw$ ($w \in \mathbb{R}^d, t \in [0, T]$), the formula (1.2) is well-defined for *all* left invariant vector fields. The solution X_t is just the one-parameter subgroup defined by

$$X_t = \exp\left(t \sum_{i=1}^d w^i V_i\right),$$

where \exp is now the exponential map for G . This aligns with the fact that $\log S(\gamma)$ has infinite R.O.C.

Initiated by the celebrated work of Lyons-Sidorova [20], it has been widely believed that *straight lines are the only (tree-reduced) BV paths whose logarithmic signatures can have infinite R.O.C.* The logic in Lyons-Sidorova's argument can be summarised as follows. Suppose that $\log S(\gamma)$ has infinite R.O.C. Then the end point X_T of Cartan's development of γ into any Lie group always admits a logarithm, namely, X_T is always an element in the image of the exponential map. In other words, if one is able to construct a special Cartan's development such that X_T turns out to be outside Im exp , this immediately suggests that $\log S(\gamma)$ only has finite R.O.C. Of course, this method requires the use of (non-compact) Lie groups with non-surjective exponential map (e.g. the special linear group $\text{SL}_2(\mathbb{R})$). Lyons and Sidorova [20] proved the above claim for two special classes of paths based on such a geometric viewpoint: paths which are strictly increasing in one particular direction (what they called *1-monotone paths*) and a generic class of piecewise linear paths (which they called *non-double paths*). They conjectured that the claim should be true in the most general BV setting. To the best of our knowledge, this problem remains unsolved.

1.2 Summary of main results and novelty

The main goal of the present article is to study the aforementioned Lyons-Sidorova (LS) conjecture in depth. Although we have not resolved the conjecture in its full generality, in what follows we summarise our main findings and discuss the novelty of our work.

Main results. The general philosophy of our results is that having infinite R.O.C. for the logarithmic signature leads to rigid algebraic constraints on signature coefficients and thus geometric constraints on the underlying path.

(A) The very first (and important) comment is that the conjecture turns out to be a *conjugacy class property*. As a result, it needs to be modified accordingly (for otherwise, it cannot be true in its original form). See Section 2.2 (in particular, Proposition 2.6) below for the details.

(B) Our first main result is that *if the logarithmic signature of a non-closed path has infinite R.O.C., the path has to satisfy a specific type of line integral identities*. Such integral identities naturally lead to geometric constraints on the underlying path, and in certain special situations, yield the conclusion of the conjecture. See Section 3 below for the precise formulation.

(C) Our second main result is a weak version of the conjecture which holds in the most general BV setting. More precisely, we prove that *if the logarithmic signature of an arbitrary BV path has infinite R.O.C. over any time interval $[s, t]$, the path must live on a straight line*. This result is a non-trivial application of Theorem 3.1 together with fine analysis of path-geometry based on winding number considerations. See Section 4 below for more details.

(D) Our last main result is the higher order extension of (B). We show that if the logarithmic signature of a non-closed path has infinite R.O.C., the path has to satisfy *an infinite system of specific iterated integral identities*. These iterated integral identities impose much stronger geometric constraints on the path than the line integral identities given by Theorem 3.1 do. In particular, they allow one to handle paths that cannot be detected by Theorem 3.1 (cf. Example 5.6 below).

Novelty. Apart from the nature of the aforementioned main results, an important aspect of novelty in the current work is the methodology we develop. The original technique of Lyons-Sidorova [20] has a geometric nature; they studied the geometric behaviour of the lifted path in the Lie group and showed that the lifted path will eventually leave the image of the exponential map under some very special choice of Cartan's development.

We also work with Cartan's development as a starting point, however, our method is quite different in nature from the one developed in [20]. Essentially, the finiteness of R.O.C. property is closely related to how the Baker-Campbell-Hausdorff (BCH) formula exhibits singularity properties in a certain sense. This problem is already quite delicate in the simplest case of a piecewise linear path with two edges, say v, w (which corresponds to the classical BCH formula for $\log e^v e^w$). The Lyons-Sidorova conjecture can be viewed as an infinite dimensional version of singularity analysis for the BCH formula.

A main novelty of the current work is that such singularity properties can be studied through *Cartan's developments into complex semisimple Lie algebras*. It turns out that the special root patterns of semisimple Lie algebras allows one to project the infinite dimensional BCH formula in a very special (finite dimensional) manner, so that one can perform singularity analysis at various explicit and quantitative levels.

Our approach thus has a strong algebraic nature in contrast to Lyons-Sidorova's

geometric perspective. Stated in vague terms, after performing suitable Cartan's development at the Lie algebra level, the logarithmic signature $\log S(\gamma)$ is transformed into linear combinations of products between certain path functionals, say $S_m(a_1, \dots, a_m)$ (defined through iterated path integrals) and meromorphic functions, say $\phi(a_1, \dots, a_m)$ (arising from the Hausdorff series). If $\log S(\gamma)$ has infinite R.O.C., its Cartan's development must define an entire function. As a consequence, at the singularities of the function $\phi(a_1, \dots, a_m)$ one must have $S_m(a_1, \dots, a_m) = 0$. This is precisely how the integral identities in our main theorems arise.

Cartan's development was first used by Hambly-Lyons [16] for studying the signature uniqueness problem in the BV setting. Such a method was used in various works related to the study of (both deterministic and random) rough paths and their signatures (see e.g. [2, 6, 10, 22]). The use of Cartan's development at the logarithmic signature level first appeared in [4] for the study of signature asymptotics for pure rough paths.

1.3 Further questions

There are a few basic questions which naturally arise from the current work but remain to be answered.

1. Our weak version of LS conjecture is proved in Theorem 4.1 for BV paths only. It is reasonable to expect that for weakly geometric rough paths, the logarithmic signature has infinite R.O.C. over all time intervals $[s, t]$ if and only if the path is a *pure rough path*, i.e. the exponential of a line segment in the space of Lie polynomials. Is such an extension true?
2. It is not particularly clear what exact geometric constraints are imposed on the path by Theorems 5.1, 6.5 and whether they eventually lead to the conclusion that the path is conjugate to a line segment. We expect that there is still a non-trivial gap which requires deeper analysis and new ideas.
3. Theorems 5.1, 6.5 only hold for non-closed paths, which is indeed a limitation of the current analysis. Can one extend the current approach to the case when the underlying path is a loop? The modified LS conjecture in the context of loops takes a particularly elegant form: the logarithmic signature of a tree-reduced, closed, BV path γ has infinite R.O.C. if and only if γ is constant.
4. Let $\xi \in G((V))$ be a group-like element. What are the exact analytic conditions on ξ so that ξ is the signature of a BV path over V ? At the signature level, having exact factorial decay seems to be the natural analytic constraint on ξ . However, at the logarithmic signature level one encounters the finiteness of R.O.C. property. Are these two analytic properties related to each other and how do they yield a suitable characterisation of the image of the signature transform?

2 The logarithmic signature transform

In this section, we review the background materials. In particular, we discuss several basic properties of the signature and logarithmic signature transforms where most details can be found in [14, 19]. We also recall the precise formulation of the Lyons-Sidorova conjecture from [20], which provides the main source of motivation for the current work.

2.1 Definitions and basic properties

Let V be a finite dimensional normed vector space over $\mathbb{K} = \mathbb{R}$ or \mathbb{C} . We define the infinite tensor algebra

$$T((V)) \triangleq \prod_{n=0}^{\infty} V^{\otimes n}$$

where $V^{\otimes 0} \triangleq \mathbb{K}$ as a convention. Addition in $T((V))$ is defined component-wisely and multiplication is defined by

$$(\xi \otimes \eta)_n \triangleq \sum_{k=0}^n \xi_k \otimes \eta_{n-k}, \quad n \geq 0$$

where $\xi = (\xi_0, \xi_1, \dots)$, $\eta = (\eta_0, \eta_1, \dots)$ are given tensor series. This makes $T((V))$ into an associative \mathbb{K} -algebra with unit $\mathbf{1} \triangleq (1, 0, 0, \dots)$. Elements in $T((V))$ are known as *formal tensor series* over V . We use

$$\pi_n : T((V)) \rightarrow V^{\otimes n}, \quad \pi^{(n)} : T((V)) \rightarrow \bigoplus_{k=0}^n V^{\otimes k}$$

to denote the canonical projections.

Any tensor series of the form $\xi = (1, \xi_1, \xi_2, \dots)$ has a multiplicative inverse given by

$$\xi^{-1} = \sum_{n=0}^{\infty} (-1)^n (\xi - \mathbf{1})^{\otimes n}.$$

The above series is well-defined since the summation is locally finite (i.e. the projection onto the m -th component for each fixed m only involves finitely many nonzero terms).

There are two basic operations over $T((V))$ that will be important to us. To define them, let us introduce two subspaces:

$$T_0((V)) \triangleq \{\xi = (\xi_0, \xi_1, \dots) \in T((V)) : \xi_0 = 0\}, \quad T_1((V)) \triangleq \{\xi \in T((V)) : \xi_0 = 1\}.$$

The *exponential* and *logarithmic* transforms are defined by

$$e^{(\cdot)} : T_0((V)) \rightarrow T_1((V)), \quad e^{\xi} \triangleq \sum_{n=0}^{\infty} \frac{1}{n!} \xi^{\otimes n};$$

$$\log(\cdot) : T_1((V)) \rightarrow T_0((V)), \quad \log \xi \triangleq \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (\xi - \mathbf{1})^{\otimes n}.$$

It can be shown that these two functions are inverse to each other. It is also easy to see that $(e^\xi)^{-1} = e^{-\xi}$.

Definition 2.1. Let $\gamma : [0, T] \rightarrow V$ be a continuous path with bounded variation. The *signature transform* of γ (or simply the signature) is the formal tensor series defined by

$$S(\gamma) \triangleq \left(1, \gamma_T - \gamma_0, \int_{0 < s < t < T} d\gamma_s \otimes d\gamma_t, \dots, \int_{0 < t_1 < \dots < t_n < T} d\gamma_{t_1} \otimes \dots \otimes d\gamma_{t_n}, \dots\right),$$

where the path integrals are defined in the Lebesgue-Stieltjes sense. The tensor series $S(\gamma)$ is known as the *logarithmic signature* of γ .

One can also consider the signature $S(\gamma)_{s,t}$ over $[s, t] \subseteq [0, T]$ by replacing the above iterated integrals with the ones over $[s, t]$. It is easily check that

$$S(\gamma)_{s,u} = S(\gamma)_{s,t} \otimes S(\gamma)_{t,u} \quad \forall s \leq t \leq u \in [0, T]. \quad (2.1)$$

This is known as *Chen's identity*. Stated in a more general form, one has

$$S(\alpha \sqcup \beta) = S(\alpha) \otimes S(\beta),$$

where $\alpha \sqcup \beta$ means the concatenation between the two paths α, β . It can also be shown that $S(\overleftarrow{\gamma}) = S(\gamma)^{-1}$ where $\overleftarrow{\gamma}$ denotes the reversal of γ .

Algebraic properties

There are universal algebraic dependencies among different signature components, which are precisely described by the so-called *shuffle product formula*:

$$X^m \otimes X^n = \sum_{\sigma \in \mathcal{P}(m,n)} P_\sigma(X^{m+n}) \quad \forall m, n \in \mathbb{N}. \quad (2.2)$$

Here X^m is the m -th component of $S(\gamma)$. $\mathcal{P}(m, n)$ denotes the set of (m, n) -shuffles, i.e. the set of permutations σ of order $m + n$ satisfying

$$\sigma(1) < \dots < \sigma(m), \quad \sigma(m+1) < \dots < \sigma(m+n).$$

Given any permutation σ , the operator P_σ is the linear transformation over $V^{\otimes(m+n)}$ induced by

$$P_\sigma(v_1 \otimes \dots \otimes v_{m+n}) \triangleq v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(m+n)}.$$

The set of permutations of order m is denoted as \mathcal{S}_m .

Definition 2.2. A tensor series $\xi = (1, \xi_1, \xi_2, \dots) \in T_1((V))$ is *group-like* if it satisfies the relation (2.2) (with X^n replaced by ξ_n). The space of group-like elements is denoted as $G((V))$.

An important reason for considering the logarithmic signature is that there are no algebraic dependencies among its components, because it takes values in the “free” Lie algebra. This is the content of the celebrated *Chen’s theorem*. To state this theorem, we first define

$$\mathcal{L}((V)) \triangleq \prod_{n=1}^{\infty} \mathcal{L}_n(V) \subseteq T_0((V))$$

where $\mathcal{L}_1(V) \triangleq V$ and

$$\mathcal{L}_{n+1}(V) \triangleq [\mathcal{L}_n(V), V] \triangleq \text{Span}\{[\xi, v] : \xi \in \mathcal{L}_n(V), v \in V\}$$

for all $n \geq 1$. Here $[\cdot, \cdot]$ is the commutator defined by $[\xi, \eta] \triangleq \xi \otimes \eta - \eta \otimes \xi$. Elements in $\mathcal{L}((V))$ are *formal Lie series* and elements in $\bigoplus_{k=1}^n \mathcal{L}_k(V)$ are *Lie polynomials* of degree n . Chen’s theorem is stated as follows (cf. [7, 26]).

Theorem 2.3. *A tensor series $\xi \in T_1((V))$ is group-like if and only if $\log \xi \in \mathcal{L}((V))$.*

One can also consider truncated versions of signatures and logarithmic signatures. Let $T^{(N)}(V)$, $G^{(N)}(V)$ and $\mathcal{L}^{(N)}(V)$ be the truncations up to level N (i.e. taking the first N components) of the previous infinite dimensional spaces. The tensor product and Lie bracket restrict to these truncated spaces in the obvious way.

Analytic properties

We also need to deal with analytic properties of the signature. For this purpose, we need to consider suitable tensor norms on the tensor product spaces. From now on, we assume that $V^{\otimes n}$ is equipped with a given norm $\|\cdot\|_n$. The norms $\{\|\cdot\|_n\}$ are *admissible* in the sense that

$$\|\xi \otimes \eta\|_{m+n} \leq \|\xi\|_m \|\eta\|_n \quad \forall m, n \in \mathbb{N}$$

and

$$\|P_\sigma(\xi)\|_n = \|\xi\|_n \quad \forall \xi \in V^{\otimes n}, \sigma \in \mathcal{S}_n.$$

An explicit example is the Hilbert-Schmidt tensor norm on $V^{\otimes n}$ induced by an Euclidean (or Hermitian if $\mathbb{K} = \mathbb{C}$) metric on V . What plays a basic role in rough path theory is the so-called projective norm.

Definition 2.4. The *projective tensor norm* of $\xi \in V^{\otimes n}$ is defined by

$$\|\xi\|_{n;\text{proj}} \triangleq \inf \left\{ \sum_j |v_1^j|_V \cdots |v_n^j|_V \right\},$$

where the infimum is taken over all possible representations of ξ as linear combinations of tensor monomials:

$$\xi = \sum_j v_1^j \otimes \cdots \otimes v_n^j, \quad v_i^j \in V.$$

Throughout the rest of this article, we will be exclusively using the projective tensor norm. For simplicity, we will also omit the subscripts if no confusion will be caused.

The signature transform can be defined for arbitrary rough paths. Let $p \geq 1$ be a given fixed number. A *p-rough path* is a continuous functional

$$\mathbf{X}_{\cdot,\cdot} = (1, X_{\cdot,\cdot}^1, \dots, X_{\cdot,\cdot}^{[p]}) : \Delta_T \triangleq \{(s, t) : 0 \leq s \leq t \leq T\} \rightarrow T^{([p])}(V)$$

such that

$$\mathbf{X}_{s,u} = \mathbf{X}_{s,t} \otimes \mathbf{X}_{t,u} \quad \forall s \leq t \leq u \in [0, T]$$

and \mathbf{X} has finite p -variation in the sense that

$$\omega_{\mathbf{X}} \triangleq \|\mathbf{X}\|_{p\text{-var}}^p \triangleq \sum_{i=1}^{[p]} \sup_{\mathcal{P}} \sum_{t_i \in \mathcal{P}} \|X_{t_{i-1}, t_i}^i\|^{p/i} < \infty, \quad (2.3)$$

where the supremum is taken over all finite partitions \mathcal{P} of $[0, T]$. A p -rough path \mathbf{X} is *weakly geometric* if it takes values in $G^{([p])}(V)$. Sometimes we just call $\gamma : [0, T] \rightarrow V$ a rough path over V but its precise meaning is the multi-level functional \mathbf{X} (γ is the first level component of \mathbf{X}).

Let \mathbf{X} be a p -rough path. According to *Lyons' extension theorem* (cf. [23]), there exists a unique extension of $\mathbf{X}_{\cdot,\cdot}$ to a functional

$$S(\mathbf{X})_{\cdot,\cdot} = (1, X_{\cdot,\cdot}^1, \dots, X_{\cdot,\cdot}^n, \dots) : \Delta_T \rightarrow T((V)),$$

such that Chen's identity (2.1) holds for $S(\mathbf{X})_{\cdot,\cdot}$ in $T((V))$ and its truncation up to any level $n \geq [p]$ has finite p -variation in $T^{(n)}(V)$. The functional $S(\mathbf{X})_{\cdot,\cdot}$ is called the *signature path* of \mathbf{X} . The tensor series $S(\mathbf{X})_{0,T}$ (respectively, $\log S(\mathbf{X})_{0,T}$) is the *signature* (respectively, *logarithmic signature*) of \mathbf{X} . If \mathbf{X} is weakly geometric, it can be shown that its signature path takes values in $G((V))$.

An important analytic property of the signature is its rapid decay with respect to the degree. More precisely, there exists a universal number $\beta > 0$, such that the following estimate

$$\|X_{0,T}^n\| \leq \frac{\omega_{\mathbf{X}}^{n/p}}{\beta \cdot (n/p)!} \quad \forall n \geq 1 \quad (2.4)$$

holds for all p -rough paths. Here $(n/p)! \triangleq \Gamma(n/p+1)$ where $\Gamma(\cdot)$ is the Gamma function. When $p = 1$, \mathbf{X} is just a classical path in V with bounded variation and (2.4) follows trivially from the triangle inequality. The estimate (2.4) for the general rough paths is contained as part of Lyons' extension theorem.

2.2 The Lyons-Sidorova conjecture

To state the main question of our study, we first introduce the following definition.

Definition 2.5. Let $\xi = (\xi_0, \xi_1, \dots) \in T((V))$ be a given tensor series. Its *radius of convergence* (R.O.C.) is the radius of convergence for the power series

$$P_\xi(z) \triangleq \sum_{n=0}^{\infty} \|\xi_n\| z^n$$

in the standard real analysis sense.

The factorial decay (2.4) of signature shows that the signature of a rough path always has infinite R.O.C. However, a highly non-trivial fact is that the logarithmic signature (generically) only decays geometrically fast and should thus have a *finite* R.O.C. Consider a line segment $\gamma_t = tv$ ($0 \leq t \leq 1$) in V . It is immediate that $\log S(\gamma) = v$ and thus $\log S(\gamma)$ has infinite R.O.C. It was conjectured by T. Lyons and N. Sidorova [20] that these are the only tree-reduced BV paths whose logarithmic signatures can have infinite R.O.C.

The Lyons-Sidorova (LS) Conjecture. The logarithmic signature of a continuous, tree-reduced BV path γ has infinite R.O.C. if and only if γ is a line segment.

In their original work [20], the conjecture was confirmed for two special classes of paths: 1-monotone paths and non-double piecewise linear paths. The general BV case is still an unsolved open problem in rough path theory.

Before developing our main results, the very first comment we shall make is that *the LS conjecture needs to be modified* to reflect the following simple observation.

Proposition 2.6. *Let \mathbf{X}, \mathbf{Y} be given rough paths. Then $\log S(\mathbf{X})$ has infinite R.O.C. if and only if $\log S(\mathbf{Y} \sqcup \mathbf{X} \sqcup \overleftarrow{\mathbf{Y}})$ has infinite R.O.C.*

Proof. Let us prove a more general claim. Suppose that $l \in T_0((V))$ and g_1, g_2 are signatures (of certain rough paths). Then l has infinite R.O.C. if and only if $g_1 \otimes l \otimes g_2$ has infinite R.O.C.

To prove this claim, since g_1, g_2 are signatures, one knows from Lyons' factorial decay estimate (2.4) that

$$\|\pi_n(g_i)\| \leq \frac{C \cdot \omega^{n/p}}{(n/p)!} \quad \forall n \geq 1, \quad i = 1, 2. \quad (2.5)$$

Here C, p, ω are suitable constants depending on the underlying paths defining g_1, g_2 . Let us denote $l' \triangleq g_1 \otimes l \otimes g_2$. Then one has

$$\pi_n(l') = \sum_{k=1}^n \sum_{r=0}^{n-k} \pi_r(g_1) \otimes l_k \otimes \pi_{n-k-r}(g_2),$$

where $l_k \triangleq \pi_k(l)$. Since the tensor norms are admissible, one finds that

$$\|\pi_n(l')\| \leq C^2 \sum_{k=1}^n \sum_{r=0}^{n-k} \frac{\omega^{r/p}}{(r/p)!} \frac{\omega^{(n-k-r)/p}}{((n-k-r)/p)!} \|l_k\|.$$

We now apply the following so-called *neo-classical inequality* (cf. [17]):

$$\sum_{i=0}^m \frac{1}{(i/p)!((m-i)/p)!} \leq \frac{p2^{m/p}}{(m/p)!} \quad \forall m \in \mathbb{N}, p \geq 1. \quad (2.6)$$

It follows that

$$\|\pi_n(l')\| \leq C^2 p \sum_{k=1}^n \frac{(2\omega)^{(n-k)/p}}{((n-k)/p)!} \|l_k\|. \quad (2.7)$$

Let $\rho > 0$ be given fixed. Since l has infinite R.O.C., there exists $K > 0$ such that $\|l_k\| \leq \rho^k$ for all $k > K$. Therefore, one has from (2.7) that

$$\begin{aligned} \|\pi_n(l')\| &\leq C^2 p \left(\sum_{k=1}^K \frac{(2\omega)^{(n-k)/p}}{((n-k)/p)!} \|l_k\| + \sum_{k=K+1}^n \frac{(2\omega)^{(n-k)/p}}{((n-k)/p)!} \rho^k \right) \\ &\leq C^2 p \left(\sum_{k=1}^K \frac{(2\omega)^{(n-k)/p}}{((n-k)/p)!} \|l_k\| + \rho^n \sum_{k=0}^{\infty} \frac{(2\omega\rho^{-p})^{k/p}}{(k/p)!} \right) \end{aligned}$$

for all $n > K$. When n is sufficiently large, one can ensure that

$$\sum_{k=1}^K \frac{(2\omega)^{(n-k)/p}}{((n-k)/p)!} \|l_k\| < \rho^n$$

and thus

$$\|\pi_n(l')\| \leq C^2 p \left(1 + \sum_{k=0}^{\infty} \frac{(2\omega\rho^{-p})^{k/p}}{(k/p)!} \right) \rho^n.$$

This shows that

$$\overline{\lim}_{n \rightarrow \infty} \|\pi_n(l')\|^{1/n} \leq \rho.$$

Since ρ is arbitrary, one concludes that l' has infinite R.O.C.

In the context of the proposition, one takes

$$l = \log S(\mathbf{X}), \quad g_1 = S(\mathbf{Y}), \quad g_2 = S(\overleftarrow{\mathbf{Y}}) = g_1^{-1}.$$

The crucial observation is that

$$S(\mathbf{Y} \sqcup \mathbf{X} \sqcup \overleftarrow{\mathbf{Y}}) = g_1 \otimes e^l \otimes g_1^{-1} = e^{g_1 \otimes l \otimes g_1^{-1}}$$

which implies that

$$\log S(\mathbf{Y} \sqcup \mathbf{X} \sqcup \overleftarrow{\mathbf{Y}}) = g_1 \otimes l \otimes g_1^{-1}.$$

The result follows from the above claim since g_1 is a signature. □

Proposition 2.6 shows that the logarithmic signature of any path that is conjugate to a line segment must have infinite R.O.C. Here we say that a path β is *conjugate to* γ if $\beta = \alpha \sqcup \gamma \sqcup \overleftarrow{\alpha}$ for some path α . This result suggests that having finite R.O.C. is a property of conjugate classes. This naturally leads to the following modification of the LS conjecture.

Modified LS Conjecture. The logarithmic signature of a continuous, tree-reduced BV path has infinite R.O.C. if and only if it is conjugate to a line segment.

Remark 2.7. *Since V is finite dimensional, the property of having infinite / finite R.O.C. does not depend on the specific choices of the (admissible) tensor norms.*

Remark 2.8. *It was proved by Chevyrev-Lyons [10] that the logarithmic signature of any rough path always has positive R.O.C.*

3 First order integral identities

In this section, we establish our first main result:

Infinite R.O.C. for log signature \implies A special type of line integral identities.

Although we could formulate the theorem in a more general form, the result is essentially stated in terms of two-dimensional projections. To make the statement as simple as possible, we will just (and always) assume that $\gamma_t = (x_t, y_t)_{0 \leq t \leq 1}$ is a weakly geometric rough path over \mathbb{R}^2 . In this section, we only work with *non-closed* paths (i.e. $\gamma_1 \neq \gamma_0$). In particular, we impose the following normalisation conditions:

$$x_0 = y_0 = 0, \quad x_1 = 1. \tag{3.1}$$

Note that γ is the first level component of the actual rough path which will not be referred to. The main result of this section is stated as follows.

Theorem 3.1. *Let γ be a weakly geometric rough path over \mathbb{R}^2 which satisfies the normalisation condition (3.1). Suppose that $\log S(\gamma)$ has infinite R.O.C. Then one has*

$$\int_0^1 e^{2k\pi i x_t} dy_t = 0 \tag{3.2}$$

for all nonzero integers k .

Remark 3.2. *If the condition $x_1 = 1$ is dropped (still assuming $x_1 \neq 0$), the conclusion (3.2) should be replaced by $\int_0^1 e^{2k\pi i x_t/x_1} dy_t = 0$. For a general path over V , the result is formulated in terms of arbitrary projections of γ onto two-dimensional subspaces of V where (x_t, y_t) are the coordinates of the projected path with respect to a suitable basis.*

Remark 3.3. *If one assumes additionally that $y_1 = 0$, then (3.2) holds for all $k \in \mathbb{Z}$. In this case, one has $\int_0^1 f(x_t) dy_t = 0$ for all smooth, 1-periodic functions f . This result is stronger than the one obtained by Lyons-Sidorova's method in [20]; the latter established the integral property (3.2) for $k \in 2\mathbb{Z} + 1$ which is difficult to be extended to cover even k 's using their method.*

A more general (and useful) formulation of Theorem 3.1 is given as follows.

Theorem 3.4. *Under the assumptions of Theorem 3.1, suppose further that $y_1 = 0$. Then one has*

$$\int_0^1 \Phi(d\gamma_t) = 0 \quad (3.3)$$

for all smooth one-forms $\Phi(x, y) = f(x, y)dx + g(x, y)dy$ on \mathbb{R}^2 which satisfy the following two conditions:

- (i) $\Phi(x + 1, y) = \Phi(x, y)$ for all $x, y \in \mathbb{R}$;
- (ii) $\int_0^1 f(x, y)dx = 0$ for all $y \in \mathbb{R}$.

To prove Theorem 3.4, we first present a basic lemma.

Lemma 3.5. *Let $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a linear map, where \mathbb{R}^2 is equipped with the standard Euclidean norm. Let $\hat{A} : (\mathbb{R}^2)^{\otimes n} \rightarrow (\mathbb{R}^2)^{\otimes n}$ be the unique linear map such that*

$$\hat{A}(v_1 \otimes \cdots \otimes v_n) = Av_1 \otimes \cdots \otimes Av_n.$$

Then one has

$$\|\hat{A}(w)\| \leq \|A\|_{\text{op}}^n \|w\|$$

for all $w \in (\mathbb{R}^2)^{\otimes n}$. Here $\|A\|_{\text{op}}$ denotes the operator norm of A and the tensor products are all equipped with the projective tensor norm.

Proof. Suppose that w admits a representation given by

$$w = \sum_{j=1}^N v_1^j \otimes \cdots \otimes v_n^j. \quad (3.4)$$

Then one has

$$\begin{aligned} \|\hat{A}(w)\| &= \left\| \sum_{j=1}^N \hat{A}(v_1^j \otimes \cdots \otimes v_n^j) \right\| \leq \left\| \sum_{j=1}^N A(v_1^j) \otimes \cdots \otimes A(v_n^j) \right\| \\ &\leq \sum_{j=1}^N \|A(v_1^j)\| \cdots \|A(v_n^j)\| \\ &\leq \sum_{j=1}^N \|A\|_{\text{op}}^n \|v_1^j\| \cdots \|v_n^j\|. \end{aligned}$$

Taking infimum over all $\{v_1^j, \dots, v_n^j\}_{j=1}^N$ satisfying (3.4) gives the Lemma. \square

Proof of Theorem 3.4. Let γ be a path satisfying $\gamma_0 = (0, 0)$ and $\gamma_1 = (1, 0)$. Suppose that $\log S(\gamma)$ has infinite R.O.C. One knows from Theorem 3.1 that $\int_0^1 e^{2k\pi i x_s} dy_s = 0$ for all $k \in \mathbb{Z}$. Given $k \in \mathbb{Z} \setminus \{0\}$ and $\lambda \in \mathbb{R}$, we define

$$\hat{\gamma}_s = \begin{pmatrix} 1 & \frac{\lambda}{2k\pi} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_s \\ y_s \end{pmatrix}. \quad (3.5)$$

Note that the n -th component of $\log S(\hat{\gamma})$ is obtained by applying \hat{A} to the n -th component of $\log S(\gamma)$ (A is the transform defined by the matrix in (3.5)). According to Lemma 3.5, $\log S(\hat{\gamma})$ has infinite R.O.C. with respect to the projective tensor norm. The new path $\hat{\gamma}$ also satisfies $\hat{\gamma}_0 = (0, 0)$ and $\hat{\gamma}_1 = (1, 0)$. As a result, one has

$$\int_0^1 e^{2k\pi i x_s + \lambda i y_s} dy_s = \int_0^1 e^{2k\pi i \hat{x}_s} d\hat{y}_s = 0 \quad (3.6)$$

for $k \in \mathbb{Z} \setminus \{0\}$ and $\lambda \in \mathbb{R}$. It is plain to check that (3.6) holds when $k = 0$ (and thus for all $k \in \mathbb{Z}$ and $\lambda \in \mathbb{R}$).

Now let

$$K_1 \triangleq \inf \{y_t : t \in [0, 1]\}, \quad K_2 \triangleq \sup \{y_t : t \in [0, 1]\}$$

and recall \mathbb{S}^1 is the topological circle obtained by identifying the points 0 with 1 on $[0, 1]$. Note that the set of functions

$$S = \{e^{2k\pi i x + \lambda i y} : k \in \mathbb{Z}, \lambda \in \mathbb{R}\}$$

separate points on the compact space $\mathbb{S}^1 \times [K_1, K_2]$ and its linear span is a sub-algebra of continuous functions. By the Stone-Weierstrass theorem, one knows that $\text{Span} S$ that is dense in $C(\mathbb{S}^1 \times [K_1, K_2], \mathbb{C})$. Since $g \rightarrow \int_0^1 g(x_s, y_s) dy_s$ is continuous with respect to the uniform norm, one concludes that $\int_0^1 g(x_s, y_s) dy_s = 0$ for any continuous function $g(x, y)$ that is 1-periodic in the x -variable. This gives the desired integral condition for the dy -integral.

To prove the corresponding relation for the dx -integral, let $k \in \mathbb{Z} \setminus \{0\}$ and $\lambda \in \mathbb{R} \setminus \{0\}$. Using integration by parts, one finds that

$$\begin{aligned} 0 &= \int_0^1 e^{2k\pi i x_s + \lambda i y_s} dy_s = \frac{1}{\lambda i} \int_0^1 e^{2k\pi i x_s} d e^{\lambda i y_s} \\ &= \frac{1}{\lambda i} \left(e^{2k\pi i x_s + \lambda i y_s} \Big|_0^1 - 2k\pi i \int_0^1 e^{2k\pi i x_s + \lambda i y_s} dx_s \right) \\ &= -\frac{k}{\lambda} \int_0^1 e^{2k\pi i x_s + \lambda i y_s} dx_s. \end{aligned}$$

Since $k \neq 0$, one obtains that

$$\int_0^1 e^{2k\pi i x_s + \lambda i y_s} dx_s = 0. \quad (3.7)$$

The continuity of the integral as a function of λ implies that the relation (3.7) holds for all $k \in \mathbb{Z} \setminus \{0\}$ and $\lambda \in \mathbb{R}$. By exactly the same Stone-Weierstrass argument, one concludes that $\int_0^1 f(x_s, y_s) dx_s = 0$ for any continuous function $f(x, y)$ which is 1-periodic in x and satisfies $\int_0^1 f(x, y) dx = 0$ for every $y \in \mathbb{R}$ (the latter constraint appears since the relation (3.7) excludes the zero Fourier mode).

Now the proof of Theorem 3.4 is complete. □

3.1 Two immediate applications

The line integral condition (3.2) (or more generally, (3.3)) implicitly leads to rigid geometric constraints on the underlying path, and in some special situations, confirms the LS conjecture. We use one class of examples to illustrate this point. A more inspiring application of Theorem 3.4 on path-geometry is given in Section 4 below where a weak version of the LS conjecture is proved.

Proposition 3.6. *Let $\gamma_t = (x_t, y_t)$ be a two-dimensional, continuous BV path satisfying the assumptions in Theorem 3.4. Suppose further that $0 \leq x_t \leq 1$ and γ is non-self-intersecting. Then there exists $\varepsilon \in \mathbb{R}$ such that $\gamma = (\varepsilon e_2) \sqcup e_1 \sqcup (-\varepsilon e_2)$, where $\{e_1, e_2\}$ is the canonical basis of \mathbb{R}^2 .*

Proof. Suppose on the contrary that γ is not of the form $(\varepsilon e_2) \sqcup e_1 \sqcup (-\varepsilon e_2)$. Then there exist two points $z = (x_1, y_1)$ and $w = (x_2, y_2)$ such that $0 < x_1, x_2 < 1$, $y_1 = y_2$ and $z \in \text{Im}\gamma$, $w \notin \text{Im}\gamma$. One can choose a small $\delta > 0$, such that by defining

$$U_\delta \triangleq \{(x, y) : |x - x_1| \vee |y - y_1| < \delta\}, \quad V_\delta \triangleq \{(x, y) : |x - x_2| \vee |y - y_2| < \delta\}$$

one has $U_\delta \cap V_\delta = \emptyset$ and $V_\delta \cap \text{Im}\gamma = \emptyset$.

By using the argument in [3], one can construct a smooth one-form φ which is compactly supported in U_δ such that $\int_0^1 \varphi(d\gamma_t) \neq 0$. We define the one-form ψ by

$$\psi(x, y) = \begin{cases} \varphi(x, y), & (x, y) \in U_\delta; \\ -\varphi(x - x_2 + x_1, y), & (x, y) \in V_\delta. \end{cases}$$

It is clear that ψ is a smooth one-form supported in $U_\delta \cup V_\delta$ and satisfies

$$\int_0^1 f(x, y) dx = 0$$

for all $y \in \mathbb{R}$, where $f(x, y)$ denotes the dx -coefficient of ψ . Let Φ denote the 1-periodic extension of ψ (in the x -direction). It follows that Φ satisfies Properties (i), (ii) in Theorem 3.4 but

$$\int_0^1 \Phi(d\gamma_t) = \int_0^1 \varphi(d\gamma_t) \neq 0.$$

This leads to a contradiction. Therefore, $\gamma = (\varepsilon e_2) \sqcup e_1 \sqcup (-\varepsilon e_2)$ for some $\varepsilon \in \mathbb{R}$. □

Remark 3.7. *The above argument fails if the image of γ is not contained in the strip $0 \leq x \leq 1$, since the geometry of γ outside the strip may affect the line integral against the one-form Φ constructed before in an unexpected way. It is possible to replace this geometric assumption by another kind of geometric restriction. But we will not pursue such discussions here.*

Another application of Theorem 3.1 is the Brownian motion case. This is an important example due to its basic role in the study of stochastic differential equations.

Proposition 3.8. *Let B_t be a standard d -dimensional Brownian motion over $[0, 1]$. Then with probability one, its logarithmic (Stratonovich) signature has finite R.O.C.*

Proof. It is obvious that one only needs to prove the claim for the first two components of B_t , say (X_t, Y_t) . According to Theorem 3.1 (see also Remark 3.2), one only needs to show that

$$\int_0^1 \sin(2\pi X_t/X_1) dY_t \neq 0 \text{ a.s.}$$

This can be obtained by proving the stronger claim that the random variable $Z \triangleq \int_0^1 \cos(2\pi X_t/X_1) dY_t$ admits a density with respect to the Lebesgue measure on \mathbb{R} .

We use Malliavin's calculus to prove such a claim. We assume that (X_t, Y_t) are realised on the canonical path space. To prove the claim, it suffices to show that the Malliavin derivative $DZ \neq 0$ a.s. (cf. [25]). Let h be any Cameron-Martin path (i.e. absolutely continuous with L^2 -derivative). Explicit calculation together with integration by parts shows that the Malliavin derivative of Z along the direction h is given by

$$D_h Z = \int_0^1 \left[\frac{2\pi(U_1 - U_t)}{X_1} - \frac{2\pi V}{X_1^2} + \sin\left(\frac{2\pi X_t}{X_1}\right) \right] dh_t,$$

where

$$U_t \triangleq \int_0^t \cos\left(\frac{2\pi X_t}{X_1}\right) dY_t, \quad V \triangleq \int_0^1 X_t \cos\left(\frac{2\pi X_t}{X_1}\right) dY_t.$$

We now choose

$$h_t \triangleq \int_0^t \left[\frac{2\pi(U_1 - U_s)}{X_1} - \frac{2\pi V}{X_1^2} + \sin\left(\frac{2\pi X_s}{X_1}\right) \right] ds.$$

It follows that

$$D_h Z = \int_0^1 \left[\frac{2\pi(U_1 - U_t)}{X_1} - \frac{2\pi V}{X_1^2} + \sin\left(\frac{2\pi X_t}{X_1}\right) \right]^2 dt.$$

The above integral is nonzero with probability one, for otherwise the integrand would be identically zero which is clearly not the case. Therefore, $DZ \neq 0$ a.s. \square

3.2 Proof of Theorem 3.1

The fundamental idea behind our proof of Theorem 3.1, as well as its higher order extensions (cf. Theorem 5.1 and Theorem 6.5 below), is based on the notion of *Cartan's path development*. Stated conceptually, we “develop” the logarithmic signature into a suitably chosen (complex) Lie algebra \mathfrak{g} , so that the Lie structure of \mathfrak{g} “projects” the logarithmic signature $\log S(\gamma)$ in a very special way to yield products between certain explicit meromorphic functions and line integrals along γ (e.g. see the relation (3.18) below). If $\log S(\gamma)$ has infinite R.O.C., its development must produce an entire function. As a result, the line integral must have zeros matching the singularities of the meromorphic function. This will naturally lead to the line integral condition (3.2) at the discrete locations $k \in \mathbb{Z} \setminus \{0\}$.

3.2.1 The Baker-Campbell-Hausdorff formula

Let $\{e_1, e_2\}$ be the standard basis of \mathbb{R}^2 . Our analysis relies crucially on the consideration of the path $\tilde{\gamma} \triangleq \gamma \sqcup \overleftarrow{e_1}$ as a starting point (here \sqcup means concatenation). Clearly, the following relation holds on $T((\mathbb{R}^2))$:

$$e^{L(\gamma)} = e^{\tilde{L}(\gamma)} \otimes e^{e_1}, \quad (3.8)$$

where $L(\gamma), \tilde{L}(\gamma)$ denote the logarithmic signatures of $\gamma, \tilde{\gamma}$ respectively. Our next step is to apply the Baker-Campbell-Hausdorff (BCH) formula to the right hand side of (3.8), so that one can write

$$e^{\tilde{L}(\gamma)} \otimes e^{e_1} = e^{B(\tilde{L}(\gamma), e_1)}, \quad (3.9)$$

where the BCH functional

$$B(\xi, \eta) = \xi + \eta + \frac{1}{2}[\xi, \eta] + \frac{1}{12}[\xi, [\xi, \eta]] - \frac{1}{12}[\eta, [\xi, \eta]] + \dots$$

can be expressed in terms of commutators in a universal way.

We now recall the precise definition of the BCH functional $B(\cdot, \cdot)$. Let v, w be two letters and let V be the vector space generated by them. The quantity $B(v, w)$ is the Lie series defined by the following formula:

$$B(v, w) = \sum_{n=1}^{\infty} H_n(v, w) \in T((V)). \quad (3.10)$$

Here

$$H_1(v, w) \triangleq \sum_{m=0}^{\infty} \frac{B_m}{m!} (\text{ad}_w)^m(v), \quad (3.11)$$

where $\{B_m\}$ are the *Bernoulli numbers* defined by the expansion

$$\frac{z}{e^z - 1} = \sum_{m=0}^{\infty} \frac{B_m}{m!} z^m$$

and

$$\text{ad}_w(v) \triangleq [w, v] \triangleq w \otimes v - v \otimes w.$$

To define the higher order terms H_n , we first recall the definition of a derivation. A linear operator $\mathfrak{D} : T((V)) \rightarrow T((V))$ is called a *derivation* if

$$\mathfrak{D}(\xi \otimes \eta) = \mathfrak{D}(\xi) \otimes \eta + \xi \otimes \mathfrak{D}(\eta).$$

Now let $H_1(v, w)\partial_w$ denote the derivation induced by

$$(H_1(v, w)\partial_w)(v) \triangleq 0, \quad (H_1(v, w)\partial_w)(w) \triangleq H_1(v, w).$$

Then the functional $H_n(v, w)$ is defined by

$$H_n(v, w) \triangleq \frac{1}{n!} (H_1(v, w)\partial_w)^n(w). \quad (3.12)$$

It is clear that the partial degree of v in $H_n(v, w)$ is n . The series H_n will be referred to as the *n-th Hausdorff series*.

To summarise, one obtains from the relations (3.8) and (3.9) that

$$L(\gamma) = B(\tilde{L}(\gamma), e_1) \in \mathcal{L}((\mathbb{R}^2)). \quad (3.13)$$

Note that here $B(\tilde{L}(\gamma), e_1)$ is defined through the substitution $(v, w) \leftrightarrow (\tilde{L}(\gamma), e_1)$ and applying the tensor product structure over $T((\mathbb{R}^2))$. We remark that $B(\tilde{L}(\gamma), e_1)$ is a well-defined Lie series in $\mathcal{L}((\mathbb{R}^2))$. This is clear from the observation that the series

$$\sum_{n=1}^{\infty} H_n(\tilde{L}(\gamma), e_1)$$

is locally finite, i.e. its projection onto the truncated tensor algebra only involves finitely many Lie polynomials arising from the above series.

3.2.2 An adjoint representation of signature

We now present a basic formula for the signature $S(\gamma)$ on which our analysis is largely based.

Lemma 3.9. *Let $S(\gamma)$ be the signature of γ . Then one has*

$$S(\gamma) = \left(\sum_{n=0}^{\infty} \int_{0 < t_1 < \dots < t_n < 1} e^{x_{t_1} \text{ad}_{e_1}}(e_2) \otimes \dots \otimes e^{x_{t_n} \text{ad}_{e_1}}(e_2) dy_{t_1} \dots dy_{t_n} \right) \otimes e^{e_1}, \quad (3.14)$$

where

$$e^{\text{ad}_{\xi}}(\eta) \triangleq \sum_{m=0}^{\infty} \frac{(\text{ad}_{\xi})^m}{m!}(\eta).$$

In particular,

$$\tilde{L}(\gamma) \triangleq \log \left(\sum_{n=0}^{\infty} \int_{0 < t_1 < \dots < t_n < 1} e^{x_{t_1} \text{ad}_{e_1}}(e_2) \otimes \dots \otimes e^{x_{t_n} \text{ad}_{e_1}}(e_2) dy_{t_1} \dots dy_{t_n} \right)$$

is a Lie series.

Proof. The signature path $t \mapsto S(\gamma)_{0,t}$ satisfies the differential equation

$$dS(\gamma)_{0,t} = S(\gamma)_{0,t} \otimes (e_1 dx_t + e_2 dy_t).$$

Let $C_t \triangleq S(\gamma)_{0,t} \otimes e^{-x_t e_1}$. Then C_t satisfies the differential equation

$$dC_t = C_t \otimes e^{x_t e_1} \otimes e_2 \otimes e^{-x_t e_1} dy_t.$$

The result follows from the relation that

$$e^{x_t e_1} \otimes e_2 \otimes e^{-x_t e_1} = e^{x_t \text{ad}_{e_1}}(e_2).$$

Since $\tilde{L}(\gamma)$ is the logarithmic signature of $\gamma \sqcup \overleftarrow{\mathfrak{e}_1}$, it is clear that $\tilde{L}(\gamma)$ is a Lie series. \square

3.2.3 Basic notions on Cartan's path development

Based on the representation (3.14), a fundamental idea in our analysis is to perform Cartan's path development in such a way that e_2 is mapped into certain (combinations of) eigenspaces of e_1 . This will enable one to evaluate $e^{x_t \text{ad}_{e_1}}(e_2)$ in an explicit way (within the new Lie algebra). Before introducing such constructions, we first recall basic notions on Cartan's path development.

Definition 3.10. A (Cartan's) path development over \mathbb{R}^2 is a pair (\mathfrak{g}, F) where \mathfrak{g} is a finite dimensional complex Lie algebra and $F : \mathbb{R}^2 \rightarrow \mathfrak{g}$ is a real linear map.

We always assume that \mathfrak{g} is embedded in the matrix algebra $\mathcal{M} = \text{Mat}(N, \mathbb{C})$ for some N . The space \mathbb{R}^2 is complexified into \mathbb{C}^2 in the canonical way and so is F . All algebraic relations are understood over the complex field. Working over \mathbb{C} rather than \mathbb{R} is an important point in our analysis.

The linear map F admits a unique extension to an algebra homomorphism

$$\hat{F} : T(\mathbb{C}^2) \triangleq \bigoplus_{n=0}^{\infty} (\mathbb{C}^2)^{\otimes n} \rightarrow \mathcal{M}. \quad (3.15)$$

The restriction of \hat{F} on Lie polynomials over \mathbb{C}^2 defines a Lie homomorphism, which is the unique extension of F to a Lie homomorphism from the free Lie algebra over \mathbb{C}^2 to \mathfrak{g} . It should be noted that \hat{F} is not always well-defined on $T((\mathbb{C}^2))$ (formal tensor series) or on $\mathcal{L}((\mathbb{C}^2))$ (formal Lie series). The extension of \hat{F} to these spaces requires analytic consideration. In what follows, we always fix a matrix norm on \mathcal{M} .

Definition 3.11. Let $\xi \in T((\mathbb{C}^2))$ be a given tensor series. We say that $\hat{F}(\xi)$ *exists* and write $\xi \in \mathcal{D}(\hat{F})$, if $\lim_{N \rightarrow \infty} \hat{F}(\pi^{(N)}\xi)$ exists in \mathcal{M} .

Let γ be a weakly geometric rough path over \mathbb{C}^2 . Since $S(\gamma)$ has infinite R.O.C., it is obvious that $S(\gamma) \in \mathcal{D}(\hat{F})$. In addition, if $L(\gamma) \triangleq \log S(\gamma)$ has infinite R.O.C., one also has $L(\gamma) \in \mathcal{D}(\hat{F})$. In general, $\hat{F}(L(\gamma))$ always exists when F is “sufficiently small”.

Corollary 3.12. *Let γ be a weakly geometric rough path over \mathbb{C}^2 . Then there exists $\delta > 0$ such that for any path development (\mathfrak{g}, F) with $\|F\|_{\mathbb{C}^2 \rightarrow \mathcal{M}} < \delta$, one has $L(\gamma) \in \mathcal{D}(\hat{F})$. In particular, under the previous setting $\tilde{L}(\gamma) \in \mathcal{D}(\hat{F})$ provided that the operator norm of F is small enough (recall that $\tilde{L}(\gamma)$ is the logarithmic signature of $\gamma \sqcup \overleftarrow{e}_1$).*

Proof. According to [10], it is enough to choose $\delta = (2\|\gamma\|_{p\text{-var}})^{-1}$ where p is the roughness of γ . \square

Remark 3.13. *The use of path development in the literature, which also justifies its name, has a geometric nature. Given a pair (\mathfrak{g}, F) , there is a canonical way of “lifting” any Euclidean path γ in \mathbb{R}^2 to a corresponding path Γ taking values in a Lie group whose Lie algebra is \mathfrak{g} . It turns out that one can study geometric and signature-related properties of the original path γ from the lifted path Γ . The viewpoint we shall take in this work is however quite different. We work with path developments at the Lie algebra level instead of considering the actual lifted path in the group. Our technique thus has an algebraic flavour in contrast to the geometric viewpoints taken in earlier works.*

3.2.4 A two-dimensional development

To prove Theorem 3.1, we are going to choose a specific path development. Namely, we take \mathfrak{g} to be the two-dimensional complex Lie algebra generated by (two symbols) A, D under the Lie structure $[A, D] = D$. For each $\lambda \in \mathbb{C}$, we define the development map $F_\lambda : \mathbb{C}^2 \rightarrow \mathfrak{g}$ to be the linear map induced by

$$F_\lambda(e_1) \triangleq \lambda A, \quad F_\lambda(e_2) = D.$$

A formal calculation

We first show at a formal level how the choice of such a development enables one to prove Theorem 3.1. Recall that \hat{F}_λ is the induced algebra homomorphism on tensors and also Lie homomorphism on Lie elements. By applying \hat{F}_λ on both sides of (3.13) one finds that

$$e^{\hat{F}_\lambda(L(\gamma))} = e^{B(\hat{F}_\lambda(\tilde{L}(\gamma)), \lambda A)} \xrightarrow{\text{formally}} \hat{F}_\lambda(L(\gamma)) = B(\hat{F}_\lambda(\tilde{L}(\gamma)), \lambda A). \quad (3.16)$$

According to Lemma 3.9 and the construction of F_λ , one has

$$\begin{aligned} e^{\hat{F}_\lambda(\tilde{L}(\gamma))} &= \hat{F}_\lambda(e^{\tilde{L}(\gamma)}) = \sum_{n=0}^{\infty} \int_{0 < t_1 < \dots < t_n < 1} e^{x_{t_1} \text{ad}_{\lambda A}(D)} \dots e^{x_{t_n} \text{ad}_{\lambda A}(D)} dy_{t_1} \dots dy_{t_n} \\ &= \sum_{n=0}^{\infty} \int_{0 < t_1 < \dots < t_n < 1} e^{\lambda(x_{t_1} + \dots + x_{t_n})} D^n dy_{t_1} \dots dy_{t_n} \\ &= \sum_{n=0}^{\infty} \frac{D^n}{n!} \left(\int_0^1 e^{\lambda x_t} dy_t \right)^n = \exp \left(D \int_0^1 e^{\lambda x_t} dy_t \right). \end{aligned}$$

This formally implies that

$$\hat{F}_\lambda(\tilde{L}(\gamma)) = D \int_0^1 e^{\lambda x_t} dy_t. \quad (3.17)$$

By substituting (3.17) into (3.16), one finds that

$$\hat{F}_\lambda(L(\gamma)) = B \left(D \int_0^1 e^{\lambda x_t} dy_t, \lambda A \right).$$

Next, recall from (3.10) that $B = \sum_{n \geq 1} H_n$. The crucial observation is that

$$H_n \left(D \int_0^1 e^{\lambda x_t} dy_t, \lambda A \right) = 0 \quad \forall n \geq 2$$

and

$$H_1 \left(D \int_0^1 e^{\lambda x_t} dy_t, \lambda A \right) = \left(\sum_{m=0}^{\infty} \frac{B_m}{m!} \lambda^m \right) \cdot \left(\int_0^1 e^{\lambda x_t} dy_t \right) D,$$

both being consequences of the relation $[A, D] = D$. As a result, one arrives at the following relation:

$$\hat{F}_\lambda(L(\gamma)) = \left(\sum_{m=0}^{\infty} \frac{B_m}{m!} \lambda^m \right) \cdot \left(\int_0^1 e^{\lambda x_t} dy_t \right) D. \quad (3.18)$$

Now suppose that $L(\gamma)$ has infinite R.O.C. Then the left hand side of (3.18) defines a \mathfrak{g} -valued entire function (as a function of $\lambda \in \mathbb{C}$). On the other hand, the power series $\sum_{m=0}^{\infty} \frac{B_m}{m!} \lambda^m$ defines the meromorphic function $\phi(\lambda) = \frac{\lambda}{e^\lambda - 1}$ on \mathbb{C} that has isolated singularities at $\lambda = 2k\pi i$ ($k \in \mathbb{Z} \setminus \{0\}$). This forces $\int_0^1 e^{\lambda x_t} dy_t = 0$ at these singularities $\lambda = 2k\pi i$. The conclusion of Theorem 3.1 thus follows.

Proof of Theorem 3.1

The above argument is only formal, since one cannot directly pass \hat{F} into the BCH function B without analytical considerations, and to justify this one needs to apply

truncations and then pass to the limit. This is only a technical matter which we now make precise.

Let \mathfrak{g} be the Lie algebra of 2×2 upper triangular matrices over \mathbb{C} with zero trace. Using the previous notation, one can take

$$A = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

so that $\mathfrak{g} = \text{Span}\{A, D\}$ and $[A, D] = D$. We consider the path development defined by

$$F_{\lambda, \mu}(e_1) \triangleq \lambda A, \quad F_{\lambda, \mu}(e_2) \triangleq \mu D \quad (\lambda, \mu \in \mathbb{C}).$$

According to Corollary 3.12, there exists $\delta > 0$ such that

$$\tilde{L}(\gamma) \in \mathcal{D}(\hat{F}_{\lambda, \mu}) \quad \forall \lambda, \mu : |\lambda| \vee |\mu| < \delta.$$

In other words, the limit

$$\hat{F}_{\lambda, \mu}(\tilde{L}(\gamma)) = \lim_{N \rightarrow \infty} \hat{F}_{\lambda, \mu}(\pi^{(N)} \tilde{L}(\gamma)) \quad (3.19)$$

exists for all such λ, μ .

On the other hand, recall from the normalisation (3.1) that the x -increment of $\tilde{\gamma} \triangleq \gamma \sqcup \overleftarrow{e_1}$ is zero. In particular, $\tilde{L}(\gamma)$ does not contain the e_1 -component. After applying the development $\hat{F}_{\lambda, \mu}$, one can thus write

$$\hat{F}_{\lambda, \mu}(\pi^{(N)} \tilde{L}(\gamma)) = C_N(\lambda, \mu) D \quad (3.20)$$

for some $C_N(\lambda, \mu) \in \mathbb{C}$. The function $(\lambda, \mu) \mapsto C_N(\lambda, \mu)$ is clearly entire on \mathbb{C}^2 for every fixed N . The relation (3.19) then implies that $C_N(\lambda, \mu)$ is convergent as $N \rightarrow \infty$ for small λ, μ , whose limit with no surprise should coincide with (3.17) if one were able to take $\mu = 1$.

Lemma 3.14. *One has*

$$\hat{F}_{\lambda, \mu}(\tilde{L}(\gamma)) = \lim_{N \rightarrow \infty} C_N(\lambda, \mu) D = \mu \left(\int_0^1 e^{\lambda x_t} dy_t \right) D, \quad |\lambda| \vee |\mu| < \delta.$$

Proof. By applying $\hat{F}_{\lambda, \mu}$ to the signature of $\tilde{\gamma}$ and using Lemma 3.9, one obtains that

$$\hat{F}_{\lambda, \mu}(e^{\tilde{L}(\gamma)}) = \exp \left(\mu \left(\int_0^1 e^{\lambda x_t} dy_t \right) D \right), \quad \forall \lambda, \mu \in \mathbb{C}.$$

Note that there is no convergence issue at the signature level due to its factorial decay. When λ, μ is small, one also has

$$\hat{F}_{\lambda, \mu}(e^{\tilde{L}(\gamma)}) = e^{\hat{F}_{\lambda, \mu}(\tilde{L}(\gamma))}.$$

For each fixed λ with $|\lambda| < \delta$, both of the functions

$$G(\mu) \triangleq \hat{F}_{\lambda,\mu}(\tilde{L}(\gamma)), \quad H(\mu) \triangleq \mu \left(\int_0^1 e^{\lambda x_t} dy_t \right) D$$

are analytic in μ for $|\mu| < \delta$. In addition, the relation (3.20) shows that $G(\mu) \in \text{Span}\{D\}$ and thus $G(\mu)$ is commutative for different μ 's. Since G and H coincide when $\mu = 0$, a standard power series argument shows that

$$e^{G(\mu)} = e^{H(\mu)} \text{ for } |\mu| < \delta \implies G(\mu) = H(\mu) \text{ for } |\mu| < \delta.$$

The result thus follows. \square

On the other hand, by using Lemma 3.14 one can also compute $C_N(\lambda, \mu)$ explicitly.

Lemma 3.15. *For any $\lambda, \mu \in \mathbb{C}$, one has*

$$C_N(\lambda, \mu) = \mu \sum_{j=0}^{N-1} \left(\int_0^1 \frac{x_t^j}{j!} dy_t \right) \lambda^j. \quad (3.21)$$

Proof. It suffices to establish the relation for small λ, μ , since both sides of the identity define entire functions in $(\lambda, \mu) \in \mathbb{C}^2$. To this end, note that the logarithmic signature $\tilde{L}(\gamma)$ admits a unique decomposition

$$\tilde{L}(\gamma) = \sum_{j=0}^{\infty} c_j \text{ad}_{e_1}^j(e_2) + L',$$

where L' consists of those tensor components with at least two e_2 's. After applying $\hat{F}_{\lambda,\mu}$, one finds that

$$C_N(\lambda, \mu) = \mu \sum_{j=0}^{N-1} c_j \lambda^j$$

for all small λ, μ . According to Lemma 3.14, the λ -power series of $\hat{F}_{\lambda,\mu}(\tilde{L}(\gamma))$ (for fixed μ) is given by

$$\mu D \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \int_0^1 x_t^j dy_t.$$

By comparing coefficients, one finds that

$$c_j = \frac{1}{j!} \int_0^1 x_t^j dy_t$$

and the result thus follows. \square

Proof of Theorem 3.1. Suppose that $L(\gamma)$ has infinite R.O.C. Then $L(\gamma) \in \mathcal{D}(\hat{F}_{\lambda,\mu})$ for all $\lambda, \mu \in \mathbb{C}$. It follows from the relation (3.13) that

$$\lim_{N \rightarrow \infty} \hat{F}_{\lambda,\mu}(\pi^{(N)} B(\tilde{L}(\gamma), e_1)) = \hat{F}_{\lambda,\mu}(L(\gamma)) \in \mathfrak{g} \subseteq \mathcal{M}.$$

According to the definition (3.10) of B , one has

$$\pi^{(N)} B(\tilde{L}(\gamma), e_1) = \pi^{(N)} \sum_{n=1}^{\infty} H_n(\tilde{L}(\gamma), e_1) = \sum_{n=1}^N \pi^{(N)} H_n(\pi^{(N)} \tilde{L}(\gamma), e_1).$$

Note that the right hand side only contains finitely many terms from the expansion of H_n . Due to the relations (3.20) and $[A, D] = D$, it is easily seen that

$$H_n(\hat{F}_{\lambda,\mu}(\pi^{(N)} \tilde{L}(\gamma)), \lambda A) = 0 \quad \forall n \geq 2.$$

As a result, according to the formulas (3.11) and (3.21) one finds that

$$\hat{F}_{\lambda,\mu}(\pi^{(N)} B(\tilde{L}(\gamma), e_1)) = D \sum_{m=0}^{N-1} \frac{B_m}{m!} \lambda^m C_{N-m}(\lambda, \mu) = \mu D \sum_{m=0}^{N-1} \frac{B_m}{m!} \lambda^m \sum_{j=0}^{N-m-1} \int_0^1 \frac{x_t^j}{j!} dy_t.$$

By taking $N \rightarrow \infty$, one arrives at the identity

$$\hat{F}_{\lambda,\mu}(L(\gamma)) = \mu D \cdot \lim_{N \rightarrow \infty} \sum_{m=0}^{N-1} \frac{B_m}{m!} \lambda^m \sum_{j=0}^{N-m-1} \frac{\lambda^j}{j!} \int_0^1 x_t^j dy_t.$$

Note that the above limit exists for all $\lambda, \mu \in \mathbb{C}$.

The next observation is that the expression

$$\sum_{m=0}^{N-1} \frac{B_m}{m!} \lambda^m \sum_{j=0}^{N-m-1} \frac{\lambda^j}{j!} \int_0^1 x_t^j dy_t = \sum_{r=0}^{N-1} \left(\sum_{l=0}^r \frac{B_l}{l!(m-l)!} \int_0^1 x_t^{r-l} dy_t \right) \lambda^r$$

is precisely the partial sum of the power series defined by expanding the product

$$\left(\sum_{m=0}^{\infty} \frac{B_m}{m!} \lambda^m \right) \left(\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \int_0^1 x_t^j dy_t \right).$$

For fixed $\mu \in \mathbb{C}$ and $|\lambda| < 2\pi$, both of the above power series are absolutely convergent, yielding that

$$\hat{F}_{\lambda,\mu}(L(\lambda)) = \mu D \cdot \left(\sum_{m=0}^{\infty} \frac{B_m}{m!} \lambda^m \right) \left(\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \int_0^1 x_t^j dy_t \right) = \frac{\lambda \mu}{e^\lambda - 1} D \times \int_0^1 e^{\lambda x_t} dy_t.$$

Therefore,

$$(e^\lambda - 1) \hat{F}_{\lambda,\mu}(L(\gamma)) = \mu D \cdot \lambda \int_0^1 e^{\lambda x_t} dy_t$$

for fixed $\mu \neq 0$ and all λ with $|\lambda| < 2\pi$. Now observe that both sides are entire functions in λ (since $L(\gamma)$ has infinite R.O.C.). As a result, the same identity holds for all $\lambda \in \mathbb{C}$. By taking $\lambda = 2k\pi i$ with $k \in \mathbb{Z} \setminus \{0\}$, one concludes that

$$\int_0^1 e^{2k\pi i \cdot x_t} dy_t = 0,$$

hence giving the desired integral property (3.2). □

4 A weak version of the LS conjecture

In this section, we prove a weak version of the LS conjecture: having infinite R.O.C. for the logarithmic signature *over all sub-intervals of time* implies that the underlying path must live on a straight line.

Theorem 4.1. *Let $\gamma : [0, 1] \rightarrow \mathbb{R}^d$ be a continuous path with bounded variation. Suppose that $\log S(\gamma)_{s,t}$ has infinite R.O.C. for all $[s, t] \subseteq [0, 1]$. Then $\gamma_t = \gamma_0 + f(t) \cdot \vec{v}$ for some real-valued function $f : [0, 1] \rightarrow \mathbb{R}$ and some fixed vector $\vec{v} \in \mathbb{R}^d$.*

Our main strategy of proving Theorem 4.1 can be summarised as follows. It is based on non-trivial applications of Theorem 3.4 together with winding number considerations.

1. We may assume $d = 2$ and parametrise the path γ at unit speed. Let s be an arbitrary point at which γ is differentiable and choose a time $T > s$ that is sufficiently close to s . We normalise (i.e. rescale and rotate) the path $\gamma|_{[s,T]}$ to obtain a new path β which starts at $(0,0)$ and ends at $(1,0)$. Since $|\gamma'_s| = 1$ and $T \approx s$, in the rescaled picture the new path β stays inside a cone region $\{(x, y) : x \in [0, 1], |y| < cx\}$ for some small number c .
2. For each $(x, y) \in \mathbb{R}^2$, let $w(x, y)$ denote the winding number of the path $\tilde{\beta} \triangleq \beta \sqcup \overleftarrow{e_1}$ around the point (x, y) . Since the path β satisfies the generalised integral condition (3.3), one can show that the function $(x, y) \mapsto w(x, y)$ depends only on y . As a result, $w(x, y)$ has to vanish identically (because for any (x, y) inside the cone there is some (x', y) at the same level which is outside the cone and the winding number of $\tilde{\beta}$ around that point is obviously zero). This further implies that β is the line segment e_1 *in the weak sense* that the line integrals along β and e_1 are identical for any smooth one-form. Transferring back to the original path γ , one concludes that $\gamma|_{[s,T]}$ is a line segment in the above weak sense. Call this line L .
3. The path $\gamma|_{[T,1]}$ has to be entirely contained in L and we prove this by contradiction. If this is not true, there exists a time $t > T$ such that $\gamma_t \notin L$. By

normalising the path $\gamma|_{[s,t]}$ to a new path ζ which starts from the origin and ends at $(1,0)$, the line L is transformed to another line L' which is not horizontal. Draw a tiny ball B_1 centered at the mid-point of $L'|_{[s,T]}$ and make a copy B_2 at the same vertical so that $B_2 \cap L'|_{[s,T]} = \emptyset$. This is possible since L' is not horizontal.

4. Choose an arbitrary one-form Ψ supported on B_1 whose line integral along $L'|_{[s,T]}$ is non-zero. Modify its values on B_2 to obtain a new one-form Φ which satisfies the two conditions of Theorem 3.4. The conclusion of Point 2 and the construction of Φ implies that

$$\int_s^t \Phi(d\zeta_u) = \int_s^T \Psi(dL'_u) \neq 0.$$

This contradicts the conclusion of Theorem 3.4 since the logarithmic signature of ζ has infinite R.O.C. Therefore, $\gamma|_{[T,1]}$ has to be contained in the line L . Since the differentiable point s is arbitrary and $T \approx s$, this implies that the entire path γ lives on a single well-defined straight line.

The rest of this section is devoted to the proof of Theorem 4.1.

4.1 An application of the integral condition to winding number

In this subsection, we derive a simple application of the generalised line integral condition (3.3) to the winding number. Such a property will play a key role in our proof of Theorem 4.1.

Let $\gamma : [0, 1] \rightarrow \mathbb{C}$ be a continuous BV path such that $\gamma_0 = \gamma_1$. Given $(x, y) \notin \gamma[0, 1]$, the *winding number* of γ around (x, y) is defined by

$$\eta(\gamma, (x, y)) \triangleq \frac{1}{2\pi i} \int \frac{1}{\gamma_t - (x + yi)} d\gamma_t. \quad (4.1)$$

The winding number of $\gamma = (\gamma^1, \gamma^2) : [0, 1] \rightarrow \mathbb{R}^2$ is defined as the winding number of the path $t \mapsto \gamma_t^1 + i\gamma_t^2$. We will use the following properties of the winding number.

1. If there is a simply-connected set K such that $\gamma[0, 1] \subseteq K$, then $\eta(\gamma, (x, y)) = 0$ for all $(x, y) \notin K$. This is a consequence of Cauchy's theorem.
2. Let $R : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a rotation and $\lambda \in \mathbb{R} \setminus \{0\}$. Then

$$\eta(\lambda R(\gamma), (x, y)) = \eta(\gamma, \lambda^{-1} R^{-1}(x, y)).$$

This follows from the definition (4.1) of winding number.

3. (Green's theorem for self-intersecting paths) For any smooth functions $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$, one has

$$\int_{\mathbb{R}^2} (\partial_x f(x, y) + \partial_y g(x, y)) \eta(\gamma, (x, y)) dx dy = \int f dy_s - g dx_s.$$

In addition, the function

$$(x, y) \rightarrow \eta(\gamma, (x, y)).$$

is square integrable with respect to the Lebesgue measure on \mathbb{R}^2 . This fact can be found in [5, Theorem 15] and references therein.

Let $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ be a continuous BV path such that $\gamma_0 = (0, 0)$ and $\gamma_1 = (1, 0)$. Recall that $\tilde{\gamma} \triangleq \gamma \sqcup \overleftarrow{e_1}$ and we parametrise it on $[0, 1]$. More precisely, it is defined by

$$\tilde{\gamma}_t \triangleq \begin{cases} \gamma_{2t}, & t \in [0, 1/2] \\ \gamma_1 - (\gamma_1 - \gamma_0)(2t - 1), & t \in [1/2, 1]. \end{cases}$$

Suppose that the generalised line integral condition (3.3) holds for γ . Then the same condition also holds for $\tilde{\gamma}$. This is due to the facts that (3.3) holds for both γ and $\beta \triangleq [t \mapsto (1 - t, 0)]$ and that

$$\int \Phi(d\tilde{\gamma}) = \int \Phi(d\gamma) + \int \Phi(d\beta).$$

We now state the main application of (3.3) to the winding number.

Lemma 4.2. *Let γ be a continuous BV path such that $\gamma_0 = (0, 0)$, $\gamma_1 = (1, 0)$ and*

$$\gamma[0, 1] \subseteq [0, 1] \times \mathbb{R},$$

with $\gamma_t \in \{1\} \times \mathbb{R}$ if and only if $t = 1$. Let η be the winding number of $\tilde{\gamma}$ around the point (x, y) . Suppose that γ satisfies the generalised line integral condition (3.3). Then for $(x, y) \in (0, 1) \times \mathbb{R} \setminus \tilde{\gamma}[0, 1]$, one has

$$\eta(\tilde{\gamma}, (x, y)) = \int_0^1 \eta(\tilde{\gamma}, (v, y)) dv.$$

Proof. Let g be a smooth function satisfying $g(x + 1, y) = g(x, y)$ and define $f(x, y) \triangleq g(x, y) - \int_0^1 g(t, y) dt$. Write $\gamma_t = (x_t, y_t)$. According to the integral condition (3.3), one has

$$\int_0^1 g(x_s, y_s) dx_s - \int_0^1 \left(\int_0^1 g(t, y_s) dt \right) dx_s = 0.$$

It follows from Green's theorem for self-intersecting paths that

$$\int_{\mathbb{R}^2} \frac{\partial g}{\partial y}(x, y) \eta(\tilde{\gamma}, (x, y)) dx dy = \int_{\mathbb{R}^2} \left(\int_0^1 \frac{\partial g}{\partial y}(t, y) dt \right) \eta(\tilde{\gamma}, (x, y)) dx dy. \quad (4.2)$$

Since $\tilde{\gamma}[0, 1] \subseteq [0, 1] \times \mathbb{R}$ and $\eta(\tilde{\gamma}, (x, y)) = 0$ for all $(x, y) \notin [0, 1] \times \mathbb{R}$, the relation (4.2) can be rewritten as

$$\int_{\mathbb{R}} \int_0^1 \frac{\partial g}{\partial y}(x, y) \eta(\tilde{\gamma}, (x, y)) dx dy = \int_{\mathbb{R}} \int_0^1 \frac{\partial g}{\partial y}(t, y) \left(\int_0^1 \eta(\tilde{\gamma}, (x, y)) dx \right) dt dy,$$

or equivalently,

$$\int_{\mathbb{R}} \int_0^1 \frac{\partial g}{\partial y}(x, y) [\eta(\tilde{\gamma}, (x, y)) - \int_0^1 \eta(\tilde{\gamma}, (t, y)) dt] dx dy = 0. \quad (4.3)$$

Let h be a smooth function such that $h(x+1, y) = h(x, y)$. By applying the relation (4.3) to the function $g(x, y) \triangleq \int_0^y h(x, t) dt$, one finds that

$$\int_{\mathbb{R}} \int_0^1 h(x, y) [\eta(\tilde{\gamma}, (x, y)) - \int_0^1 \eta(\tilde{\gamma}, (t, y)) dt] dx dy = 0.$$

Now fix any $(x_0, y_0) \in (0, 1) \times \mathbb{R} \setminus \tilde{\gamma}[0, 1]$ and let $(h_\varepsilon)_{\varepsilon > 0}$ be a standard mollifier. Define

$$\tilde{h}_\varepsilon(x, y) \triangleq h_\varepsilon((x, y) - (x_0, y_0))$$

and modify it to be 1-periodic in x . Then one has

$$\int_{\mathbb{R}} \int_0^1 \tilde{h}_\varepsilon((x, y) - (x_0, y_0)) \left(\eta(\tilde{\gamma}, (x, y)) - \int_0^1 \eta(\tilde{\gamma}, (t, y)) dt \right) dx dy = 0.$$

By taking $\varepsilon \rightarrow 0$, it follows that

$$\eta(\tilde{\gamma}, (x_0, y_0)) - \int_0^1 \eta(\tilde{\gamma}, (t, y_0)) dt = 0.$$

The result follows since (x_0, y_0) is arbitrary. \square

4.2 A zero winding lemma

Due to the conditions in Theorem 3.4, it is necessary to rescale and rotate a path γ so that it satisfies $\gamma_0 = (0, 0)$ and $\gamma_1 = (1, 0)$. We first introduce such a normalising operation.

Definition 4.3. Let $\gamma : [0, T] \rightarrow \mathbb{R}^2$ be a BV path and let $s < t$ be two fixed times such that $\gamma_s \neq \gamma_t$. We define the associated *normalisation operator* $\mathcal{A}_{s,t}$ by

$$\mathcal{A}_{s,t}(x) \triangleq \frac{1}{|\gamma_t - \gamma_s|} R_{s,t}(x - \gamma_s), \quad x \in \mathbb{R}^2.$$

Here $R_{s,t} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ denotes the rotation that maps $\gamma_t - \gamma_s$ to $(|\gamma_t - \gamma_s|, 0)$. We use $\mathcal{A}_{s,t}\gamma$ to denote the path $[s, t] \ni u \mapsto \mathcal{A}_{s,t}(\gamma_u)$.

In this subsection, we will prove the following key result. As usual, we use $\tilde{\mathcal{A}}_{s,t}\gamma$ to denote the path $\mathcal{A}_{s,t}\gamma \sqcup \overleftarrow{e_1}$ (we also assume that it is parametrised on $[s, t]$). Recall that $\eta(\tilde{\mathcal{A}}_{s,t}\gamma, (x, y))$ is the winding number of $\tilde{\mathcal{A}}_{s,t}\gamma$ around the point (x, y) .

Proposition 4.4. *Let $\gamma : [0, K] \rightarrow \mathbb{R}^2$ be a continuous BV path which is parametrised at unit speed. Suppose that $\log S(\gamma)_{u,v}$ has infinite R.O.C. on all time intervals $[u, v] \subseteq [0, K]$. Let s be a given fixed time at which γ is differentiable and let $\varepsilon \in (0, 1/6]$. Let $T \in (s, K]$ be chosen such that*

$$\begin{aligned} |\gamma_t - \gamma_s - \gamma'_s(t-s)| &\leq \varepsilon(t-s), \\ \left| \left| \frac{\gamma_t - \gamma_s}{t-s} \right| - 1 \right| &\leq \varepsilon \end{aligned} \tag{4.4}$$

for all $t \in (s, T]$ and that $|\gamma_t - \gamma_s| < |\gamma_T - \gamma_s|$ for all $s \in (t, T)$. Then one has

$$\eta(\tilde{\mathcal{A}}_{s,T}\gamma, (x, y)) = 0$$

for all $(x, y) \in (0, 1) \times \mathbb{R} \setminus (\tilde{\mathcal{A}}_{s,T}\gamma)[s, T]$.

The proof of Proposition 4.4 relies on the following key observation which asserts that the path $\mathcal{A}_{s,T}\gamma$ lies inside a cone region.

Lemma 4.5. *Let $\gamma, s, T, \varepsilon$ be as in Proposition 4.4. Then*

$$(\mathcal{A}_{s,T}\gamma)[s, T] \subseteq \left\{ (x, y) : |y| < \frac{2\varepsilon}{1-3\varepsilon}x \right\}.$$

Proof. By using (4.4), one has

$$\left| \frac{1}{|\gamma_T - \gamma_s|} R_{s,T}(\gamma_t - \gamma_s) - \frac{R_{s,T}\gamma'_s}{|\gamma_T - \gamma_s|}(t-s) \right| \leq \varepsilon \frac{(t-s)}{|\gamma_T - \gamma_s|}$$

and hence

$$|\mathcal{A}_{s,T}(\gamma_t) - \mathcal{A}_{s,T}(\gamma'_s)(t-s)| \leq \varepsilon \frac{(t-s)}{|\gamma_T - \gamma_s|}.$$

In particular, by taking $t = T$ and dividing the equation by $T - s$, one obtains that

$$\left| \frac{1}{T-s}(1, 0) - \mathcal{A}_{s,T}(\gamma'_s) \right| \leq \varepsilon \frac{1}{|\gamma_T - \gamma_s|}.$$

Therefore,

$$\begin{aligned} &\left| (\mathcal{A}_{s,T}\gamma)_t - (t-s) \left(\frac{1}{T-s}, 0 \right) \right| \\ &\leq |(\mathcal{A}_{s,T}\gamma)_t - \mathcal{A}_{s,T}(\gamma'_s)(t-s)| + \left| \left(\frac{1}{T-s}(1, 0) - (\mathcal{A}_{s,T}\gamma)'_s \right) (t-s) \right| \\ &\leq \frac{2\varepsilon(t-s)}{(1-\varepsilon)(T-s)}. \end{aligned}$$

Let us use x^1 to denote the first coordinate of $x \in \mathbb{R}^2$ and similarly for x^2 . Then one has

$$\begin{aligned} \left| (\mathcal{A}_{s,T}\gamma)_t^1 - \frac{t-s}{T-s} \right| &\leq \frac{2\varepsilon(t-s)}{(1-\varepsilon)(T-s)}, \\ \left| (\mathcal{A}_{s,T}\gamma)_t^2 \right| &\leq \frac{2\varepsilon(t-s)}{(1-\varepsilon)(T-s)}. \end{aligned}$$

Note that for $\varepsilon \in (0, 1/6]$, the first inequality implies that

$$t-s \leq \frac{(\mathcal{A}_{s,T}\gamma)_t^1}{1 - \frac{2\varepsilon}{1-\varepsilon}}(T-s)$$

and therefore

$$\left| (\mathcal{A}_{s,T}\gamma)_t^2 \right| \leq \frac{2\varepsilon}{1-3\varepsilon} (\mathcal{A}_{s,T}\gamma)_t^1.$$

This shows that $\mathcal{A}_{s,T}\gamma$ lies in the desired cone region. \square

Proof of Proposition 4.4. By Lemma 4.5, one has

$$(\tilde{\mathcal{A}}_{s,T}\gamma)[s, T] \subseteq \left\{ (x, y) : |y| < \frac{2\varepsilon}{1-3\varepsilon}x \right\} =: \mathcal{C}.$$

As a consequence of Property 1 for the winding number,

$$\eta(\tilde{\mathcal{A}}_{s,T}\gamma, (x, y)) = 0 \tag{4.5}$$

for all $(x, y) \notin \mathcal{C}$. Now suppose that

$$(x, y) \in \mathcal{C} \cap \left((0, 1) \times \mathbb{R} \setminus (\tilde{\mathcal{A}}_{s,T}\gamma)[s, T] \right).$$

There exists some $x' \in (0, x)$ such that $(x', y) \notin \mathcal{C}$. According to Lemma 4.2, one concludes that

$$\eta(\tilde{\mathcal{A}}_{s,T}\gamma, (x, y)) = \eta(\tilde{\mathcal{A}}_{s,T}\gamma, (x', y)) = 0.$$

The result of the proposition thus follows. \square

The following result is an immediate application of Proposition 4.4. It shows that infinite R.O.C. for the logarithmic signature implies that the path contains a line segment in a weak sense.

Lemma 4.6. *Let $\gamma : [0, K] \rightarrow \mathbb{R}^2$ be a continuous BV path with unit speed parametrisation. Let s be a point at which γ is differentiable. Let $T \in (s, K]$ be chosen such that*

$$\left| \gamma_t - \gamma_s - \gamma'_s(t-s) \right| \leq \frac{1}{6}(t-s), \quad \left| \left| \frac{\gamma_t - \gamma_s}{t-s} \right| - 1 \right| \leq \frac{1}{6}$$

for all $t \in (s, T]$ and

$$|\gamma_t - \gamma_s| < |\gamma_T - \gamma_s|$$

for all $t \in (s, T)$. Suppose that $\log S(\gamma)$ has infinite R.O.C. on the interval $[s, T]$. Then the following two statements hold true.

(i) One has

$$\eta(\tilde{\gamma}, (x, y)) = 0$$

for all $(x, y) \in (0, 1) \times \mathbb{R} \setminus \tilde{\gamma}[s, T]$. Here $\tilde{\gamma} \triangleq \gamma|_{[s, T]} \sqcup (\gamma_s - \gamma_T)$.

(ii) Let L be the line segment joining γ_s to γ_T (also parametrised on $[s, T]$). Then for any smooth one-form $\Phi(x, y) = f(x, y)dx + g(x, y)dy$, one has

$$\int_s^T \Phi(d\gamma_v) = \int_s^T \Phi(dL_v).$$

Proof. (i) This follows from Proposition 4.4, $\tilde{\mathcal{A}}_{s, T}\gamma = \mathcal{A}_{s, T}\tilde{\gamma}$ (up to reparametrisation), that $\mathcal{A}_{s, T}$ is a composition of rotation and scaling, as well as Property 2 for the winding number which describes how the winding number behaves under rotation and scaling.

(ii) According to Green's Theorem for self-intersecting paths,

$$\int \Phi(d\tilde{\gamma}_v) = \int_{\mathbb{R}^2} \left(\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} \right) \eta(\tilde{\gamma}, (x, y)) dx dy.$$

It follows from Part (i) that $\int \Phi(d\tilde{\gamma}_v) = 0$, or equivalently,

$$\int_s^T \Phi(d\gamma_v) = \int_s^T \Phi(dL_v).$$

This proves the desired claim. □

4.3 Proof of Theorem 4.1

In this subsection, we develop the proof of Theorem 4.1. We first state a technical lemma.

Lemma 4.7. *Let $\gamma : [\tau_1, \tau_2] \rightarrow \mathbb{R}^2$ be a continuous BV path. There exist $\delta_1, \delta_2, r \in (0, 1)$ such that the following properties hold true. Let $a < b < d$ be elements of $[\tau_1, \tau_2]$ such that $\gamma_a \neq \gamma_b$*

$$|\gamma_b - \gamma_c| < \delta_1 |\gamma_b - \gamma_a|.$$

for all $c \in [b, d]$.

(i) One has

$$|\mathcal{A}_{a, d}(\gamma_b) - \mathcal{A}_{a, d}(\gamma_c)| < \frac{\delta_1}{1 - \delta_1}, \quad |\mathcal{A}_{a, d}(\gamma_c) - (1, 0)| < \frac{2\delta_1}{1 - \delta_1} \quad (4.6)$$

for all $c \in [b, d]$;

(ii) One has

$$\mathcal{A}_{a,d}(m_{a,b}) \subseteq \left(\frac{1}{2} - \frac{\delta_1/2}{1-\delta_1}, \frac{1}{2} + \frac{\delta_1/2}{1-\delta_1} \right) \times \mathbb{R} \quad (4.7)$$

and

$$\overline{B(\mathcal{A}_{a,d}(m_{a,b}), r)} \cup \overline{B(\mathcal{A}_{a,d}(m_{a,b}) - (\delta_2, 0), r)} \subseteq (0, 1) \times \mathbb{R}, \quad (4.8)$$

where $m_{a,b} \triangleq (\gamma_a + \gamma_b)/2$ and $B(x, r)$ denotes the ball centered at x with radius r ;

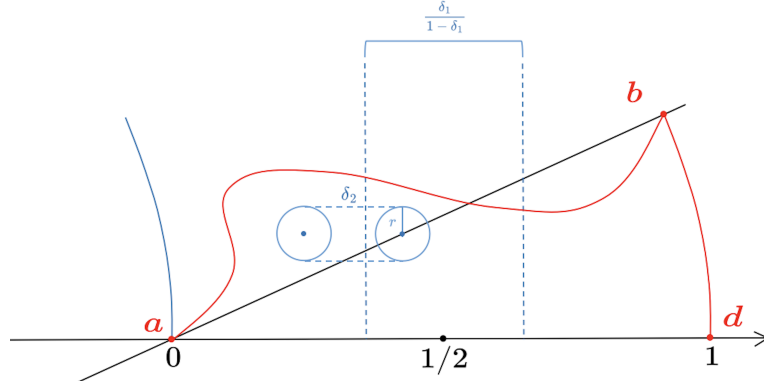
(iii) The four sets

$$\mathcal{A}_{a,d}\gamma|_{[b,d]}, \mathcal{A}_{a,d}\gamma|_{[b,d]} - (1, 0), \overline{B(\mathcal{A}_{a,d}(m_{a,b}), r)}, \overline{B(\mathcal{A}_{a,d}(m_{a,b}) - (\delta_2, 0), r)}$$

are disjoint.

(iv) Let $L^{a,b}$ denote the straight line joining $\mathcal{A}_{a,d}(\gamma_a)$ and $\mathcal{A}_{a,d}(\gamma_b)$. Suppose that $L^{a,b}$ is not horizontal. Then

$$\overline{B(\mathcal{A}_{a,d}(m_{a,b}) - (\delta_2, 0), r)} \cap L^{a,b} = \emptyset. \quad (4.9)$$



Proof. (i) Note that $\mathcal{A}_{a,d}\gamma$ is well-defined since

$$|\gamma_d - \gamma_a| \geq |\gamma_b - \gamma_a| - |\gamma_b - \gamma_d| \geq (1 - \delta_1)|\gamma_b - \gamma_a| > 0.$$

Now for all $c \in [b, d]$, one has

$$|\mathcal{A}_{a,d}(\gamma_b) - \mathcal{A}_{a,d}(\gamma_c)| = \frac{|\gamma_b - \gamma_c|}{|\gamma_d - \gamma_a|} < \frac{\delta_1 |\gamma_b - \gamma_a|}{|\gamma_b - \gamma_a| - |\gamma_d - \gamma_b|} \leq \frac{\delta_1}{1 - \delta_1}$$

and

$$\begin{aligned} |\mathcal{A}_{a,d}(\gamma_c) - (1, 0)| &= |\mathcal{A}_{a,d}(\gamma_c) - \mathcal{A}_{a,d}(\gamma_d)| \\ &\leq |\mathcal{A}_{a,d}(\gamma_b) - \mathcal{A}_{a,d}(\gamma_c)| + |\mathcal{A}_{a,d}(\gamma_b) - \mathcal{A}_{a,d}(\gamma_d)| < \frac{2\delta_1}{1 - \delta_1}. \end{aligned}$$

The desired estimate (4.6) thus follows.

(ii) The relation (4.7) follows since

$$\begin{aligned}\mathcal{A}_{a,d}(m_{a,b}) &= \frac{1}{2}(\mathcal{A}_{a,d}(\gamma_a) + \mathcal{A}_{a,d}(\gamma_b)) \\ &= \frac{1}{2}(\mathcal{A}_{a,d}(\gamma_b) - \mathcal{A}_{a,d}(\gamma_d)) + \frac{1}{2}\mathcal{A}_{a,d}(\gamma_d) \\ &= \frac{1}{2}(\mathcal{A}_{a,d}(\gamma_b) - \mathcal{A}_{a,d}(\gamma_d)) + \frac{1}{2}(1, 0).\end{aligned}$$

and one also knows from (4.6) that

$$|\mathcal{A}_{a,d}(\gamma_b) - \mathcal{A}_{a,d}(\gamma_d)| \leq \frac{\delta_1}{1 - \delta_1}.$$

As a result of (4.7), one can obviously make (4.8) valid by choosing δ_1, δ_2, r to be small enough.

(iii) According to (i),

$$\begin{aligned}\mathcal{A}_{a,d}\gamma|_{[b,d]} &\subseteq B\left((1, 0), \frac{2\delta_1}{1 - \delta_1}\right), \\ \mathcal{A}_{a,d}\gamma|_{[b,d]} - (1, 0) &\subseteq B\left((0, 0), \frac{2\delta_1}{1 - \delta_1}\right),\end{aligned}$$

and by (ii),

$$\mathcal{A}_{a,d}(m_{a,b}) \subseteq \left(\frac{1}{2} - \frac{\delta_1/2}{1 - \delta_1}, \frac{1}{2} + \frac{\delta_1/2}{1 - \delta_1}\right) \times \mathbb{R}.$$

The claim thus follows if one takes δ_1, δ_2, r to be sufficiently small so that the three sets

$$\left(\frac{1}{2} - \frac{\delta_1/2}{1 - \delta_1}, \frac{1}{2} + \frac{\delta_1/2}{1 - \delta_1}\right) \times \mathbb{R}, \quad B\left((1, 0), \frac{2\delta_1}{1 - \delta_1}\right), \quad B\left((0, 0), \frac{2\delta_1}{1 - \delta_1}\right)$$

are disjoint.

(iv) Since $L^{a,b}$ is not horizontal and $\mathcal{A}_{a,d}(m_{a,b}) \in L^{a,b}$, one knows that

$$\mathcal{A}_{a,d}(m_{a,b}) - (\delta_2, 0) \notin L^{a,b}.$$

By further reducing r if necessary, one can ensure that the relation (4.9) holds. \square

Next, by using Lemma 4.7 we prove a key lemma that is needed for the later proof of Theorem 4.1.

Lemma 4.8. *Let $\gamma : [0, K] \rightarrow \mathbb{R}^2$ be a continuous BV path with unit speed parametrisation. Suppose that $\log S(\gamma)$ has infinite R.O.C. on all sub-intervals $[u, v] \subseteq [0, K]$. Let s, T be chosen fixed as in Lemma 4.6. Then one has*

$$\gamma[T, K] \subseteq \{\gamma_s + \lambda(\gamma_T - \gamma_s) : \lambda \in \mathbb{R}\}.$$

Proof. Let us denote the straight line $\{\gamma_s + \lambda(\gamma_T - \gamma_s) : \lambda \in \mathbb{R}\}$ by l . Define

$$T^* = \sup \{t \in [T, K] : \gamma [T, t] \subseteq l\}.$$

Suppose on the contrary that $T^* < K$. We will reach a contradiction by considering the following two cases.

Case 1: $\gamma_{T^*} \neq \gamma_s$.

There exists $t > T^*$ such that $\gamma_t \notin l$ and t is chosen to be close enough to T^* so that

$$|\gamma_u - \gamma_{T^*}| < \delta_1 |\gamma_{T^*} - \gamma_s|$$

for all $u \in [T^*, t]$ (with δ_1 to be specified later). By choosing $a = s, b = T^*, d = t$ in Lemma 4.7, one can find $\delta_1, \delta_2, r > 0$ such that the four closed sets

$$\mathcal{A}_{s,t}\gamma|_{[T^*,t]}, \quad \mathcal{A}_{s,t}\gamma|_{[T^*,t]} - (1, 0), \quad \overline{B(\mathcal{A}_{s,t}(m_{s,T^*}), r)}, \quad \overline{B(\mathcal{A}_{s,t}(m_{s,T^*}) - (\delta_2, 0), r)}$$

are all disjoint, where $m_{s,T^*} \triangleq (\gamma_s + \gamma_{T^*})/2$. The same lemma shows that

$$\overline{B(\mathcal{A}_{s,t}(m_{s,T^*}), r)} \cup \overline{B(\mathcal{A}_{s,t}(m_{s,T^*}) - (\delta_2, 0), r)} \subseteq (0, 1) \times \mathbb{R},$$

and since $\gamma_t \notin l$, L^{s,T^*} cannot be horizontal so one also knows that

$$\overline{B(\mathcal{A}_{s,t}(m_{s,T^*}) - (\delta_2, 0), r)} \cap L^{s,T^*} = \emptyset. \quad (4.10)$$

Here L^{s,T^*} denotes the line segment joining the origin to $\mathcal{A}_{s,t}(\gamma_{T^*})$. In what follows, we assume that L^{s,T^*} is parametrised on $[s, T^*]$.

Let Ψ be a smooth one-form supported on $B(\mathcal{A}_{s,t}(m_{s,T^*}), r)$ such that

$$\int_s^{T^*} \Psi(dL_v^{s,T^*}) = 1. \quad (4.11)$$

Such a Ψ clearly exists. Let us define

$$\Phi(x, y) = \begin{cases} \Psi(x, y), & (x, y) \in B(\mathcal{A}_{s,t}(m_{s,T^*}), r) := B_1 \\ -\Psi(x + \delta_2, y), & (x, y) \in B(\mathcal{A}_{s,t}(m_{s,T^*}) - (\delta_2, 0), r) := B_2, \\ 0, & (x, y) \in (0, 1) \times \mathbb{R} \setminus (B_1 \cup B_2) \\ \Phi(x \bmod 1, y), & (x, y) \notin (0, 1) \times \mathbb{R} \setminus (B_1 \cup B_2). \end{cases}$$

Note that the smooth one-form Φ satisfies the two conditions in Theorem 3.4. One can now write

$$\int_s^t \Phi(d[\mathcal{A}_{s,t}\gamma]_v) = \int_s^T \Phi(d[\mathcal{A}_{s,t}\gamma]_v) + \int_T^{T^*} \Phi(d[\mathcal{A}_{s,t}\gamma]_v) + \int_{T^*}^t \Phi(d[\mathcal{A}_{s,t}\gamma]_v).$$

According to Lemma 4.6, one has

$$\int_s^T \Phi(d[\mathcal{A}_{s,t}\gamma]_v) = \int_s^T \Phi(dL_v^{s,T*}).$$

Since $\Phi = 0$ on $(0, 1) \times \mathbb{R} \setminus (B_1 \cup B_2)$ and $\mathcal{A}_{s,t}\gamma|_{[T^*,t]}$ is disjoint from $B_1 \cup B_2$, one sees that

$$\int_{T^*}^t \Phi(d[\mathcal{A}_{s,t}\gamma]_v) = 0.$$

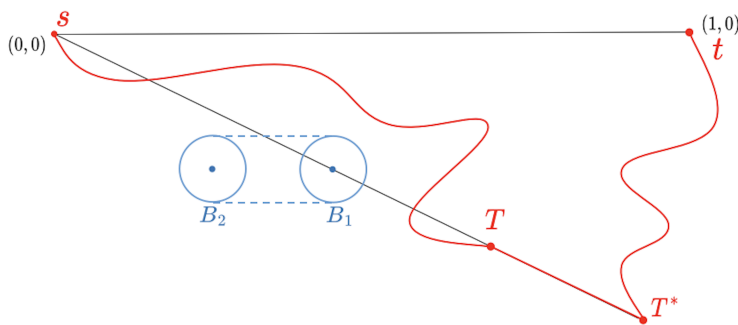
In addition, since $\gamma[T, T^*] \subseteq l$ one also knows that

$$\int_T^{T^*} \Phi(d[\mathcal{A}_{s,t}\gamma]_v) = \int_T^{T^*} \Phi(dL_v^{s,T*}).$$

Therefore, one concludes that

$$\int_s^t \Phi(d[\mathcal{A}_{s,t}\gamma]_v) = \int_s^{T^*} \Phi(dL_v^{s,T*}).$$

Note that L^{s,T^*} is disjoint from B_2 . By the construction of Φ , the last integral is thus also equal to the same integral along Ψ whose value is $1 \neq 0$ (cf. 4.11). This leads to a contradiction to the generalised line integral condition (3.3) in Theorem 3.4.



Case 2: $\gamma_{T^*} = \gamma_s$.

In this case, $\gamma_T \neq \gamma_{T^*}$ because $\gamma_T \neq \gamma_s$. Then there exists $t > T^*$ such that $\gamma_t \notin l$ and t is close enough to T^* so that

$$|\gamma_u - \gamma_{T^*}| < \delta_1 |\gamma_{T^*} - \gamma_T|$$

for all $u \in [T^*, t]$ (with δ_1 to be specified later). By choosing $a = T, b = T^*, d = t$ in Lemma 4.7, one can now proceed in exactly the same way as in Case 1 to construct a one-form Φ which satisfies the two conditions in Theorem 3.4 but

$$\int_T^t \Phi(d[\mathcal{A}_{T,t}\gamma]_v) = 1 \neq 0.$$

This again contradicts the generalised integral condition (3.3). □

The proof of Theorem 4.1 can be easily reduced to the two-dimensional situation which we first handle.

Theorem 4.9. *Let $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ be a continuous BV path. Suppose that $\log S(\gamma)_{s,t}$ has infinite R.O.C. for all $s < t \in [0, 1]$. Then $\gamma_t = f(t) \cdot \vec{v}$ for some real-valued function $f : [0, 1] \rightarrow \mathbb{R}$ and some fixed vector $\vec{v} \in \mathbb{R}^2$.*

Proof. By Appendix B in the arXiv version of [2], there exist a $K \in [0, \infty)$, a non-decreasing function $L : [0, 1] \rightarrow [0, K]$ and $\hat{\gamma} : [0, K] \rightarrow \mathbb{R}^2$ such that $\hat{\gamma}_{L(\cdot)} = \gamma$ and $\hat{\gamma}$ is Lipschitz continuous with unit speed. Since the signature is invariant under reparametrisation, one knows that $\log S(\hat{\gamma})_{s,t}$ has infinite R.O.C. for all $s < t \in [0, K]$. Let

$$\tau = \sup \{t : \hat{\gamma}'_u = 0 \text{ for almost all } u \in [0, t]\}.$$

Given $\varepsilon > 0$, there exist a differentiable point $s \in [\tau, \tau + \varepsilon/2)$ and a sequence $(t_n)_{n=1}^\infty$ such that $t_n \downarrow s$ and $\hat{\gamma}_s \neq \hat{\gamma}_{t_n}$ for all n . Let $\delta \in (0, \varepsilon/2)$ be chosen such that

$$|\hat{\gamma}_u - \hat{\gamma}_s - \hat{\gamma}'_s(u - s)| \leq \frac{1}{6}(u - s), \quad \left| \left| \frac{\hat{\gamma}_u - \hat{\gamma}_s}{u - s} \right| - 1 \right| \leq \frac{1}{6}$$

whenever $|u - s| < \delta$. Take N to be such that $|t_N - s| < \delta$ and let

$$T = \inf \{u \geq s : \hat{\gamma}_u = \hat{\gamma}_{t_N}\}.$$

Clearly, the numbers s, T satisfy the assumptions of Lemma 4.8. According to Lemma 4.8, one knows that

$$\hat{\gamma}[T, K] \subseteq \{\hat{\gamma}_s + \lambda(\hat{\gamma}_T - \hat{\gamma}_s) : \lambda \in \mathbb{R}\} =: l_\varepsilon$$

and in particular, $\hat{\gamma}[\tau + \varepsilon, K] \subseteq l_\varepsilon$. Since this is true for all $\varepsilon > 0$, the straight lines $\{l_\varepsilon\}_{\varepsilon > 0}$ are clearly consistent and thus all identical. Let us call this line l . It follows that $\hat{\gamma}[\tau, K] \subseteq l$. Since $\hat{\gamma}'_t = 0$ for almost all $t \in [0, \tau]$, one has $\hat{\gamma}_t = \hat{\gamma}_\tau$ for all $t \in [0, \tau]$. As a result, $\hat{\gamma}[0, K] \subseteq l$. This provides that $\hat{\gamma}$ (and thus γ) lives on a straight line. \square

Proof of Theorem 4.1. Suppose that $\gamma_t = (x_t^1, x_t^2, \dots, x_t^d)$ and assume that $\gamma_0 = \mathbf{0}$. The conclusion is trivial if the image of γ is a single point. Otherwise, let us assume that x_t^1 is not identically zero. For each $i \neq 1$, the logarithmic signature of the path $\gamma_t^i \triangleq (x_t^1, x_t^i)$ has infinite R.O.C. on $[s, t]$ for all $s < t \in (0, 1)$. By Theorem 4.9, one has $\gamma_t^i = f_i(t)(a_i, b_i)$ for some real-valued function $f_i(t)$ and some vector (a_i, b_i) with $a \neq 0$. It is clear that $x_t^1 = f_i(t)a_i$ and thus $\gamma_t^i = x_t^1 \cdot (1, c_i)$ where $c_i \triangleq b_i/a_i$. It follows that $\gamma_t = x_t^1 \cdot (1, c_2, \dots, c_d)$. This completes the proof of the theorem. \square

5 Second order integral identities

We continue to assume that $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ is a weakly geometric rough path satisfying the normalisation condition (3.1). In Section 3, we showed that if the logarithmic signature of γ has infinite R.O.C., the path γ must satisfy the line integral condition (3.2). In this and the next sections, by using the method of Cartan's path development we will show that γ has to satisfy *an infinite system of iterated integral identities* where the relation (3.2) appears to be the first level of them.

To better illustrate the essential idea, it is helpful to first discuss the derivation of *second order* iterated integral identities. This is our main goal in the current section and the main result is stated as follows.

Theorem 5.1. *Suppose that $\log S(\gamma)$ has infinite R.O.C. Then the following identity*

$$\begin{aligned} & (1 - \cosh b) \int_{0 < s < t < 1} (e^{2k\pi i \cdot x_s + b(x_t - x_s)} - e^{2k\pi i \cdot x_t + b(x_s - x_t)}) dy_s dy_t \\ & + (\sinh b) \left(\int_0^1 e^{bx_s} dy_s \right) \left(\int_0^1 e^{(2k\pi i - b)x_s} dy_s \right) = 0 \end{aligned} \quad (5.1)$$

holds true for all $k \in \mathbb{Z} \setminus \{0\}$ and $b \in \mathbb{C}$.

5.1 A second order extension of the line integral condition

In contrast to the line integral condition (3.2), Theorem 5.1 yields a *continuum* family of integral identities from which one can extract various information. The following result is a consequence of Theorem 5.1 which can be viewed as a second order extension of (3.2).

Corollary 5.2. *Under the assumption of Theorem 5.1, one has*

$$\int_{0 < s < t < 1} (e^{2\pi i(px_s + qx_t)} - e^{2\pi i(px_t + qx_s)}) dy_s dy_t = 0 \quad (5.2)$$

for all $p, q \in \mathbb{Z} \setminus \{0\}$ with $p + q \neq 0$.

Proof. Let us denote

$$\varphi(b) \triangleq \int_{0 < s < t < 1} (e^{2k\pi i \cdot x_s + b(x_t - x_s)} - e^{2k\pi i \cdot x_t + b(x_s - x_t)}) dy_s dy_t$$

and

$$\psi(b) \triangleq \left(\int_0^1 e^{bx_s} dy_s \right) \left(\int_0^1 e^{(2k\pi i - b)x_s} dy_s \right).$$

Differentiating (5.1) with respect to b gives that

$$-\sinh b \cdot \varphi(b) + (1 - \cosh b)\varphi'(b) + \cosh b \cdot \psi(b) + \sinh b \cdot \psi'(b) = 0. \quad (5.3)$$

By further differentiating (5.3) and taking $b = 2l\pi i$, one obtains that

$$-\varphi(2l\pi i) + 2\psi'(2l\pi i) = 0.$$

Now suppose that $l \neq k$ and $l \neq 0$. The line integral condition (3.2) yields that

$$\begin{aligned} \psi'(2l\pi i) &= \left(\int_0^1 x_s e^{2l\pi i \cdot x_s} dy_s \right) \left(\int_0^1 e^{2(k-l)\pi i \cdot x_s} dy_s \right) \\ &\quad - \left(\int_0^1 e^{2l\pi i \cdot x_s} dy_s \right) \left(\int_0^1 x_s e^{2(k-l)\pi i \cdot x_s} dy_s \right) = 0. \end{aligned}$$

As a result, one concludes that

$$\varphi(2l\pi i) = \int_{0 < s < t < 1} \left(e^{2(k-l)\pi i \cdot x_s + 2l\pi i \cdot x_t} - e^{2(k-l)\pi i \cdot x_t + 2l\pi i \cdot x_s} \right) dy_s dy_t = 0.$$

This gives the desired identity (5.2) with $p = k - l$ and $q = l$. \square

Remark 5.3. *The line integral identity (3.2) is also a consequence of Theorem 5.1. Indeed, let k be a nonzero even integer in (5.1). By taking $b = k\pi i$ in the relation (5.3), one finds that*

$$0 = \psi(k\pi i) = \left(\int_0^1 e^{k\pi i x_s} dy_s \right)^2 = 0 \iff \int_0^1 e^{k\pi i x_s} dy_s = 0.$$

This is exactly the previous line integral identity (3.2) since k is even.

Remark 5.4. *According to (3.2), one has*

$$\int_{0 < s < t < 1} \left(e^{2\pi i(p x_s + q x_t)} + e^{2\pi i(p x_t + q x_s)} \right) dy_s dy_t = \left(\int_0^1 e^{2\pi i p x_s} dy_s \right) \left(\int_0^1 e^{2\pi i q x_t} dy_t \right) = 0.$$

In particular, the relation (5.2) can also be rewritten as

$$\int_{0 < s < t < 1} e^{2\pi i(p x_s + q x_t)} dy_s dy_t = 0$$

for all $p, q \in \mathbb{Z} \setminus \{0\}$ with $p + q \neq 0$.

Remark 5.5. *The information encoded by the condition (5.1) is larger than the one encoded in (5.2). For instance, by further differentiating (5.3) and setting $b = 0$ one obtains that*

$$-\varphi(0) + \psi(0) + \psi'(0) = 0. \tag{5.4}$$

Note that $\psi(0) = 0$ as a consequence of (3.2). Therefore, the relation (5.4) becomes

$$-\int_{0 < s < t < 1} \left(e^{2k\pi i \cdot x_s} - e^{2k\pi i \cdot x_t} \right) dy_s dy_t - y_1 \int_0^1 x_s e^{2k\pi i \cdot x_s} dy_s = 0.$$

After simplification, this can be rewritten as the following identity:

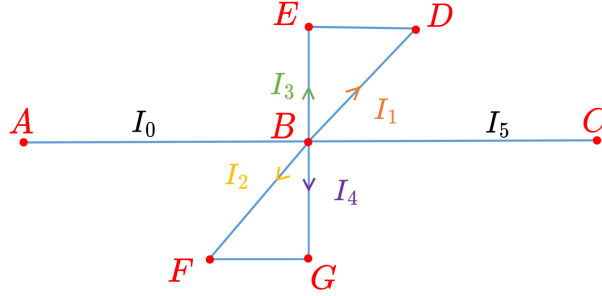
$$\int_0^1 y_t e^{2k\pi i \cdot x_t} dy_t - y_1 \int_0^1 x_t e^{2k\pi i \cdot x_t} dy_t = 0.$$

This new identity cannot be implied by (3.2) or (5.2). One can also show by a similar type of calculation that

$$\int_{0 < s < t < 1} e^{\pi i (px_s + qx_t)} dy_s dy_t = 0$$

for all odd integers p, q with $p + q \neq 0$, as a consequence of Theorem 5.1.

Example 5.6. We use one example to illustrate the usefulness of the second order integral identity (5.2). Consider the following piecewise linear path defined by a ‘‘Figure Eight’’ trajectory:



Here A is the origin, $B = (1/2, 0)$, $C = (1, 0)$ and

$$\vec{BD} = \left(\frac{1}{4}, \frac{1}{4}\right), \quad \vec{BE} = \left(0, \frac{1}{4}\right), \quad \vec{BG} = \left(0, -\frac{1}{4}\right), \quad \vec{BF} = \left(-\frac{1}{4}, -\frac{1}{4}\right).$$

The path γ is defined by

$$I_0 \rightarrow I_1 \rightarrow I_2 \rightarrow I_3 \rightarrow I_4 \rightarrow I_5,$$

where I_1 (respectively, I_3) denotes the loop $BDEB$ (respectively, $BEDE$) and I_2 (respectively, I_4) denotes the loop $BFGB$ (respectively, $BGFB$). Explicit calculation shows that

$$\int_{0 < s < t < 1} e^{2\pi i x_s + 4\pi i x_t} dy_s dy_t \approx -0.05 - 0.08i \neq 0.$$

Therefore, one concludes from Corollary 5.2 and Remark 5.4 that $\log S(\gamma)$ has finite R.O.C. Note that this example *cannot* be handled by Theorem 3.4; it is easily seen that

$$\int_0^1 \Phi(d\gamma_t) = \int_{I_0 \cup I_5} \Phi(d\gamma_t) = 0$$

for any Φ satisfying the conditions in that theorem.

In the following subsections, we develop the proof of Theorem 5.1. The key insight is to make use of path developments into *complex semisimple* Lie algebras.

5.2 Homogeneous projection formulae

Our strategy relies on certain tensor projections as a starting point. We first define these operations. Let $\{e_1, e_2\}$ be the standard basis of \mathbb{R}^2 .

Definition 5.7. For each $m \geq 1$, we define $D_m : T((\mathbb{R}^2)) \rightarrow T((\mathbb{R}^2))$ to be linear operator induced by

$$D_m(e_{j_1} \otimes \cdots \otimes e_{j_n}) \triangleq \begin{cases} e_{j_1} \otimes \cdots \otimes e_{j_n}, & \#\{l : j_l = 2\} = m; \\ 0, & \text{otherwise.} \end{cases}$$

In other words, D_m projects a tensor series onto the sub-series whose components have precisely m of e_2 's.

Suppose that γ is a two-dimensional weakly geometric rough path whose logarithmic signature $L \triangleq \log S(\gamma)$ has infinite R.O.C. It is easy to see that $D_m L$ has infinite R.O.C. for all $m \geq 1$. As we will show, the main philosophy is that

$D_m L$ has infinite R.O.C. \implies suitable m -th order iterated integral identities.

When $m = 1$, this gives back the line integral condition (3.2). Theorem 5.1 follows from the consideration of $D_2 L$. The higher order iterated integral identities we will derive in Section 6 are based on the consideration of general $D_m L$ ($m \geq 3$).

In order to prove Theorem 5.1, we first derive the expressions of $D_1 L$ and $D_2 L$, leaving the more general situation to Section 6. This is contained in the lemma below. Recall that $\{B_m\}$ are the Bernoulli numbers arising from the first Hausdorff series H_1 (cf. (3.11)).

Lemma 5.8. *The following formulae for $D_1 L$ and $D_2 L$ hold true:*

$$D_1 L = \sum_{m=0}^{\infty} \frac{B_m}{m!} \text{ad}_{e_1}^m \left(\int_0^1 e^{x_t \text{ad}_{e_1}}(e_2) dy_t \right) \quad (5.5)$$

and

$$\begin{aligned} D_2 L &= \frac{1}{2} \sum_{m=0}^{\infty} \frac{B_m}{m!} \text{ad}_{e_1}^m \left(\int_{0 < s < t < 1} [e^{x_s \text{ad}_{e_1}}(e_2), e^{x_t \text{ad}_{e_1}}(e_2)] dy_s dy_t \right) \\ &\quad + \frac{1}{2} \sum_{m, l \geq 0} \frac{B_m B_l}{m! l!} \sum_{k=1}^m \text{ad}_{e_1}^{k-1} \left([\text{ad}_{e_1}^l \left(\int_0^1 e^{x_s \text{ad}_{e_1}}(e_2) dy_s \right), \text{ad}_{e_1}^{m-k} \left(\int_0^1 e^{x_t \text{ad}_{e_1}}(e_2) dy_t \right)] \right). \end{aligned} \quad (5.6)$$

Proof. Let \tilde{L} denote the logarithmic signature of $\tilde{\gamma} \triangleq \gamma \sqcup \overleftarrow{e_1}$. Recall from (3.13) that

$$L = B(\tilde{L}, e_1) = \sum_{n=1}^{\infty} H_n(\tilde{L}, e_1).$$

By applying projections D_1 and D_2 to both sides, one finds that

$$D_1 L = H_1(D_1 \tilde{L}, e_1), \quad D_2 L = H_1(D_2 \tilde{L}, e_1) + H_2(D_1 \tilde{L}, e_1). \quad (5.7)$$

Here we used the fact that the first level of \tilde{L} does not contain the e_1 -component (due to the normalisation (3.1)).

We first compute $D_1 \tilde{L}$ and $D_2 \tilde{L}$. Recall from Lemma 3.9 that

$$\begin{aligned} e^{\tilde{L}} &= \sum_{n=0}^{\infty} \int_{0 < t_1 < \dots < t_n < 1} e^{x_{t_1} \text{ad}_{e_1}(e_2)} \otimes \dots \otimes e^{x_{t_n} \text{ad}_{e_1}(e_2)} dy_{t_1} \dots dy_{t_n} \\ &= 1 + \tilde{L} + \frac{1}{2} \tilde{L}^{\otimes 2} + \dots = 1 + (D_1 \tilde{L} + D_2 \tilde{L} + \dots) + \frac{1}{2} (D_1 \tilde{L} + D_2 \tilde{L} + \dots)^{\otimes 2} + \dots \end{aligned}$$

By applying D_i ($i = 1, 2$) to both sides, one finds that

$$D_1 \tilde{L} = \int_0^1 e^{x_t \text{ad}_{e_1}(e_2)} dy_t \quad (5.8)$$

and

$$\begin{aligned} D_2 \tilde{L} &= \int_{0 < s < t < 1} e^{x_s \text{ad}_{e_1}(e_2)} \otimes e^{x_t \text{ad}_{e_1}(e_2)} dy_s dy_t - \frac{1}{2} D_1 \tilde{L} \otimes D_1 \tilde{L} \\ &= \frac{1}{2} \int_{0 < s < t < 1} [e^{x_s \text{ad}_{e_1}(e_2)}, e^{x_t \text{ad}_{e_1}(e_2)}] dy_s dy_t. \end{aligned} \quad (5.9)$$

Next, we compute the Hausdorff series H_1 and H_2 . According to the definition (3.11) of H_1 , one has

$$H_1(D_i \tilde{L}, e_1) = \sum_{m=0}^{\infty} \frac{B_m}{m!} \text{ad}_{e_1}^m(D_m \tilde{L}) \quad (i = 1, 2). \quad (5.10)$$

Respectively, the formula (3.12) for H_n (with $n = 2$) yields that

$$\begin{aligned} H_2(D_1 \tilde{L}, e_1) &= \frac{1}{2} \left(H_1 \frac{\partial}{\partial e_1} \right) H_1(D_1 \tilde{L}, e_1) \\ &= \frac{1}{2} \sum_{m=0}^{\infty} \frac{B_m}{m!} \left(H_1 \frac{\partial}{\partial e_1} \right) (\text{ad}_{e_1}^m(D_1 \tilde{L})) \\ &= \frac{1}{2} \sum_{m=0}^{\infty} \frac{B_m}{m!} \sum_{k=1}^m \text{ad}_{e_1}^{k-1} \circ \text{ad}_{H_1(D_1 \tilde{L}, e_1)} \circ \text{ad}_{e_1}^{m-k}(D_1 \tilde{L}) \\ &= \frac{1}{2} \sum_{m, l \geq 0} \frac{B_m B_l}{m! l!} \sum_{k=1}^m \text{ad}_{e_1}^{k-1}([\text{ad}_{e_1}^l(D_1 \tilde{L}), \text{ad}_{e_1}^{m-k}(D_1 \tilde{L})]). \end{aligned} \quad (5.11)$$

After substituting the expressions (5.8) and (5.9) into (5.10) and (5.11), the lemma follows from the relation (5.7). \square

Recapturing Theorem 3.1 through the D_1 -projection

The formula (5.5) easily recovers Theorem 3.1. Indeed, by taking the development

$$F_\lambda : e_1 \mapsto A, e_2 \mapsto D \text{ with Lie structure } [A, D] = \lambda D. \quad (5.12)$$

The relation (5.5) yields

$$\hat{F}_\lambda(D_1L) = \sum_{m=0}^{\infty} \frac{B_m}{m!} \lambda^m \left(\int_0^1 e^{\lambda x_t} dy_t D \right) = \frac{\lambda}{e^\lambda - 1} \times \left(\int_0^1 e^{\lambda x_t} dy_t \right) D.$$

Since the left hand side defines an entire function in $\lambda \in \mathbb{C}$, one must have

$$\int_0^1 e^{2k\pi i \cdot x_t} dy_t = 0 \quad \forall k \in \mathbb{Z} \setminus \{0\}. \quad (5.13)$$

Of course, this argument is essentially the same as our earlier proof of Theorem 3.1.

An interesting point is that in the D_1L case, one can achieve more by showing that the line integral condition (3.2) is also *sufficient* for D_1L to have infinite R.O.C. Indeed, one can rewrite (5.5) as

$$D_1L = \sum_{m,p=0}^{\infty} \frac{B_m}{m!p!} \text{ad}_{e_1}^{m+p}(e_2) \int_0^1 x_t^p dy_t = \sum_{N=0}^{\infty} \left(\sum_{m=0}^N \frac{B_m}{m!(N-m)!} \int_0^1 x_t^{N-m} dy_t \right) \text{ad}_{e_1}^N(e_2).$$

Since the Lie polynomials $\text{ad}_{e_1}^N(e_2)$ ($N \geq 0$) all have different degrees and

$$C_1 \|\text{ad}_{e_1}^N(e_2)\| \leq 2^N,$$

one has for every $\lambda > 0$ that

$$\begin{aligned} F(\lambda) &\triangleq \sum_{N=0}^{\infty} \|\pi_{N+1}(D_1L)\| \lambda^N \\ &\leq \sum_{N=0}^{\infty} \left| \sum_{m=0}^N \frac{B_m}{m!(N-m)!} \int_0^1 x_t^{N-m} dy_t \right| \lambda^N 2^N := \sum_{N=0}^{\infty} |c_N| (2\lambda)^N. \end{aligned} \quad (5.14)$$

Observe that

$$\sum_{N=0}^{\infty} c_N (2\lambda)^N = \left(\sum_m \frac{B_m}{m!} (2\lambda)^m \right) \left(\sum_p \frac{1}{p!} \int_0^1 x_t^p dy_t \right) = \frac{2\lambda}{e^{2\lambda} - 1} \times \int_0^1 e^{2\lambda x_t} dy_t.$$

Now suppose that the line integral condition (3.2) holds. Then the function

$$z \mapsto \frac{z}{e^z - 1} \times \int_0^1 e^{z x_t} dy_t$$

is entire. As a result, the power series $\sum_{N \geq 0} c_N z^N$ has infinite R.O.C. and so does

$\sum_{N \geq 0} |c_N| z^N$. In view of (5.14), this implies that D_1L has infinite R.O.C.

To summarise, one has obtained the following neat result.

Proposition 5.9. *Let $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ be a weakly geometric rough path satisfying the normalisation condition (3.1). Then $D_1 \log S(\gamma)$ has infinite R.O.C. if and only if $\int_0^1 e^{2k\pi i x_t} dy_t = 0$ for all nonzero integers k .*

5.3 Complex semisimple Lie algebras

In order to prove Theorem 5.1, one needs to investigate the D_2 -projection. Our method relies on path developments into a particular type of Lie algebras: *complex semisimple Lie algebras*. We first provide a quick review on relevant concepts.

Definition 5.10. A (finite dimensional) complex Lie algebra \mathfrak{g} is said to be *semisimple* if it can be decomposed into a direct sum $\mathfrak{g} \cong \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_r$, where each \mathfrak{g}_i is a *simple* Lie subalgebra in the sense that it does not contain non-trivial proper ideals.

A key concept in semisimple Lie theory is the notion of Cartan subalgebras. Let \mathfrak{g} be a complex semisimple Lie algebra.

Definition 5.11. A *Cartan subalgebra* \mathfrak{h} of \mathfrak{g} is a subspace which satisfies the following two properties:

- (i) \mathfrak{h} is a maximal abelian subalgebra;
- (ii) $\text{ad}_H \in \text{End}(\mathfrak{g})$ is diagonalisable (over \mathbb{C}) for each $H \in \mathfrak{h}$.

It is well known that a Cartan subalgebra always exists and is unique up to conjugation in \mathfrak{g} . Let \mathfrak{h} be a Cartan subalgebra of \mathfrak{g} . Given any representation $\rho : \mathfrak{g} \rightarrow \text{End}(W)$ over a finite dimensional complex vector space W , all elements of \mathfrak{h} are simultaneously diagonalisable when they are viewed as linear transformations over W . As a result, W admits a decomposition into common \mathfrak{h} -eigenspaces. To be precise, a complex linear functional $\mu \in \mathfrak{h}^*$ is said to be a *weight* for ρ if the subspace

$$W^\mu \triangleq \{w \in W : \rho(H)(w) = \mu(H)w \ \forall H \in \mathfrak{h}\} \quad (5.15)$$

is non-trivial. There are at most finitely many weights for ρ due to finite dimensionality and their collection is denoted as $\Pi(\rho)$. The space W admits a decomposition (simultaneous diagonalisation)

$$W = \bigoplus_{\mu \in \Pi(\rho)} W^\mu,$$

where for each $H \in \mathfrak{h}$, W^μ is an eigenspace of $\rho(H)$ with eigenvalue $\mu(H)$ ($\mu \in \Pi(\rho)$).

The above general consideration, being applied to the adjoint representation $\text{ad} : \mathfrak{g} \rightarrow \text{End}(\mathfrak{g})$, gives the so-called *root space decomposition* of \mathfrak{g} . Let $\alpha \in \mathfrak{h}^*$. Similar to (5.15) we define the subspace

$$\mathfrak{g}^\alpha \triangleq \{X \in \mathfrak{g} : \text{ad}_H(X) = \alpha(H)X \ \forall H \in \mathfrak{h}\}.$$

It is readily checked that $\mathfrak{g}^0 = \mathfrak{h}$ and $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] \subseteq \mathfrak{g}^{\alpha+\beta}$ for all $\alpha, \beta \in \mathfrak{h}^*$. A complex linear functional $\alpha \in \mathfrak{h}^*$ is called a *root* of \mathfrak{g} with respect to \mathfrak{h} if it is a weight for the adjoint

representation (i.e. if $\mathfrak{g}^\alpha \neq \{0\}$). In this case, \mathfrak{g}^α is called the *root space* associated with α . It is clear that there are at most finitely many roots of \mathfrak{g} .

Theorem 5.12 (Root Space Decomposition). *Let Δ be the set of all nonzero roots with respect to a given Cartan subalgebra \mathfrak{h} . Then \mathfrak{g} can be written as the direct sum*

$$\mathfrak{g} = \mathfrak{h} + \sum_{\alpha \in \Delta} \mathfrak{g}^\alpha. \quad (5.16)$$

In this decomposition, one has $\dim \mathfrak{g}^\alpha = 1$ for each $\alpha \in \Delta$. Moreover, if $\alpha, \beta, \alpha + \beta \in \Delta$ then $[\mathfrak{g}^\alpha, \mathfrak{g}^\beta] = \mathfrak{g}^{\alpha+\beta}$.

The following basic example will play an essential role in the proof of Theorem 5.1 (and more significantly, of Theorem 6.5 below).

Example 5.13. Let $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ be the Lie algebra of $n \times n$ matrices over \mathbb{C} with zero trace. A Cartan subalgebra \mathfrak{h} can be taken as the subspace of diagonal matrices in \mathfrak{g} . One has $\dim \mathfrak{h} = n - 1$. For each $1 \leq i \leq n$, let $\mu_i \in \mathfrak{h}^*$ be the linear functional defined by taking the i -th diagonal entry of $H \in \mathfrak{h}$. Then the set of nonzero roots is given by

$$\Delta = \{\lambda_{ij} \triangleq \mu_i - \mu_j : 1 \leq i \neq j \leq n\}.$$

Respectively, the root spaces are given by

$$\mathfrak{g}^{\lambda_{ij}} = \mathbb{C} \cdot E_{ij}, \quad (i \neq j)$$

where E_{ij} denotes the matrix whose (i, j) -entry is one and all other entries are zero.

5.4 Proof of Theorem 5.1 through $\mathfrak{sl}_3(\mathbb{C})$ -development

A nice feature about the root space decomposition is that the adjoint actions by elements of the Cartan subalgebra are simultaneously diagonalised into root spaces (eigenspaces). In particular, if one applies a path development into a complex semisimple Lie algebra with e_1 being mapped into a Cartan element, the formulae (5.5, 5.6) for D_2L could potentially be projected along suitable root spaces to yield *scalar* equations similar to the one obtained in (3.18). This will allow one to treat D_2L as a scalar entire function and to obtain suitable integral conditions in a similar fashion as in the proof of Theorem 3.1.

To make this idea precise, let \mathfrak{g} be a complex semisimple Lie algebra with root space decomposition (5.16). We consider a path development $F : \mathbb{C}^2 \rightarrow \mathfrak{g}$ induced by

$$e_1 \mapsto A, \quad e_2 \mapsto D$$

where A, D are chosen to be such that

$$A \in \mathfrak{h}, \quad D = \sum_{\alpha \in \Delta} c_\alpha E_\alpha \in \sum_{\alpha \in \Delta} \mathfrak{g}^\alpha.$$

Here E_α is a generator of \mathfrak{g}^α (recall that $\dim \mathfrak{g}^\alpha = 1$) and $c_\alpha \in \mathbb{C}$. Under such a development, one has

$$\hat{F}([e^{x_s \text{ad}_{e_1}}(e_2), e^{x_t \text{ad}_{e_1}}(e_2)]) = \sum_{\alpha, \beta \in \Delta} c_\alpha c_\beta e^{\alpha(A)x_s + \beta(A)x_t} [E_\alpha, E_\beta].$$

A crucial point is that $[E_\alpha, E_\beta] \neq 0$ if $\alpha + \beta \in \Delta$ and one can also make $\alpha(A), \beta(A)$ arbitrary by varying A .

Proof of Theorem 5.1. As the simplest non-trivial example, in what follows we choose $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{C})$. According to Example 5.13, the Cartan subalgebra \mathfrak{h} consists of diagonal matrices in \mathfrak{g} and the roots are given by

$$\lambda_{ij}(H) = H_{ii} - H_{jj}, \quad H \in \mathfrak{h}. \quad (1 \leq i \neq j \leq 3)$$

Let a, b be two given complex numbers. We construct a path development $F_{a,b} : \mathbb{C}^2 \rightarrow \mathfrak{g}$ by specifying

$$e_1 \mapsto A \triangleq \begin{pmatrix} \frac{2a+b}{3} & 0 & 0 \\ 0 & \frac{b-a}{3} & 0 \\ 0 & 0 & -\frac{a+2b}{3} \end{pmatrix}, \quad e_2 \mapsto D \triangleq E_{12} + E_{23} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

In particular, one has

$$D \in \mathfrak{g}^{\lambda_{12}} + \mathfrak{g}^{\lambda_{2,3}}; \quad \lambda_{12}(A) = a, \quad \lambda_{23}(A) = b; \quad [E_{12}, E_{23}] = E_{13} \in \mathfrak{g}^{\lambda_{13}}.$$

It follows that

$$\begin{aligned} \hat{F}_{a,b}([e^{x_s \text{ad}_{e_1}}(e_2), e^{x_t \text{ad}_{e_1}}(e_2)]) &= [e^{ax_s} E_{12} + e^{bx_s} E_{23}, e^{ax_t} E_{12} + e^{bx_t} E_{23}] \\ &= (e^{ax_s + bx_t} - e^{ax_t + bx_s}) E_{13}. \end{aligned}$$

According to the formula (5.10) for $H_1(D_2 \tilde{L}, e_1)$, one finds that

$$\begin{aligned} \hat{F}_{a,b}(H_1(D_2 \tilde{L}, e_1)) &= \frac{1}{2} \sum_{m=0}^{\infty} \frac{B_m}{m!} \text{ad}_A^m \left(\int_{0 < s < t < 1} (e^{ax_s + bx_t} - e^{ax_t + bx_s}) dy_s dy_t \right) E_{13} \\ &= \frac{1}{2} \sum_{m=0}^{\infty} \frac{B_m}{m!} (a+b)^m \int_{0 < s < t < 1} (e^{ax_s + bx_t} - e^{ax_t + bx_s}) dy_s dy_t \cdot E_{13} \\ &= \frac{1}{2} \cdot \frac{a+b}{e^{a+b} - 1} \cdot \int_{0 < s < t < 1} (e^{ax_s + bx_t} - e^{ax_t + bx_s}) dy_s dy_t \cdot E_{13}. \end{aligned}$$

Similarly, by using the formula (5.11) for $H_2(D_1\tilde{L}, e_1)$ one has

$$\begin{aligned}\hat{F}_{a,b}(H_2(D_1\tilde{L}, e_1)) &= \frac{1}{2} \int_0^1 \int_0^1 dy_s dy_t \times \sum_{m,l \geq 0} \frac{B_m B_l}{m!l!} \\ &\quad \sum_{k=1}^m \text{ad}_A^{k-1} [a^l e^{ax_s} E_{12} + b^l e^{bx_s} E_{23}, a^{m-k} e^{ax_t} E_{12} + b^{m-k} e^{bx_t} E_{23}] \\ &= \frac{1}{2} \int_0^1 \int_0^1 dy_s dy_t \times \sum_{m,l \geq 0} \frac{B_m B_l}{m!l!} \\ &\quad \times \sum_{k=1}^m (a^l b^{m-k} e^{ax_s+bx_t} - a^{m-k} b^l e^{ax_t+bx_s}) (a+b)^{k-1} \cdot E_{13}.\end{aligned}$$

After evaluating the summations explicitly, one finds that

$$\begin{aligned}\hat{F}_{a,b}(H_2(D_1\tilde{L}, e_1)) &= \frac{1}{2} \left(\frac{(a+b)(e^b - e^a)}{(e^{a+b} - 1)(e^a - 1)(e^b - 1)} - \frac{b-a}{(e^a - 1)(e^b - 1)} \right) \\ &\quad \times \left(\int_0^1 e^{ax_t} dy_t \right) \left(\int_0^1 e^{bx_t} dy_t \right) E_{1,3}.\end{aligned}$$

Consequently,

$$\begin{aligned}\hat{F}_{a,b}(D_2L) &= \hat{F}_{a,b}(H_1(D_2\tilde{L}, e_1)) + \hat{F}_{a,b}(H_2(D_1\tilde{L}, e_1)) \\ &= \frac{1}{2} E_{1,3} \times \left(\frac{a+b}{e^{a+b} - 1} \cdot \int_{0 < s < t < 1} (e^{ax_s+bx_t} - e^{ax_t+bx_s}) dy_s dy_t \right. \\ &\quad \left. + \left(\frac{(a+b)(e^b - e^a)}{(e^{a+b} - 1)(e^a - 1)(e^b - 1)} - \frac{b-a}{(e^a - 1)(e^b - 1)} \right) \left(\int_0^1 e^{ax_t} dy_t \right) \left(\int_0^1 e^{bx_t} dy_t \right) \right).\end{aligned}$$

By multiplying $(e^a - 1)(e^b - 1)(e^{a+b} - 1)$ on both sides, one obtains that

$$\begin{aligned}(e^a - 1)(e^b - 1)(e^{a+b} - 1) \hat{F}_{a,b}(D_2L) &= \frac{1}{2} E_{13} \times \left((a+b)(e^a - 1)(e^b - 1) \int_{0 < s < t < 1} (e^{ax_s+bx_t} - e^{ax_t+bx_s}) dy_s dy_t \right. \\ &\quad \left. + ((a+b)(e^b - e^a) - (b-a)(e^{a+b} - 1)) \left(\int_0^1 e^{ax_t} dy_t \right) \left(\int_0^1 e^{bx_t} dy_t \right) \right).\end{aligned}\quad (5.17)$$

Under the assumption that D_2L has infinite R.O.C., the left hand side of (5.17) defines an entire function in $(a, b) \in \mathbb{C}^2$. Let $k \in \mathbb{Z} \setminus \{0\}$ be given fixed and let $a, b \in \mathbb{C}$ satisfy $a + b = 2k\pi i$. Then (5.17) becomes

$$\begin{aligned}0 &= 2k\pi i E_{13} \times \left((1 - \cosh b) \int_{0 < s < t < 1} (e^{2k\pi i \cdot x_s + b(x_t - x_s)} - e^{2k\pi i \cdot x_t + b(x_s - x_t)}) dy_s dy_t \right. \\ &\quad \left. + (\sinh b) \left(\int_0^1 e^{(2k\pi i - b)x_t} dy_t \right) \left(\int_0^1 e^{bx_t} dy_t \right) \right).\end{aligned}$$

Since $k \neq 0$, one concludes the desired integral identity (5.1). This completes the proof of Theorem 5.1. \square

Remark 5.14. *It is certainly possible to use other types of complex semisimple Lie algebras in the above analysis. We are not sure if the use of other root structures would lead to new integral identities that are not covered by (5.1).*

6 Higher order iterated integral identities

In this section, we establish the higher order counterpart of Corollary 5.2. To state our main theorem, we first introduce the following definitions.

Definition 6.1. A finite sequence of numbers (a_1, \dots, a_m) is said to be *non-degenerate* if it does not contain zero consecutive sums, i.e. if

$$a_j + a_{j+1} + \dots + a_k \neq 0$$

for all $1 \leq j \leq k \leq m$.

Definition 6.2. Let $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ be a weakly geometric rough path. For each $m \geq 1$, we define an analytic function $S_m^\gamma : \mathbb{C}^m \rightarrow \mathbb{C}$ in the following way. Given $a_1, \dots, a_m \in \mathbb{C}$, let $B : [0, 1] \rightarrow \mathbb{C}^m$ be the path defined by

$$B_t \triangleq \sum_{j=1}^m \left(\int_0^t e^{a_j x_s} dy_s \right) e_j$$

where $\{e_1, \dots, e_m\}$ denotes the standard basis of \mathbb{C}^m . We define $S_m^\gamma(a_1, \dots, a_m)$ to be the coefficient of the logarithmic signature of B with respect to the tensor $e_1 \otimes \dots \otimes e_m$, i.e.

$$S_m^\gamma(a_1, \dots, a_m) \triangleq (\log S(B))^{(1, \dots, m)}.$$

Example 6.3. By explicit calculation, one finds that

$$S_1^\gamma(a) = \int_0^1 e^{ax_s} dy_s, \quad S_2^\gamma(a, b) = \frac{1}{2} \int_{0 < s < t < 1} (e^{ax_s + bx_t} - e^{bx_s + ax_t}) dy_s dy_t$$

and

$$\begin{aligned} S_3^\gamma(a, b, c) = \int_{0 < s < t < r < 1} & \left(\frac{1}{3} e^{ax_s + bx_t + cx_r} - \frac{1}{6} e^{ax_s + cx_t + bx_r} - \frac{1}{6} e^{bx_s + ax_t + cx_r} \right. \\ & \left. - \frac{1}{6} e^{bx_s + cx_t + ax_r} - \frac{1}{6} e^{cx_s + ax_t + bx_r} + \frac{1}{3} e^{cx_s + bx_t + ax_r} \right) dy_s dy_t dy_r. \end{aligned}$$

Remark 6.4. *One can consider a general d -dimensional path*

$$B_t \triangleq \sum_{i=1}^d \left(\int_0^t e^{p_i x_s} dy_s \right) e_i$$

It is not difficult to see that the logarithmic signature coefficient of B with respect to $e_{i_1} \otimes \dots \otimes e_{i_m}$ is given by $S_m^\gamma(p_{i_1}, \dots, p_{i_m})$.

Our main theorem of this section is stated as follows. As before, we assume that $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ is a weakly geometric rough path satisfying the normalisation condition (3.1).

Theorem 6.5. *Suppose that $\log S(\gamma)$ has infinite R.O.C. Then one has*

$$S_m^\gamma(a_1, \dots, a_m) = 0$$

for all $m \geq 1$ and all non-degenerate sequences (a_1, \dots, a_m) satisfying $a_j \in 2\pi i\mathbb{Z}$ for each j .

Remark 6.6. *When $m = 1, 2$, Theorem 6.5 reduces to Theorem 3.1 and Corollary 5.2 respectively. It is possible to establish higher order versions of the stronger Theorem 5.1. However, since the general formulae become significantly more involved we decide not to pursue this generality.*

Inspired by the second order case, our strategy for proving Theorem 6.5 will be based on path developments into $\mathfrak{sl}_{m+1}(\mathbb{C})$. The general spirit is not-so-different from the second order argument. However, the underlying algebraic structure becomes much subtler and the argument involves several non-trivial combinatorial considerations.

In the following subsections, we develop the proof of Theorem 6.5 in a precise mathematical way. We will continue to use the notation introduced in Section 5.

6.1 A Chen-Strichartz type formula for $D_m \tilde{L}$

Recall that \tilde{L} is the logarithmic signature of the path $\tilde{\gamma} \triangleq \gamma \sqcup \overleftarrow{e}_1$ and D_m is the projection operator defined in Definition 5.7. As a starting point, we first derive an explicit formula for $D_m \tilde{L}$ which generalises (5.8, 5.9). Throughout the rest, we use Δ_m to denote the standard simplex $0 < t_1 < \dots < t_m < 1$.

Proposition 6.7. *For any $m \geq 1$, one has*

$$D_m \tilde{L} = \sum_{\sigma \in \mathcal{S}_m} \frac{(-1)^{e(\sigma)}}{m^2 \binom{m-1}{e(\sigma)}} \int_{\Delta_m} [e^{x_{t_{\sigma(1)}} \text{ad}_{e_1}}(e_2), \dots [e^{x_{t_{\sigma(m-1)}} \text{ad}_{e_1}}(e_2), e^{x_{t_{\sigma(m)}} \text{ad}_{e_1}}(e_2)] \dots] dy_{t_1} \dots dy_{t_m}. \quad (6.1)$$

Here we define

$$e(\sigma) \triangleq \#\{j = 1, \dots, m-1 : \sigma(j) > \sigma(j+1)\}$$

for each permutation $\sigma \in \mathcal{S}_m$.

Proof. Consider the path

$$\Gamma_t \triangleq \int_0^t e^{x_s \text{ad}_{e_1}}(e_2) dy_s \in W \triangleq T((\mathbb{C}^2)).$$

Its signature is given by

$$S(\Gamma) = \mathbf{1} + \sum_{m=1}^{\infty} \int_{\Delta_m} d\Gamma_{t_1} \boxtimes \cdots \boxtimes d\Gamma_{t_m} \in T((W)),$$

where \boxtimes denotes another tensor product that is independent of \otimes . According to the Chen-Strichartz formula (cf. [15, Theorem 3.27]), one has

$$\begin{aligned} \log S(\Gamma) = \sum_{m=1}^{\infty} \sum_{\sigma \in \mathcal{S}_m} \frac{(-1)^{e(\sigma)}}{m^2 \binom{m-1}{e(\sigma)}} \int_{\Delta_m} \llbracket e^{x_{t_{\sigma(1)}} \text{ad}_{e_1}}(e_2), \llbracket \cdots \\ \llbracket e^{x_{t_{\sigma(m-1)}} \text{ad}_{e_1}}(e_2), e^{x_{t_{\sigma(m)}} \text{ad}_{e_1}}(e_2) \rrbracket \cdots \rrbracket \rrbracket dy_{t_1} \cdots dy_{t_m} \in \mathcal{L}((W)), \end{aligned}$$

where $\llbracket \cdot, \cdot \rrbracket$ denotes the commutator for \boxtimes . Now recall from Lemma 3.9 that the signature of $\tilde{\gamma}$ is given by the formula 3.14 (without the last e^{e_1} -term). By applying the algebra homomorphism $\boxtimes \mapsto \otimes$ as well as the projection D_m to the above relation, the left hand side becomes $D_m \tilde{L}$ and the right hand side is precisely (6.1). \square

6.2 $\mathfrak{sl}_{m+1}(\mathbb{C})$ -development of the logarithmic signature

Our next step is to show that the logarithmic signature coefficients $S_m^\gamma(a_1, \dots, a_m)$ (cf. Definition 6.2) can be realised through suitable path developments. Let us formulate this fact in a slightly more general setting. Suppose that $p_1, \dots, p_d \in \mathbb{C}$ are given fixed numbers. We define the path

$$B_t \triangleq \sum_{j=1}^d \left(\int_0^t e^{p_j x_s} dy_s \right) e_j \in \mathbb{C}^d. \quad (6.2)$$

Let $m \geq 1$ and $I = (i_1, \dots, i_m) \in \{1, \dots, d\}^m$ be a given fixed word. We consider the path development $F : \mathbb{C}^2 \rightarrow \mathfrak{sl}_{m+1}(\mathbb{C})$ (cf. Example 5.13 for the relevant notation) induced by

$$F_I(e_1) \triangleq A \in \mathfrak{h}, \quad F_I(e_2) \triangleq E_{12} + E_{23} + \cdots + E_{m,m+1}, \quad (6.3)$$

where the Cartan element A is chosen to satisfy

$$[A, E_{k,k+1}] = p_{i_k} E_{k,k+1}, \quad k = 1, \dots, m. \quad (6.4)$$

It is a simple linear algebra exercise to see that such an F_I exists. Recall from Section 3.2.3 that the induced homomorphisms at the tensor and Lie series levels are both denoted as \hat{F}_I . The main result for this part stated as follows.

Proposition 6.8. *One has*

$$\hat{F}_I(D_m \tilde{L}) = (\log S(B))^I E_{1,m+1}, \quad (6.5)$$

where $(\cdot)^I$ denotes the coefficient of a tensor series over \mathbb{C}^d with respect to the monomial $e_{i_1} \otimes \cdots \otimes e_{i_m}$.

Our proof of Proposition 6.8, which has a combinatorial nature, reveals a surprising connection between two very different Lie structures ($\mathfrak{sl}_{m+1}(\mathbb{C})$ and free Lie algebra). As we will see, they are connected through a shuffle product relation. We first introduce some notation that is needed for the proof and then establish two key lemmas connecting the two sides with a specific shuffle product relation. The identity (6.5) will thus follow easily.

Some notation

(i) Suppose that (w_1, \dots, w_m) and $(w_{m+1}, \dots, w_{m+n})$ are two given words. We define their *shuffle product* by

$$(w_1, \dots, w_m) \sqcup (w_{m+1}, \dots, w_{m+n}) \triangleq \sum_{\sigma \in \mathcal{P}(m,n)} (w_{\sigma^{-1}(1)}, \dots, w_{\sigma^{-1}(m+n)}).$$

By definition, it is obvious that

$$a\mathbf{w} \sqcup b\mathbf{u} = a(\mathbf{w} \sqcup b\mathbf{u}) + b(a\mathbf{w} \sqcup \mathbf{u}) \quad (6.6)$$

for all letters a, b and words \mathbf{w}, \mathbf{u} .

(ii) Let $a_j \in \mathbb{C}$ be given numbers. The notation $(e^{a_p}, \dots, e^{a_q})$ simply represents a word and for given times s_j we define

$$(e^{a_p}, \dots, e^{a_q})(s_p, \dots, s_q) \triangleq \exp(a_p x_{s_p} + \dots + a_q x_{s_q}).$$

The tensor product between two words is simply defined by concatenation. The Lie bracket $[e^a, e^b]$ is defined by

$$[e^a, e^b] \triangleq (e^a, e^b) - (e^b, e^a).$$

When acting on a pair of times (s, t) , one has

$$[e^a, e^b](s, t) \triangleq (e^a, e^b)(s, t) - (e^b, e^a)(s, t) = e^{ax_s + bx_t} - e^{bx_s + ax_t}.$$

(iii) Let $a_1, \dots, a_m \in \mathbb{C}$ be given fixed. We introduce the word notation

$$(e^{a_p}, \overset{\nearrow}{\dots}, e^{a_q}) \triangleq \begin{cases} (e^{a_p}, e^{a_{p+1}}, \dots, e^{a_q}), & \text{if } p \leq q; \\ \mathbf{1}, & \text{if } p = q + 1; \\ \mathbf{0}, & \text{if } p > q, \end{cases}$$

and

$$(e^{a_p}, \overset{\searrow}{\dots}, e^{a_q}) \triangleq \begin{cases} (e^{a_p}, e^{a_{p-1}}, \dots, e^{a_q}), & \text{if } p \geq q; \\ \mathbf{1}, & \text{if } p = q - 1; \\ \mathbf{0}, & \text{if } p < q. \end{cases}$$

Here $\mathbf{1}$ means the multiplicative unit (for both concatenation and shuffle product) and $\mathbf{0}$ means the empty word whose (tensor or shuffle) product with any other word is still $\mathbf{0}$. In other words, conventionally one has

$$\mathbf{1} \sqcup \mathbf{w} = \mathbf{w} \sqcup \mathbf{1} = \mathbf{w}, \quad \mathbf{0} \sqcup \mathbf{w} = \mathbf{w} \sqcup \mathbf{0} = \mathbf{0}.$$

Expression of $\hat{F}_I(D_m\tilde{L})$ in terms of a shuffle identity

The first key lemma is related to computing the development of a generic term in the formula (6.1) for $D_m\tilde{L}$. Recall that the development F_I is defined by (6.3) where A is chosen according to (6.4) and $a_j = p_{i_j}$.

Lemma 6.9. *For any $k = 1, \dots, m-1$, one has*

$$\begin{aligned} & [e^{x_{t_{\sigma(m-k)}} \text{ad}A}(F_I(e_2)), [\dots [e^{x_{t_{\sigma(m-1)}} \text{ad}A}(F_I(e_2)), e^{x_{t_{\sigma(m)}} \text{ad}A}(F_I(e_2))]]]] \\ &= \sum_j \sum_{i=1}^{m-1} (-1)^{j+k-i-1} (((e^{a_j}, \overset{\nearrow}{\dots}, e^{a_{i-1}}) \sqcup (e^{a_{j+k}}, \overset{\searrow}{\dots}, e^{a_{i+2}})) \\ & \quad \otimes [e^{a_i}, e^{a_{i+1}}])(t_{\sigma(m-k)}, \dots, t_{\sigma(m)}) E_{j,j+k+1}. \end{aligned} \quad (6.7)$$

Here the Lie bracket is taken in $\mathfrak{sl}_{m+1}(\mathbb{C})$ and we adopt the convention that $E_{p,q} = 0$ if $p, q \leq 0$ or $\geq m+2$.

Proof. We prove the claim by induction on k . Let us denote the right hand side of (6.7) by $G(k)$. For the base step $k = 1$, by the definition of F_I one has

$$\begin{aligned} & [e^{x_{t_{\sigma(m-1)}} \text{ad}A}(F_I(e_2)), e^{x_{t_{\sigma(m)}} \text{ad}A}(F_I(e_2))] \\ &= [e^{x_{t_{\sigma(m-1)}} \text{ad}A}(E_{12} + \dots + E_{m,m+1}), e^{x_{t_{\sigma(m)}} \text{ad}A}(E_{12} + \dots + E_{m,m+1})] \\ &= \left[\sum_{j=1}^m e^{a_j x_{t_{\sigma(m-1)}}} E_{j,j+1}, \sum_{k=1}^m e^{a_k x_{t_{\sigma(m)}}} E_{k,k+1} \right]. \end{aligned} \quad (6.8)$$

By using the explicit commutator relation

$$[E_{ij}, E_{kl}] = \delta_{jk} E_{il} - \delta_{il} E_{kj}, \quad (6.9)$$

one easily finds that the right hand side of (6.8) equals

$$\begin{aligned} & \sum_{i=1}^{m-1} (e^{a_i x_{t_{\sigma(m-1)}}} e^{a_{i+1} x_{t_{\sigma(m)}}} - e^{a_{i+1} x_{t_{\sigma(m-1)}}} e^{a_i x_{t_{\sigma(m)}}}) E_{i,i+2} \\ &= \sum_{i=1}^{m-1} [e^{a_i}, e^{a_{i+1}}](t_{\sigma(m-1)}, t_{\sigma(m)}) E_{i,i+2} = G(1). \end{aligned}$$

This concludes the base step.

Now suppose that the claim is true for $k \leq n$. By using the induction hypothesis,

one can write

$$\begin{aligned}
& \left[e^{x_{t_{\sigma(m-n-1)} \text{ad} A}} F_I(e_2), \left[e^{x_{t_{\sigma(m-n)} \text{ad} A}} (F_I(e_2)) \right. \right. \\
& \quad \left. \left. \dots \left[e^{x_{t_{\sigma(m-1)} \text{ad} A}} (F_I(e_2)), e^{x_{t_{\sigma(m)} \text{ad} A}} (F_I(e_2)) \right] \right] \right] \\
&= \sum_j \sum_{i=1}^{m-1} (-1)^{j+n-i-1} \left[e^{a_1 x_{t_{\sigma(m-n-1)}}} E_{12} + \dots + e^{a_m x_{t_{\sigma(m-n-1)}}} E_{m,m+1}, \right. \\
& \quad \left((e^{a_j}, \overset{\nearrow}{\dots}, e^{a_{i-1}}) \sqcup (e^{a_{j+n}}, \overset{\searrow}{\dots}, e^{a_{i+2}}) \right) \otimes [e^{a_i}, e^{a_{i+1}}] \\
& \quad \left. (t_{\sigma(m-n)}, \dots, t_{\sigma(m)}) E_{j,j+n+1} \right].
\end{aligned}$$

Again by applying the commutator relation (6.9) and adjusting the j -index, the above expression is equal to

$$\begin{aligned}
& \sum_j \sum_{i=1}^{m-1} (-1)^{j+n-i} \left(e^{a_j} \otimes \left((e^{a_{j+1}}, \overset{\nearrow}{\dots}, e^{a_{i-1}}) \sqcup (e^{a_{j+n+1}}, \overset{\searrow}{\dots}, e^{a_{i+2}}) \right) \otimes [e^{a_i}, e^{a_{i+1}}] \right. \\
& \quad \left. + (e^{a_{j+n+1}} \otimes \left((e^{a_j}, \overset{\nearrow}{\dots}, e^{a_{i-1}}) \sqcup (e^{a_{j+n}}, \overset{\searrow}{\dots}, e^{a_{i+2}}) \right) \otimes [e^{a_i}, e^{a_{i+1}}] \right) \\
& \quad (t_{\sigma(m-n-1)}, \dots, t_{\sigma(m)}) E_{j,j+n+2} \\
&= \sum_j \sum_{i=1}^{m-1} (-1)^{j+n-i} \left((e^{a_j}, \overset{\nearrow}{\dots}, e^{a_{i-1}}) \sqcup (e^{a_{j+n+1}}, \overset{\searrow}{\dots}, e^{a_{i+2}}) \right) \otimes [e^{a_i}, e^{a_{i+1}}] \\
& \quad (t_{\sigma(m-n-1)}, \dots, t_{\sigma(m)}) E_{j,j+n+2} = G(n+1),
\end{aligned}$$

where the first equality follows from the relation (6.6). This completes the induction step. \square

Computation of $(\log S(B))^l$

Before stating the second key lemma, we derive a general formula for basic Lie elements which may be of independent interest. We consider the tensor algebra $T((\mathbb{C}^d))$. Let $J = (j_1, \dots, j_m)$ be a given word. We define

$$e_{[J]} \triangleq [e_{j_1}, [e_{j_2}, \dots [e_{j_{m-1}}, e_{j_m}]]], \quad e_J \triangleq e_{j_1} \otimes \dots \otimes e_{j_m}$$

respectively. Given a word $K = (k_1, \dots, k_r) \subseteq (1, \dots, m)$, we denote

$$J_K \triangleq (j_{k_1}, \dots, j_{k_r}), \quad \overleftarrow{J}_K \triangleq (j_{k_r}, \dots, j_{k_1}).$$

As a convention, we also set $e_\emptyset \triangleq \mathbf{1}$ and $\overleftarrow{\emptyset} \triangleq \emptyset$.

Lemma 6.10. *For any word $J = (j_1, \dots, j_m)$, one has*

$$e_{[J]} = \sum_{K \subseteq (1, \dots, m-2)} (-1)^{|K|} e_{J \setminus (J_K, j_{m-1}, j_m)} \otimes [e_{j_{m-1}}, e_{j_m}] \otimes e_{\overleftarrow{J}_K}. \quad (6.10)$$

Proof. We prove the claim by induction. The case when $m = 2$ is obvious. Suppose that the claim is true for any word with length $\leq m - 1$ and let $J = (j_1, \dots, j_m)$. By the induction hypothesis, one has

$$\begin{aligned}
e_{[J]} &= [e_{j_1}, e_{[J']}] \quad (J' \triangleq J \setminus \{j_1\}) \\
&= \sum_{K \subseteq (2, \dots, m-2)} (-1)^{|K|} e_{(j_1, J') \setminus (J'_K, j_{m-1}, j_m)} \otimes [e_{j_{m-1}}, e_{j_m}] \otimes e_{\overleftarrow{J'_K}} \\
&\quad - \sum_{K \subseteq (2, \dots, m-2)} (-1)^{|K|} e_{J' \setminus (J'_K, j_{m-1}, j_m)} \otimes [e_{j_{m-1}}, e_{j_m}] \otimes e_{\overleftarrow{(j_1, J'_K)}} \\
&= \sum_{L \subseteq (1, \dots, m-2)} (-1)^{|L|} e_{J \setminus (J_L, j_{m-1}, j_m)} \otimes [e_{j_{m-1}}, e_{j_m}] \otimes e_{\overleftarrow{J_L}}
\end{aligned}$$

The last equality follows from the observation that the two sums in the second last equality correspond to the cases $1 \notin L$ and $1 \in L$ respectively. \square

We are now able to establish the second key lemma for the proof of Proposition 6.8. This connects the logarithmic signature coefficient $(\log S(B))^I$ with the same kind of shuffle product identity appearing in Lemma 6.9. Recall that B_t is the path in \mathbb{C}^d defined by (6.2) with given $p_1, \dots, p_d \in \mathbb{C}$. To ease notation, we write

$$dB_{t_\sigma}^J \triangleq dB_{t_{\sigma(1)}}^{j_1} \cdots dB_{t_{\sigma(m)}}^{j_m}.$$

We denote $\pi_I : T((\mathbb{C}^d)) \rightarrow \mathbb{C}$ as the projection map which extracts the coefficient of a tensor with respect to e_I .

Lemma 6.11. *For any word $I = (i_1, \dots, i_m)$, one has*

$$\begin{aligned}
&\pi_I \left(\sum_{J: |J|=m} \int_{\Delta_m} dB_{t_\sigma}^J e_{[J]} \right) \\
&= \sum_{i=1}^{m-1} (-1)^{m-1-i} \int_{\Delta_m} \left(((e^{a_1}, \overset{\nearrow}{\cdots}, e^{a_{i-1}}) \sqcup (e^{a_m}, \overset{\searrow}{\cdots}, e^{a_{i+2}})) \right. \\
&\quad \left. \otimes [e^{a_i}, e^{a_{i+1}}] \right) (t_{\sigma(1)}, \dots, t_{\sigma(m)}) dy_{t_1} \cdots dy_{t_m}, \tag{6.11}
\end{aligned}$$

where $a_j \triangleq p_{i_j}$ ($j = 1, \dots, m$).

Proof. According to Lemma 6.10, one has

$$\pi_I(e_{[J]})e_I = \sum_{K \subseteq (1, \dots, m-2)} (-1)^{|K|} \pi_I(e_{J \setminus (J_K, j_{m-1}, j_m)} \otimes [e_{j_{m-1}}, e_{j_m}] \otimes e_{\overleftarrow{J_K}}). \tag{6.12}$$

Let $K = (k_1, \dots, k_r) \subseteq (1, \dots, m-2)$ be given fixed. For any word J with length m , one has

$$\pi_I(e_{J \setminus (J_K, j_{m-1}, j_m)} \otimes [e_{j_{m-1}}, e_{j_m}] \otimes e_{\overleftarrow{J_K}}) \neq 0$$

if and only if

$$(J \setminus (J_K, j_{m-1}, j_m), j_{m-1}, j_m, \overleftarrow{J}_K) = I$$

or

$$(J \setminus (J_K, j_{m-1}, j_m), j_m, j_{m-1}, \overleftarrow{J}_K) = I.$$

The first (respectively, second) case comes with a plus (respectively, minus) sign arising from the Lie bracket. As a result, the word J in the first case is uniquely determined by

$$J = \begin{pmatrix} \cdots & i_m & \cdots & i_{m-1} & \cdots & i_{m-r+1} & \cdots & i_{m-r-1} & i_{m-r} \\ \cdots & k_1 & \cdots & k_2 & \cdots & k_r & \cdots & m-1 & m \end{pmatrix} \quad (6.13)$$

where the “ \cdots ” positions are filled by (i_1, \dots, i_{m-r-2}) in its natural order. Respectively, the word J in the second case is determined by swapping i_{m-r-1}, i_{m-r} in the last two positions in (6.13). We denote these two uniquely determined words by $J_1(K)$ and $J_2(K)$ respectively.

It follows from the above computation that

$$\sum_{J:|J|=m} \int_{\Delta_m} dB_{\mathbf{t}_\sigma}^J \pi_I(e_{[J]}) = \sum_{K \subseteq \{1, \dots, m-2\}} (-1)^{|K|} \int_{\Delta_m} (dB_{\mathbf{t}_\sigma}^{J_1(K)} - dB_{\mathbf{t}_\sigma}^{J_2(K)}).$$

From the explicit shapes of $J_1(K), J_2(K)$, it is not hard to see that the above expression is precisely equal to

$$\begin{aligned} & \sum_{r=0}^{m-2} (-1)^r \int_{\Delta_m} (((e^{a_1}, \overset{\nearrow}{\dots}, e^{a_{m-r-2}}) \sqcup (e^{a_m}, \overset{\searrow}{\dots}, e^{a_{m-r+1}})) \\ & \quad \otimes [e^{a_{m-r-1}}, e^{a_{m-r}}])(t_{\sigma(1)}, \dots, t_{\sigma(m)}) dy_{t_1} \cdots dy_{t_m}. \end{aligned}$$

The desired relation follows from the change of indices $i \triangleq m-1-r$. \square

We are now in a position to finish the proof of Proposition 6.8.

Proof of Proposition 6.8. According to Proposition 6.7 and Lemma 6.9 with $k = m-1$, one has

$$\begin{aligned} \hat{F}_I(D_m \tilde{L}) &= \sum_{\sigma \in \mathcal{S}_m} \frac{(-1)^{e(\sigma)}}{m^2 \binom{m-1}{e(\sigma)}} \sum_{i=1}^{m-1} (-1)^{m-i-1} \\ & \quad \int_{\Delta_m} (((e^{a_1}, \overset{\nearrow}{\dots}, e^{a_{i-1}}) \sqcup (e^{a_m}, \overset{\searrow}{\dots}, e^{a_{i+2}})) \\ & \quad \otimes [e^{a_i}, e^{a_{i+1}}])(t_{\sigma(1)}, \dots, t_{\sigma(m)}) dy_{t_1} \cdots dy_{t_m} E_{1,m+1}. \end{aligned} \quad (6.14)$$

On the other hand, the Chen-Strichartz formula gives that

$$\log S(B) = \sum_{m=1}^{\infty} \sum_{\sigma \in \mathcal{S}_m} \frac{(-1)^{e(\sigma)}}{m^2 \binom{m-1}{e(\sigma)}} \sum_{J:|J|=m} \int_{\Delta_m} dB_{\mathbf{t}_\sigma}^J e_{[J]}. \quad (6.15)$$

It follows from (6.15) and Lemma 6.11 that the right hand side of (6.14) is precisely $(\log S(B))^I E_{1,m+1}$. This completes the proof of the proposition. \square

We conclude this part by presenting a general formula for the development of $D_k \tilde{L}$, which will be important to us later on. Recall that S_m^γ is the analytic function defined by Definition 6.2. Throughout the rest, we will omit the superscript γ for simplicity.

Corollary 6.12. *Let $a_1, \dots, a_m \in \mathbb{C}$ be given fixed numbers. Consider the $\mathfrak{sl}_{m+1}(\mathbb{C})$ -development defined by (6.3) where $A \in \mathfrak{h}$ is chosen to be such that $[A, E_{k,k+1}] = a_k E_{k,k+1}$ for all $k = 1, \dots, m$. Then one has*

$$\hat{F}(D_k \tilde{L}) = \sum_{j=1}^{m-k+1} S_k(a_j, \dots, a_{j+k-1}) E_{j,j+k} \quad (6.16)$$

for all $k = 1, \dots, m$.

Proof. The case when $k = m$ is just Proposition 6.7 applied to the path

$$B_t \triangleq \left(\int_0^t e^{a_1 x_s} dy_s, \dots, \int_0^t e^{a_m x_s} dy_s \right) \in \mathbb{C}^m$$

with the word $I = (1, \dots, m)$. For a general k , the same argument as the proof of Lemma 6.9 gives that

$$\begin{aligned} \hat{F}(D_k \tilde{L}) &= \sum_{\sigma \in \mathcal{S}_k} \frac{(-1)^{e(\sigma)}}{k^2 \binom{k-1}{e(\sigma)}} \sum_{j=1}^{m-k+1} \sum_{i=1}^{m-1} (-1)^{j+k-i} \\ &\quad \int_{\Delta_k} \left(((e^{a_j}, \overset{\nearrow}{\dots}, e^{a_{i-1}}) \sqcup (e^{a_{j+k-1}}, \overset{\searrow}{\dots}, e^{a_{i+2}})) \right. \\ &\quad \left. \otimes [e^{a_i}, e^{a_{i+1}}] \right) (t_{\sigma(1)}, \dots, t_{\sigma(k)}) dy_{t_1} \cdots dy_{t_k} E_{j,j+k}. \end{aligned} \quad (6.17)$$

In view of Lemma 6.11, after a change of indices $l \triangleq i - j + 1$ the right hand side of (6.17) is precisely

$$\sum_{j=1}^{m-k+1} S_k(a_j, \dots, a_{j+k-1}) E_{j,j+k}.$$

The relation (6.16) thus follows. \square

6.3 Proof of Theorem 6.5

Our proof of Theorem 6.5 is based on a key lemma regarding the analyticity of path developments for the Hausdorff series. Since its proof has a rather involved combinatorial nature, we decide to only state this analyticity lemma here and then use it to

complete the proof of Theorem 6.5. In Section 6.4 below, we will give the proof of this key lemma.

Recall that the n -th Hausdorff series H_n is defined by (3.12). In our context, we need to consider a more specific series which is defined in terms of the homogeneous projections of \tilde{L} .

Definition 6.13. Let $n \geq 1$ and $k_1, \dots, k_n \geq 1$ be given fixed. We define

$$H_n(k_1, \dots, k_n) \triangleq \frac{1}{n!} (H_1(k_1) \partial_{e_1}) \circ \dots \circ (H_1(k_n) \partial_{e_1})(e_1),$$

where we used the shorthand notation $H_1(k_i) \triangleq H_1(D_{k_i} \tilde{L}, e_1)$. It is clear that $H_n(k_1, \dots, k_n)$ depends only on the vector $K \triangleq (k_1, \dots, k_n)$ and we will thus also use the alternative shorthand notation H_K .

Remark 6.14. In the above definition, we view $D_{k_i} \tilde{L}$ as a fixed symbol and the derivation $H_1(k_j) \partial_{e_1}$ does not act on it. As a result, $H_n(k_1, \dots, k_n)$ is expressed as a formal Lie series over the $n+1$ independent symbols $\{e_1, D_{k_1} \tilde{L}, \dots, D_{k_n} \tilde{L}\}$. It is then regarded as a Lie series over \mathbb{R}^2 through the substitution $D_{k_i} \tilde{L} \in \mathcal{L}(\langle e_1, e_2 \rangle)$.

Lemma 6.15. Let $N \geq 2$ be a given fixed integer. Suppose that $S_n(a_1, \dots, a_n) = 0$ for all $n \leq N-1$ and all non-degenerate sequences (a_1, \dots, a_n) with $a_j \in 2\pi i\mathbb{Z}$. Let (b_1^*, \dots, b_N^*) be a fixed non-degenerate sequence with $b_j \in 2\pi i\mathbb{Z}$. Given $w \in \mathbb{C}$, we define the path development $F_w : \mathbb{C}^2 \rightarrow \mathfrak{sl}_{N+1}(\mathbb{C})$ by

$$F_w(e_1) \triangleq A_w, \quad F_w(e_2) \triangleq E_{12} + \dots + E_{N,N+1}, \quad (6.18)$$

where $A_w \in \mathfrak{h}$ is chosen to satisfy

$$[A_w, E_{k,k+1}] = (b_k^* + w) E_{k,k+1} \quad (k = 1, \dots, N). \quad (6.19)$$

Then for any vector $K = (k_1, \dots, k_n)$ satisfying $|K| = N, n \geq 2$ or $|K| \leq N-1, n \geq 1$ and any $j = 1, \dots, N$, there exists a meromorphic function $\Psi_{K,j} : \mathbb{C}^{|K|} \rightarrow \mathbb{C}$ (i.e. the quotient of two holomorphic functions) in $|K|$ complex variables, such that

$$\hat{F}_w(H_K) = \sum_j \Psi_{K,j}(b_j^* + w, \dots, b_{j+|K|-1}^* + w) E_{j,j+|K|}$$

and the function $w \mapsto \Psi_{K,j}(b_j^* + w, \dots, b_{j+|K|-1}^* + w)$ is analytic in a neighbourhood of $w = 0$ for all (K, j) . Here \hat{F}_w is the induced Lie homomorphism on $\mathcal{L}(\mathbb{C}^2)$ (cf. (3.15)).

Now we prove Theorem 6.5 presuming the correctness of Lemma 6.15.

Proof of Theorem 6.5. We argue by induction on the length m of the sequence (a_1, \dots, a_m) . The cases when $m = 1, 2$ are just Theorem 3.1 and Corollary 5.2 respectively. Suppose that the claim is true for all non-degenerate sequences with length $\leq N-1$. Let

(b_1^*, \dots, b_N^*) be a non-degenerate sequence with $b_j^* \in 2\pi i\mathbb{Z}$ for all j . Our goal is to show that $S_N(b_1^*, \dots, b_N^*) = 0$.

First of all, by using the degree N version of (5.7) one finds that

$$D_N L = D_N \left(\sum_{n=1}^{\infty} H_n(\tilde{L}, e_1) \right) = H_1(N) + \sum_{n=2}^N \sum_{k_1 + \dots + k_n = N} H_n(k_1, \dots, k_n). \quad (6.20)$$

Given $w \in \mathbb{C}$, let $F_w : \mathbb{C}^2 \rightarrow \mathfrak{sl}_{N+1}(\mathbb{C})$ denote the path development defined by (6.18) under the condition (6.19). By applying \hat{F}_w to both sides of (6.20), one finds that

$$\begin{aligned} \hat{F}_w(D_N L) &= H_1(\hat{F}_w(D_N \tilde{L}), A) + \sum_{n=2}^N \sum_{k_1 + \dots + k_n = N} \hat{F}_w(H_n(k_1, \dots, k_n)) \\ &= S_N(b_1^* + w, \dots, b_N^* + w) \phi(b_1^* + \dots + b_N^* + Nw) E_{1, N+1} \\ &\quad + \sum_{n=2}^N \sum_{k_1 + \dots + k_n = N} \hat{F}_w(H_n(k_1, \dots, k_n)). \end{aligned}$$

The second equality follows from Corollary 6.12 and the definition (3.11) of H_1 , where $\phi(z) \triangleq \frac{z}{e^z - 1}$. Since the induction hypothesis holds for all non-degenerate sequences with length $< N$, one can apply Lemma 6.15 to conclude that the function

$$w \mapsto \hat{F}_w(H_n(k_1, \dots, k_n))$$

is analytic near $w = 0$. Note that $w \mapsto \hat{F}_w(D_N L)$ is also an analytic function (on the whole space \mathbb{C} due to the infinite R.O.C. for $D_N L$). It follows that the function

$$w \mapsto S_N(b_1^* + w, \dots, b_N^* + w) \phi(b_1^* + \dots + b_N^* + Nw)$$

must be analytic near $w = 0$. Recall that $b^* \triangleq b_1^* + \dots + b_N^*$ is a nonzero integer multiple of $2\pi i$ and is thus a pole of ϕ . As a consequence, the function $w \mapsto S_N(b_1^* + w, \dots, b_N^* + w)$ must vanish at $w = 0$. In other words, one has $S_N(b_1^*, \dots, b_N^*) = 0$, which completes the induction step. \square

6.4 The analyticity lemma

It remains to prove the analyticity lemma which will be the main task of this subsection.

6.4.1 A recursive formula for the Hausdorff series

Our strategy relies on induction on the total degree $|K|$ of the vector $K = (k_1, \dots, k_n)$. For this purpose, we shall derive a recursive formula for computing $H_n(k_1, \dots, k_n)$ in terms of ‘‘symmetrised products’’ of H_m ’s ($m < n$). We first define such type of products.

Definition 6.16. Let A_1, \dots, A_r, e_1, v be given symbols. We define the linear operator

$$A_1 \partial_{e_1} \hat{\otimes}_s \cdots \hat{\otimes}_s A_r \partial_{e_1} : T(\langle e_1, v \rangle) \rightarrow T(\langle e_1, v, A_1, \dots, A_r \rangle)$$

in the following way. Let $\mathbf{f}_1 \otimes \cdots \otimes \mathbf{f}_N$ be a given monomial where $\mathbf{f}_i = e_1$ or v . We first define

$$\begin{aligned} & A_1 \partial_{e_1} \hat{\otimes} \cdots \hat{\otimes} A_r \partial_{e_1} (\mathbf{f}_1 \otimes \cdots \otimes \mathbf{f}_N) \\ & \triangleq \sum \mathbf{f}_1 \otimes \cdots \otimes \mathbf{f}_{i_1-1} \otimes A_1 \otimes \mathbf{f}_{i_1+1} \otimes \cdots \otimes \mathbf{f}_{i_r-1} \otimes A_r \otimes \mathbf{f}_{i_r+1} \otimes \cdots \otimes \mathbf{f}_N, \end{aligned}$$

where the summation is taken over all r -subsets $\{i_1, \dots, i_r\}$ with $\mathbf{f}_{i_j} = e_1$. In other words, the action is given by replacing r e_1 's by (A_1, \dots, A_r) and summing over all such possibilities. The definition is extended to the tensor algebra $T(\langle e_1, v \rangle)$ by linearity. We then set

$$A_1 \partial_{e_1} \hat{\otimes}_s \cdots \hat{\otimes}_s A_r \partial_{e_1} \triangleq \sum_{\sigma \in \mathcal{S}_r} A_{\sigma(1)} \partial_{e_1} \hat{\otimes} \cdots \hat{\otimes} A_{\sigma(r)} \partial_{e_1}.$$

The following simple property of the operator $A_1 \partial_{e_1} \hat{\otimes}_s \cdots \hat{\otimes}_s A_r \partial_{e_1}$ will be useful to us.

Lemma 6.17. *One has*

$$\begin{aligned} & A_1 \partial_{e_1} \hat{\otimes}_s \cdots \hat{\otimes}_s A_r \partial_{e_1} (\text{ad}_{e_1}^l(v)) \\ & = \sum_{\sigma \in \mathcal{S}_r} \sum_{\xi_1 + \cdots + \xi_{r+1} = l - r} \text{ad}_{e_1}^{\xi_1} \circ \text{ad}_{A_{\sigma(1)}} \circ \text{ad}_{e_1}^{\xi_2} \circ \cdots \circ \text{ad}_{e_1}^{\xi_r} \circ \text{ad}_{A_{\sigma(r)}} \circ \text{ad}_{e_1}^{\xi_{r+1}}(v) \end{aligned}$$

for any $l \geq r \geq 1$.

Proof. This follows by induction on r based on Definition 6.16 (one first treats the case when $l = r$ by induction and then the case when $l \geq r$ follows by another step of induction from cases $(l-1, r)$ and $(l-1, r-1)$). \square

To get some feeling about the shape of $H_n(k_1, \dots, k_n)$, we first look at the small- n cases.

Example 6.18. We have seen the computation for H_2 in (5.11). Let us consider the case when $n = 3$. By definition, one has

$$\begin{aligned} & H_3(k_1, k_2, k_3) \\ & = \frac{1}{3!} (H_1(k_1) \partial_{e_1}) \left(\sum_l \frac{B_l}{l!} (H_1(k_2) \partial_{e_1}) (\text{ad}_{e_1}^l(D_{k_3} \tilde{L})) \right) \\ & = \frac{1}{3!} (H_1(k_1) \partial_{e_1}) \left(\sum_l \frac{B_l}{l!} \sum_{\xi_1 + \xi_2 = l-1} \text{ad}_{e_1}^{\xi_1} \circ \text{ad}_{H_1(k_2)} \circ \text{ad}_{e_1}^{\xi_2} (D_{k_3} \tilde{L}) \right) \\ & = \frac{1}{3!} \sum_{\tau \in \mathcal{S}_2} \sum_l \frac{B_l}{l!} \sum_{\xi_1 + \xi_2 + \xi_3 = l-2} \text{ad}_{e_1}^{\xi_1} \circ \text{ad}_{H_1(k_{\tau(1)})} \circ \text{ad}_{e_1}^{\xi_2} \circ \text{ad}_{H_1(k_{\tau(2)})} \circ \text{ad}_{e_1}^{\xi_3} (D_{k_3} \tilde{L}) \\ & \quad + \frac{2!}{3!} \sum_l \frac{B_l}{l!} \sum_{\xi_1 + \xi_2 = l-1} \text{ad}_{e_1}^{\xi_1} \circ \text{ad}_{H_2(k_1, k_2)} \circ \text{ad}_{e_1}^{\xi_2} (D_{k_3} \tilde{L}). \end{aligned}$$

A crucial observation is that the above expression can be rewritten in terms of the symmetrised tensor product in Definition 6.16. In fact, according to Lemma 6.17 one has

$$H_3(k_1, k_2, k_3) = \frac{1}{3!} \left((H_1(k_1)\partial_{e_1}) \hat{\otimes}_s (H_1(k_2)\partial_{e_1}) \right) (H_1(k_3)) \\ + \frac{2!}{3!} (H_2(k_1, k_2)\partial_{e_1}) (H_1(k_3)).$$

Our aim is to generalise the above formula to arbitrary $H_n(k_1, \dots, k_n)$.

To state the recursive formula for general $H_n(k_1, \dots, k_n)$, we need to introduce one more definition to ease notation.

Definition 6.19. Let $P = \{I_1, \dots, I_r\}$ be a given (unordered) partition of $\{1, \dots, n\}$ (i.e. $I_p \neq \emptyset$, $I_p \cap I_q = \emptyset$ for $p \neq q$ and $\cup I_p = \{1, \dots, n\}$). We define

$$\hat{H}_P(k_1, \dots, k_n) \triangleq H_{K_1} \partial_{e_1} \hat{\otimes}_s \dots \hat{\otimes}_s H_{K_r} \partial_{e_1},$$

where $I_p = (i_1^p < \dots < i_{l_p}^p)$ and $K_p \triangleq (k_{i_1^p}, \dots, k_{i_{l_p}^p})$ ($p = 1, \dots, r$).

The main recursive formula for $H_n(k_1, \dots, k_n)$ is stated as follows.

Proposition 6.20. For any $n \geq 1$ and $k_1, \dots, k_n \geq 1$, one has

$$H_n(k_1, \dots, k_n) = \sum_{P=\{I_1, \dots, I_r\}} \frac{l_1! \dots l_r!}{n!} \hat{H}_P(k_1, \dots, k_{n-1}) (H_1(k_n)),$$

where the summation is taken over all (unordered) partitions $\{I_1, \dots, I_r\}$ of the set $\{1, \dots, n-1\}$ and l_s denotes the cardinality of I_s ($s = 1, \dots, r$).

Proof. We prove the claim by induction on n . To simplify notation, we get rid of the factorials by setting $\bar{H}_n \triangleq n! H_n$ and $\hat{\bar{H}}_P \triangleq l_1! \dots l_r! \hat{H}_P$. The base case $n = 1$ is clear. Suppose that the claim is true for \bar{H}_{n-1} . By the definition of $H_n(k_1, \dots, k_n)$ and the induction hypothesis, one has

$$\bar{H}_n(k_1, \dots, k_n) = (H_1(k_1)\partial_{e_1}) (\bar{H}_{n-1}(k_2, \dots, k_n)) \\ = (H_1(k_1)\partial_{e_1}) \left[\sum_P (\hat{\bar{H}}_P(k_2, \dots, k_{n-1}) (H_1(k_n))) \right], \quad (6.21)$$

where the summation is taken over all partitions of $\{2, \dots, n-1\}$.

Given any such partition $P = \{I_1, \dots, I_r\}$, by using Lemma 6.17 (and the notation in Definition 6.19) one can write

$$\hat{\bar{H}}_P(k_2, \dots, k_{n-1}) (H_1(k_n)) \\ = \sum_{\tau \in \mathcal{S}_r} \sum_l \frac{B_l}{l!} \sum_{\xi_1 + \dots + \xi_{r+1} = l-r} \text{ad}_{e_1}^{\xi_1} \circ \text{ad}_{\bar{H}_{K_{\tau(1)}}} \circ \text{ad}_{e_1}^{\xi_2} \circ \\ \dots \circ \text{ad}_{\bar{H}_{K_{\tau(r)}}} \circ \text{ad}_{e_1}^{\xi_{r+1}} (D_{k_n} \tilde{L}).$$

As a derivation, the outer $H_1(k_1)\partial_{e_1}$ action in (6.21) will either apply to the ad_{e_1} 's or to the $\bar{H}_{K_{\tau(i)}}$'s. The former case yields

$$\hat{H}_{\{1\} \oplus P}(k_1, \dots, k_{n-1})(H_1(k_n))$$

where $\{1\} \oplus P$ is the partition $\{\{1\}, I_1, \dots, I_r\}$ of $\{1, \dots, n-1\}$, while the latter case yields

$$\sum_{s=1}^r \hat{H}_{\{I_1, \dots, \{1\} \cup I_s, \dots, I_r\}}(k_1, \dots, k_{n-1})(H_1(k_n))$$

where $\{I_1, \dots, \{1\} \cup I_s, \dots, I_r\}$ is the partition of $\{1, \dots, n-1\}$ obtained by adding the element "1" into I_s . As a consequence, one finds that

$$\begin{aligned} \bar{H}_n(k_1, \dots, k_n) &= \sum_{P=\{I_1, \dots, I_r\}} [\hat{H}_{\{1\} \oplus P}(k_1, \dots, k_{n-1})(H_1(k_n))] \\ &\quad + \sum_{s=1}^r \hat{H}_{\{I_1, \dots, \{1\} \cup I_s, \dots, I_r\}}(k_1, \dots, k_{n-1})(H_1(k_n)). \end{aligned}$$

The right hand side is exactly the sum of $\hat{H}_Q(k_1, \dots, k_{n-1})(H_1(k_n))$ over all possible partitions of $\{1, \dots, n-1\}$. Indeed, any such partition Q corresponds to a partition P of $\{2, \dots, n-1\}$ together with a specific way of adding the element "1" (either outside P or into one of the members of P) and vice versa. This completes the induction step. \square

6.4.2 A combinatorial identity

The following combinatorial identity plays a key role in the proof. It allows one to express certain quotients defined by non-consecutive sums of a sequence in terms of consecutive sums.

Lemma 6.21. *Let $0 \leq s \leq R$ be given fixed and let $\{c_k\}_{1 \leq k \leq R+1}$ be a given sequence of numbers. Then one has*

$$\sum_{\eta \in \mathcal{T}(s, R-s)} \frac{1}{c_{\eta^{-1}(2)} c_{\eta^{-1}(3)} \cdots c_{\eta^{-1}(R+1)}} = \left(\prod_{k=1}^s c_k^s \right)^{-1} \left(\prod_{k=s+2}^{R+1} c_{s+2}^k \right)^{-1}. \quad (6.22)$$

Here $\mathcal{T}(s, R-s)$ denotes the collection of permutations $\eta \in \mathcal{S}_{R+1}$ such that

$$\eta(s+1) = 1; \quad \eta(s) < \cdots < \eta(1); \quad \eta(s+2) < \cdots < \eta(R+1). \quad (6.23)$$

We also denote

$$c_{\eta^{-1}(2)}^{\eta^{-1}(k)} \triangleq c_{\eta^{-1}(2)} + c_{\eta^{-1}(3)} + \cdots + c_{\eta^{-1}(k)}, \quad c_p^q \triangleq c_p + c_{p+1} + \cdots + c_q, \quad (6.24)$$

and conventionally we set $c_1^0 = c_{R+2}^{R+1} = 1$.

Proof. The case when $R = 2, 3$ can be checked explicitly. Suppose that the claim is true for $R - 1$. Let $0 \leq s \leq R$ and $\{c_k\}_{1 \leq k \leq R+1}$ be given. By the definition of $\eta \in \mathcal{T}(s, R - s)$, either $\eta(1) = R + 1$ or $\eta(R + 1) = R + 1$. One can thus write

$$\sum_{\eta \in \mathcal{T}(s, R-s)} \frac{1}{c_{\eta^{-1}(2)}^{\eta^{-1}(2)} c_{\eta^{-1}(3)}^{\eta^{-1}(3)} \cdots c_{\eta^{-1}(R+1)}^{\eta^{-1}(R+1)}} = \sum_{\eta: \eta(1)=R+1} + \sum_{\eta: \eta(R+1)=R+1}.$$

According to the induction hypothesis,

$$\sum_{\eta: \eta(1)=R+1} \frac{1}{c_{\eta^{-1}(2)}^{\eta^{-1}(2)} c_{\eta^{-1}(3)}^{\eta^{-1}(3)} \cdots c_{\eta^{-1}(R+1)}^{\eta^{-1}(R+1)}} = \frac{1}{(c_1^s + c_{s+2}^{R+1})} \times \left(\prod_{k=2}^s c_k^s \right)^{-1} \left(\prod_{k=s+2}^{R+1} c_{s+2}^k \right)^{-1}$$

and similarly

$$\sum_{\eta: \eta(R+1)=R+1} \frac{1}{c_{\eta^{-1}(2)}^{\eta^{-1}(2)} c_{\eta^{-1}(3)}^{\eta^{-1}(3)} \cdots c_{\eta^{-1}(R+1)}^{\eta^{-1}(R+1)}} = \frac{1}{(c_1^s + c_{s+2}^{R+1})} \times \left(\prod_{k=1}^s c_k^s \right)^{-1} \left(\prod_{k=s+2}^R c_{s+2}^k \right)^{-1}.$$

It is easily seen that the sum of the above two expressions is equal to the right hand side of (6.22). \square

6.4.3 Proof of Lemma 6.15

We are now in a position to develop the precise proof of Lemma 6.15. We first recall the standing assumptions of the lemma which will be imposed throughout the rest of this subsection.

Assumption (\mathbf{A}_{N-1}). Let $N \geq 2$ be given fixed. We assume that $S_m(a_1, \dots, a_m) = 0$ for all $m \leq N - 1$ and all non-degenerate sequences (a_1, \dots, a_m) with $a_j \in 2\pi i\mathbb{Z}$.

Now let (b_1^*, \dots, b_N^*) be a given fixed non-degenerate sequence with $b_j^* \in 2\pi i\mathbb{Z}$. For each $w \in \mathbb{C}$, we consider the path development F_w into $\mathfrak{sl}_{N+1}(\mathbb{C})$ defined by (6.18, 6.19). Recall that \hat{F}_w is the induced Lie homomorphism \hat{F}_w on $\mathcal{L}((\mathbb{C}^2))$.

We are going to prove Lemma 6.15 by induction on the degree $|K| = k_1 + \dots + k_n$ of the vector $K = (k_1, \dots, k_n)$. To be precise, we assume as the induction hypothesis that for any vector K with degree $|K| < k \leq N$ and $j = 1, \dots, N$, there exists a meromorphic function $\Psi_{K,j}$ in $|K|$ variables such that

$$\hat{F}_w(H_K) = \sum_j \Psi_{K,j}(b_j^* + w, \dots, b_{j+|K|-1}^* + w) E_{j,j+|K|}. \quad (6.25)$$

In addition, the function

$$w \mapsto \Psi_{K,j}(b_j^* + w, \dots, b_{j+|K|-1}^* + w)$$

is analytic near $w = 0$. We want to prove the same assertion for any vector $K = (k_1, \dots, k_n)$ of degree $|K| = k$. Here we should emphasise that $n \geq 2$ if $k = N$ and $n \geq 1$ if $k < N$.

We first consider the easier case when $n = 1$ and $k < N$. According to Corollary 6.12 and the H_1 -formula (3.11), one has

$$\begin{aligned}\hat{F}_w(H_1(k)) &= \sum_l \frac{B_l}{l!} \text{ad}_A^l(\hat{F}_w(D_k \tilde{L})) \\ &= \sum_j S_k(b_j^* + w, \dots, b_{j+k-1}^* + w) \sum_l \frac{B_l}{l!} (b_j^* + \dots + b_{j+k-1}^* + kw)^l E_{j,j+k} \\ &= \sum_j S_k(b_j^* + w, \dots, b_{j+k-1}^* + w) \phi(b_j^* + \dots + b_{j+k-1}^* + kw) E_{j,j+k}.\end{aligned}$$

Since $k < N$, one knows from Assumption (\mathbf{A}_{N-1}) that $S_k(b_j^*, \dots, b_{j+k-1}^*) = 0$. On the other hand, $w = 0$ is a simple pole of the function $\phi(b_j^* + \dots + b_{j+k-1}^* + kw)$. As a result, the function

$$\Psi_{(k),j}(w) \triangleq S_k(b_j^* + w, \dots, b_{j+k-1}^* + w) \phi(b_j^* + \dots + b_{j+k-1}^* + kw)$$

is analytic at $w = 0$. This proves the assertion in the current case.

We now focus on the more difficult case when $n \geq 2$ ($k \leq N$). According to Proposition 6.20, it suffices to prove the following claim.

Claim. Let K_1, \dots, K_r be given vectors (of positive integers) and let $n_0 \geq 1$ be such that $|K_r| + \dots + |K_1| + n_0 = k \leq N$. Then one can write

$$\begin{aligned}\hat{F}_w(H_{K_r} \partial_{e_1} \hat{\otimes}_s \dots \hat{\otimes}_s H_{K_1} \partial_{e_1}(H_1(n_0))) \\ = \sum_j \Phi_{K_1, \dots, K_r, n_0, j}(b_j^* + w, \dots, b_{j+k-1}^* + w) E_{j,j+k},\end{aligned}\tag{6.26}$$

where the $\Phi_{K_1, \dots, K_r, n_0, j}$'s are suitable meromorphic functions in k variables whose one-dimensional reductions

$$w \mapsto \Phi_{K_1, \dots, K_r, n_0, j}(b_j^* + w, \dots, b_{j+k-1}^* + w)$$

are all analytic near $w = 0$.

By using Lemma 6.17, the H_1 -formula (3.11) and Corollary 6.12, one can first rewrite the left hand side of (6.26) as

$$\begin{aligned}\hat{F}_w(H_{K_r} \partial_{e_1} \hat{\otimes}_s \dots \hat{\otimes}_s H_{K_1} \partial_{e_1}(H_1(n_0))) \\ = \sum_{l=0}^{\infty} \frac{B_l}{l!} \sum_{\sigma \in S_r} \sum_{\xi_0 + \xi_1 + \dots + \xi_r = l-r} \text{ad}_A^{\xi_r} \circ \text{ad}_{\hat{F}_w(H_{K_{\sigma(r)}})} \circ \text{ad}_A^{\xi_{r-1}} \\ \circ \dots \circ \text{ad}_A^{\xi_1} \circ \text{ad}_{\hat{F}_w(H_{K_{\sigma(1)}})} \circ \text{ad}_A^{\xi_0}(\hat{F}_w(D_{n_0} \tilde{L})).\end{aligned}\tag{6.27}$$

The main effort is to evaluate the above summation over the ξ_i 's by using the induction hypothesis on the $\hat{F}_w(H_{K_{\sigma(j)}})$'s (note that $|K_{\sigma(j)}| < k$). We first state the key lemma for this purpose and then explain the heavy notation involved.

Lemma 6.22. *Let L_1, \dots, L_r be given vectors of positive integers and let $n_0 \geq 1$. Suppose that $|L_1| + \dots + |L_r| + n_0 = k \leq N$. Then one has*

$$\begin{aligned}
& \sum_{\xi_0 + \xi_1 + \dots + \xi_r = l-r} \text{ad}_A^{\xi_r} \circ \text{ad}_{\hat{F}_w(H_{L_r})} \circ \dots \circ \text{ad}_A^{\xi_1} \circ \text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0} (\hat{F}_w(D_{n_0} \tilde{L})) \\
&= \sum_i \sum_{\mu=1}^{r+1} \Theta_{L_{\mu-1}, \dots, L_1, n_0}^i \cdot (b_i^{i+n_0^{\mu-1}-1})^{l+1-\mu} \\
& \quad \sum_{I \in \mathcal{W}_{\mu, r}} \frac{(-1)^{\varepsilon(I)} \Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_{\mathbf{p}, \mathbf{q}}^{\mu, r}}}{b_{(i_\mu), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu, r}} b_{(i_\mu, i_{\mu+1}), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu, r}} \dots b_{I, n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu, r}}} E_{i-|\mathbf{n}_{\mathbf{p}}^{\mu, r}|, i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r}|}. \tag{6.28}
\end{aligned}$$

Here $\mathcal{W}_{\mu, r}$ denotes the set of words $I = (i_\mu, i_{\mu+1}, \dots, i_r) \in \{\pm 1\}^{r-\mu+1}$. Given such a word I , the vector \mathbf{p} denotes the subword of (μ, \dots, r) which records the locations of -1 's in I and $\mathbf{q} \triangleq (\mu, \dots, r) \setminus \mathbf{p}$. The quantity $\varepsilon(I)$ denotes the length of \mathbf{p} (i.e. number of -1 's in I). The function $\Theta_{L_{\mu-1}, \dots, L_1, n_0}^i$ is defined by

$$\Theta_{L_{\mu-1}, \dots, L_1, n_0}^i \triangleq \Theta_{L_{\mu-1}, \dots, L_1, n_0, i}(b_i, b_{i+1}, \dots, b_{i+k-1}),$$

where $\Theta_{L_{\mu-1}, \dots, L_1, n_0, i}$ is a suitable meromorphic function in m complex variables ($m \triangleq n_0 + |L_1| + \dots + |L_{\mu-1}|$) whose shape depends on $(L_1, \dots, L_{\mu-1}), n_0$ and i . The family $\{\Theta_{L_1, \dots, L_{\mu-1}, n_0}^i\}$ of functions satisfy the following property: for any integer $i, 1 \leq \nu \leq r$ and any sequence of vectors $\mathbf{V} = (V_1, \dots, V_r)$, under the one-dimensional reduction $b_j \triangleq b_j^* + w$ the function

$$w \mapsto \sum_{\sigma \in \mathcal{S}_\nu} \Theta_{V_{\sigma(\nu)}, \dots, V_{\sigma(1)}, n_0}^i \tag{6.29}$$

is analytic near $w = 0$ and vanishes at $w = 0$. When $\mu = 1$, the function $\Theta_{L_{\mu-1}, \dots, L_1, n_0}^i$ depends only on n_0 and i . As a convention, if $\mu = r + 1$ the summation $\sum_{I \in \mathcal{W}(\mu, r)}$ in (6.22) is set to be one and $|\mathbf{n}_{\mathbf{p}}^{\mu, r}| = |\mathbf{n}_{\mathbf{q}}^{\mu, r}| \triangleq 0$.

Remark 6.23. *The functions $\Theta_{V_\nu, \dots, V_1, n_0}^i$ may be singular at $w = 0$; it is their symmetrisation (6.29) which will eliminate the possible singularity.*

Explanation of notation

Here we explain the notation involved in the expression (6.28).

(i) Let $\mathbf{L} \triangleq (L_1, \dots, L_r)$ and $n_i \triangleq |L_i|$ ($1 \leq i \leq r$). Given $0 \leq i \leq j \leq r$, we set $n_i^j \triangleq n_i + n_{i+1} + \dots + n_j$. For each $1 \leq \mu \leq r$, we set $\mathbf{n}^{\mu, r} \triangleq (n_\mu, \dots, n_r)$. Let $\mathbf{p} = (p_1, \dots, p_l)$ be a given subword of $\mathbf{n}^{\mu, r}$ and $\mathbf{q} \triangleq \mathbf{n}^{\mu, r} \setminus \mathbf{p} = (q_1, \dots, q_m)$. Elements in \mathbf{p}, \mathbf{q} are arranged in the natural order. We write

$$\mathbf{n}_{\mathbf{p}}^{\mu, r} \triangleq (n_{p_1}, \dots, n_{p_l}), \quad \mathbf{n}_{\mathbf{q}}^{\mu, r} \triangleq (n_{q_1}, \dots, n_{q_m})$$

respectively and set $|\mathbf{n}_p^{\mu,r}| \triangleq n_{p_1} + \cdots + n_{p_l}$.

(ii) Given $i \leq j$, we define $b_i^j \triangleq b_i + b_{i+1} + \cdots + b_j$. We always use $\Psi_{L,j}$ to denote a function in $|L|$ variables whose shape depends on a given vector L and a number j . We write

$$\Psi_L^j \triangleq \Psi_{L,j}(b_j, b_{j+1}, \dots, b_{j+|K|-1}).$$

Here $b_j \triangleq b_j^* + w$ is viewed as a function of w .

(iii) Let \mathbf{L} , n_i , μ , \mathbf{p} , \mathbf{q} be given as in Part (i). Let $I = (i_\mu, i_{\mu+1}, \dots, i_r)$ be a given word with $i_j = \pm 1$. Denote s (respectively, t) as the number of -1 's (respectively, 1 's) in I . Suppose that $s \leq l$ and $t \leq m$. We define

$$\mathbf{b}_{I, n_0^{\mu-1}}^{\mathbf{n}_p^{\mu,r}, \mathbf{q}} \triangleq b_{i_{-n_{p_1}} \dots - n_{p_s}}^{i-1} + b_{i+n_0^{\mu-1}}^{i+n_0^{\mu-1}+n_{q_1}+\dots+n_{q_t}-1}. \quad (6.30)$$

and

$$\begin{aligned} \Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_p^{\mu,r}, \mathbf{q}} &\triangleq \Psi_{L_{p_1}}^{i-n_{p_1}} \Psi_{L_{p_2}}^{i-n_{p_1}-n_{p_2}} \dots \Psi_{L_{p_s}}^{i-n_{p_1}-\dots-n_{p_s}} \\ &\quad \times \Psi_{L_{q_1}}^{i+n_0^{\mu-1}} \Psi_{L_{q_2}}^{i+n_0^{\mu-1}+q_1} \dots \Psi_{L_{q_t}}^{i+n_0^{\mu-1}+q_1+\dots+q_{t-1}}. \end{aligned} \quad (6.31)$$

Remark 6.24. The superscripts in \mathbf{b} and Ψ mean that the movement steps are taken from the word $\mathbf{n}^{\mu,r}$, where the indices \mathbf{p} and \mathbf{q} record backward and forward index movements with sizes n_{p_j} and n_{q_k} respectively. The subscript μ means that the index gap between the initial backward and forward movements is $n_0^\mu (= n_0 + n_1 + \cdots + n_\mu)$. In other words, the backward movement starts from index $i-1$ and the forward movement starts from $i+n_0^\mu$. In the above definitions, i is fixed and is not reflected in the notation.

Remark 6.25. The quantities $\mathbf{b}_{I, n_0^{\mu-1}}^{\mathbf{n}_p^{\mu,r}, \mathbf{q}}$ and $\Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_p^{\mu,r}, \mathbf{q}}$ depend on the word I only through the numbers s, t (numbers of ∓ 1 's respectively).

Proof of Lemma 6.22

Consider the following slightly more general claim which depends on the positive integer r .

Claim **P**(r): The following representation holds

$$\begin{aligned} &\sum_{\xi_0+\xi_1+\dots+\xi_r=v} \text{ad}_A^{\xi_r} \circ \text{ad}_{\hat{F}_w(H_{L_r})} \circ \cdots \circ \text{ad}_A^{\xi_1} \circ \text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0} (\hat{F}_w(D_{n_0} \tilde{L})) \\ &= \sum_i \sum_{\mu=1}^{r+1} \Theta_{L_{\mu-1}, \dots, L_1, n_0}^i \cdot (b_i^{i+n_0^{\mu-1}-1})^{v+r+1-\mu} \\ &\quad \sum_{I \in \mathcal{W}_{\mu,r}} \frac{(-1)^{\varepsilon(I)} \Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_p^{\mu,r}, \mathbf{q}}}{b_{(i_\mu), n_0^{\mu-1}}^{\mathbf{n}_p^{\mu,r}} b_{(i_\mu, i_{\mu+1}), n_0^{\mu-1}}^{\mathbf{n}_p^{\mu,r}} \cdots b_{I, n_0^{\mu-1}}^{\mathbf{n}_p^{\mu,r}}} E_{i-|\mathbf{n}_p^{\mu,r}|, i+n_0^{\mu-1}+|\mathbf{n}_q^{\mu,r}|} \end{aligned} \quad (6.32)$$

for any sequence of vectors $\mathbf{L} = (L_1, \dots, L_r)$ and $v \in \mathbb{N}$. Here the functions $\{\Theta_{L_{\mu-1}, \dots, L_1, n_0}^i\}$ satisfy the properties stated in Lemma 6.22.

Note that the result of Lemma 6.22 follows immediately by taking $v = l - r$. We now prove the claim $\mathbf{P}(r)$ by induction on r .

Step 1: Base case.

For the base case $r = 1$, one first recalls from Corollary 6.12 that

$$\hat{F}_w(D_{n_0} \tilde{L}) = \sum_i S_{n_0}^i E_{i, i+n_0}$$

where $S_{n_0}^i \triangleq S_{n_0}(b_i, b_{i+1}, \dots, b_{i+n_0-1})$. Therefore, one has

$$\text{ad}_A^{\xi_0}(\hat{F}_w(D_{n_0} \tilde{L})) = \sum_i S_{n_0}^i (b_i^{i+n_0-1})^{\xi_0} E_{i, i+n_0}. \quad (6.33)$$

Next, recall that the assumption (6.25) is valid since we are under the induction step for proving of the main Lemma 6.15 (all the L_j 's are assumed to satisfy $|L_j| < k$ where k is defined in the Claim (6.26)). In particular, one has

$$\hat{F}_w(H_{L_1}) = \sum_j \Psi_{L_1}^j E_{j, j+n_1} \quad (6.34)$$

where $\Psi_{L_1}^j$ are analytic near $w = 0$. It follows from (6.33, 6.34) that

$$\text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0}(\hat{F}_w(D_{n_0} \tilde{L})) = \sum_i S_{n_0}^i \cdot (b_i^{i+n_0-1})^{\xi_0} \sum_j \Psi_{L_1}^j [E_{j, j+n_1}, E_{i, i+n_0}].$$

There are at most two j -cases to make the above Lie bracket nonzero: $j = i - n_1$ and $j = i + n_0$. The former case is interpreted as a ‘‘backward movement’’ while the latter is a ‘‘forward movement’’. In other words, one can write

$$\text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0}(\hat{F}_w(D_{n_0} \tilde{L})) = \sum_i S_{n_0}^i (b_i^{i+n_0-1})^{\xi_0} (\Psi_{L_1}^{i-n_1} E_{i-n_1, i+n_0} - \Psi_{L_1}^{i+n_0} E_{i, i+n_0+n_1}).$$

The additional $\text{ad}_A^{\xi_1}$ -action yields that

$$\begin{aligned} \text{ad}_A^{\xi_1} \circ \text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0}(\hat{F}_w(D_{n_0} \tilde{L})) &= \sum_i S_{n_0}^i (b_i^{i+n_0-1})^{\xi_0} [\Psi_{L_1}^{i-n_1} \cdot (b_{i-n_1}^{i+n_0-1})^{\xi_1} E_{i-n_1, i+n_0} \\ &\quad - \Psi_{L_1}^{i+n_0} \cdot (b_i^{i+n_0+n_1-1})^{\xi_1} E_{i, i+n_0+n_1}]. \end{aligned}$$

Now we perform the summation over $\xi_0 + \xi_1 = v$ by using the elementary formula

$$\sum_{\xi_0 + \xi_1 = v} x^{\xi_0} y^{\xi_1} = \frac{y^{v+1} - x^{v+1}}{y - x}. \quad (6.35)$$

This gives that

$$\begin{aligned}
& \sum_{\xi_0+\xi_1=v} \text{ad}_A^{\xi_1} \circ \text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0} (F(D_{n_0}\tilde{L})) \\
&= \sum_i S_{n_0}^i \frac{\Psi_{L_1}^{i-n_1}}{b_{i-n_1}^{i-1}} [(b_{i-n_1}^{i+n_0-1})^{v+1} - (b_i^{i+n_0-1})^{v+1}] E_{i-n_1, i+n_0} \\
&\quad - \sum_i S_{n_0}^i \frac{\Psi_{L_1}^{i+n_0}}{b_{i+n_0}^{i+n_0+n_1-1}} [(b_i^{i+n_0+n_1-1})^{v+1} - (b_i^{i+n_0-1})^{v+1}] E_{i, i+n_0+n_1} \\
&= \sum_i S_{n_0}^i (b_i^{i+n_0-1})^{v+1} \left(-\frac{\Psi_{L_1}^{i-n_1}}{b_{i-n_1}^{i-1}} E_{i-n_1, i+n_0} + \frac{\Psi_{L_1}^{i+n_0}}{b_{i+n_0}^{i+n_0+n_1-1}} E_{i, i+n_0+n_1} \right) \\
&\quad + \sum_i \left[(S_{n_0}^{i+n_1} \frac{\Psi_{L_1}^i}{b_i^{i+n_1-1}} - S_{n_0}^i \frac{\Psi_{L_1}^{i+n_0}}{b_{i+n_0}^{i+n_0+n_1-1}}) (b_i^{i+n_0+n_1-1})^{v+1} \right] E_{i, i+n_0+n_1}, \tag{6.36}
\end{aligned}$$

where the last line follows from a change of indices $i - n_1 \leftrightarrow i$. The desired identity (6.32) thus follows by taking $\Theta_{n_0}^i \triangleq S_{n_0}^i$ ($\mu = 1$) and

$$\Theta_{L_1, n_0}^i \triangleq \left(S_{n_0}^{i+n_1} \frac{\Psi_{L_1}^i}{b_i^{i+n_1-1}} - S_{n_0}^i \frac{\Psi_{L_1}^{i+n_0}}{b_{i+n_0}^{i+n_0+n_1-1}} \right) (b_i^{i+n_0+n_1-1}). \tag{6.37}$$

The function Θ_{L_1, n_0}^i is analytic near $w = 0$ since the sequence (b_1^*, \dots, b_N^*) is non-degenerate (the denominators appearing in (6.37) are nonzero near $w = 0$). It is clear that both $\Theta_{n_0}^i$ and Θ_{L_1, n_0}^i vanishes at $w = 0$ due to the presence of the S_{n_0} -terms (cf. Assumption (\mathbf{A}_{N-1})).

Step 2: Evaluation of (6.32) based on induction hypothesis.

Now suppose that Claim $\mathbf{P}(r-1)$ is true and we want to prove Claim $\mathbf{P}(r)$. Let (L_1, \dots, L_r) be a given sequence of vectors and let $v \geq 0$. By applying (6.25) and

Claim **P**($r - 1$) to $v - \xi_r$, one finds that

$$\begin{aligned}
& \sum_{\xi_0 + \dots + \xi_{r-1} + \xi_r = v} \text{ad}_A^{\xi_r} \circ \text{ad}_{\hat{F}_w(H_{L_r})} \circ \text{ad}_A^{\xi_{r-1}} \circ \dots \circ \text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0} (\hat{F}_w(D_{n_0} \tilde{L})) \\
&= \sum_{\xi_r=0}^v \text{ad}_A^{\xi_r} \left(\left[\sum_j \Psi_{L_r}^j E_{j, j+n_r}, \sum_{\xi_0 + \dots + \xi_{r-1} = v - \xi_r} \text{ad}_A^{\xi_{r-1}} \circ \text{ad}_{\hat{F}_w(H_{L_{r-1}})} \circ \dots \circ \text{ad}_A^{\xi_1} \circ \text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0} (\hat{F}_w(D_{n_0} \tilde{L})) \right] \right) \\
&= \sum_{\xi_r=0}^v \sum_i \sum_{\mu=1}^r \Theta_{L_{\mu-1}, \dots, L_1, n_0}^i \cdot \sum_{I \in \mathcal{W}_{\mu, r-1}} \frac{(-1)^{\varepsilon(I)} \Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_{\mathbf{p}, \mathbf{q}}^{\mu, r-1}} \cdot \Psi_{L_r}^{i - |\mathbf{n}_{\mathbf{p}}^{\mu, r-1}| - n_r}}{b_{(i_\mu), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu, r-1}} \cdots b_{I, n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu, r-1}}} \\
&\quad \times (b_i^{i+n_0^{\mu-1}-1})^{v-\xi_r+r-\mu} (b_{i-|\mathbf{n}_{\mathbf{p}}^{\mu, r-1}| - n_r}^{i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r-1}|-1})^{\xi_r} E_{i-|\mathbf{n}_{\mathbf{p}}^{\mu, r-1}| - n_r, i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r-1}|} \quad (6.38)
\end{aligned}$$

$$\begin{aligned}
& - \sum_{\xi_r=0}^v \sum_i \sum_{\mu=1}^r \Theta_{L_{\mu-1}, \dots, L_1, n_0}^i \cdot \sum_{I \in \mathcal{W}_{\mu, r-1}} \frac{(-1)^{\varepsilon(I)} \Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_{\mathbf{p}, \mathbf{q}}^{\mu, r-1}} \cdot \Psi_{L_r}^{i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r-1}|}}{b_{(i_\mu), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu, r-1}} \cdots b_{I, n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu, r-1}}} \\
&\quad \times (b_i^{i+n_0^{\mu-1}-1})^{v-\xi_r+r-\mu} (b_{i-|\mathbf{n}_{\mathbf{p}}^{\mu, r-1}|}^{i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r-1}|+n_r-1})^{\xi_r} E_{i-|\mathbf{n}_{\mathbf{p}}^{\mu, r-1}|, i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r-1}|+n_r}. \quad (6.39)
\end{aligned}$$

Just like the base case, one performs the summation over ξ_r in (6.38) and (6.39) respectively. This gives that

$$\begin{aligned}
& \sum_{\xi_r=0}^v (b_i^{i+n_0^{\mu-1}-1})^{v-\xi_r+r-\mu} (b_{i-|\mathbf{n}_{\mathbf{p}}^{\mu, r-1}| - n_r}^{i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r-1}|-1})^{\xi_r} \\
&= (b_i^{i+n_0^{\mu-1}-1})^{r-\mu} \times \frac{(b_{i-|\mathbf{n}_{\mathbf{p}, r}^{\mu, r}|}^{i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r-1}|-1})^{v+1} - (b_i^{i+n_0^{\mu-1}-1})^{v+1}}{b_{(I, -1), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, r}^{\mu, r}, \mathbf{q}}} \quad (6.40)
\end{aligned}$$

and respectively,

$$\begin{aligned}
& \sum_{\xi_r=0}^v (b_i^{i+n_0^{\mu-1}-1})^{v-\xi_r+r-\mu} (b_{i-|\mathbf{n}_{\mathbf{p}}^{\mu, r-1}|}^{i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu, r-1}|+n_r-1})^{\xi_r} \\
&= (b_i^{i+n_0^{\mu-1}-1})^{r-\mu} \times \frac{(b_{i-|\mathbf{n}_{\mathbf{p}, r}^{\mu, r}|}^{i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}, r}^{\mu, r}|-1})^{v+1} - (b_i^{i+n_0^{\mu-1}-1})^{v+1}}{b_{(I, 1), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, r}^{\mu, r}, (\mathbf{q}, r)}}. \quad (6.41)
\end{aligned}$$

By substituting (6.40, 6.41) into (6.38, 6.39) respectively, one can write

$$\begin{aligned}
& \sum_{\xi_0 + \dots + \xi_{r-1} + \xi_r = v} \text{ad}_A^{\xi_r} \circ \text{ad}_{\hat{F}_w(H_{L_r})} \circ \dots \circ \text{ad}_A^{\xi_1} \circ \text{ad}_{\hat{F}_w(H_{L_1})} \circ \text{ad}_A^{\xi_0} (\hat{F}_w(D_{n_0} \tilde{L})) \\
&= \sum_i \sum_{\mu=1}^r \Theta_{L_{\mu-1}, \dots, L_1, n_0}^i (b_i^{i+n_0^{\mu-1}-1})^{v+r+1-\mu} \sum_{I \in \mathcal{W}_{\mu,r}} \frac{(-1)^{\varepsilon(I)} \Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_{\mathbf{p}, \mathbf{q}}^{\mu,r}}}{b_{(i_\mu), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}} \dots b_{I, n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}}} \\
&\quad \times E_{i-|\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}|, i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu,r}|} - \sum_i \sum_{\mu=1}^r \Theta_{L_{\mu-1}, \dots, L_1, n_0}^i (b_i^{i+n_0^{\mu-1}-1})^{r-\mu} \\
&\quad \sum_{I \in \mathcal{W}_{\mu,r}} \frac{(-1)^{\varepsilon(I)} \Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_{\mathbf{p}, \mathbf{q}}^{\mu,r}}}{b_{(i_\mu), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}} \dots b_{I, n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}}} (b_{i-|\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}|}^{i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu,r}|})^{v+1} E_{i-|\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}|, i+n_0^{\mu-1}+|\mathbf{n}_{\mathbf{q}}^{\mu,r}|} \\
&=: I + J. \tag{6.42}
\end{aligned}$$

Now we define

$$\begin{aligned}
\Theta_{L_r, \dots, L_1, n_0}^j &\triangleq - (b_j^{j+n_0^r}) \sum_{\mu=1}^r \sum_{i, I: i-|\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}|=j} \Theta_{L_{\mu-1}, \dots, L_1, n_0}^i \\
&\quad \times (b_i^{i+n_0^{\mu-1}-1})^{r-\mu} \frac{(-1)^{\varepsilon(I)} \Psi_{I, n_0^{\mu-1}}^{\mathbf{L}_{\mathbf{p}, \mathbf{q}}^{\mu,r}}}{b_{(i_\mu), n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}} \dots b_{I, n_0^{\mu-1}}^{\mathbf{n}_{\mathbf{p}, \mathbf{q}}^{\mu,r}}}. \tag{6.43}
\end{aligned}$$

It follows that

$$J = \sum_j \Theta_{L_r, \dots, L_1, n_0}^j (b_j^{j+n_0^r-1})^v E_{j, j+n_0^r}.$$

This is regarded as the term corresponding to $\mu = r + 1$ in the induction claim $\mathbf{P}(r)$ (cf. (6.32) and the I -term corresponds to the summation $\sum_{\mu=1}^r$ in the claim (6.32)). It is clear from its expression that $\Theta_{L_r, \dots, L_1, n_0}^j$ is a meromorphic function in the complex variables (b_j, b_{j+1}, \dots) (before the one-dimensional reduction $b_j = b_j^* + w$).

Step 3: Analyticity of the symmetrisation of J .

To finish the proof, it remains to show that the function

$$\sum_{\sigma \in \mathcal{S}_r} \Theta_{L(\sigma(r)), \dots, L(\sigma(1)), n_0}^j$$

is (as a function of w) analytic near $w = 0$. To this end, it is convenient to use the original expression of J defined in (6.42). Equivalently, we aim at showing that the

$\mathfrak{sl}_{N+1}(\mathbb{C})$ -valued function

$$\begin{aligned}
\sum_{\sigma \in \mathcal{S}_r} J &= - \sum_i \sum_{\mu=1}^r \sum_{\sigma \in \mathcal{S}_r} \sum_{I \in \mathcal{W}_{\mu,r}} \Theta_{L_{\sigma(\mu-1)}, \dots, L_{\sigma(1)}, n_0}^i (b_i^{i+(n^\sigma)_0^{\mu-1}-1})^{r-\mu} \\
&\quad \times \frac{(-1)^{\varepsilon(I)} \Psi_{I, (n^\sigma)_0^{\mu-1}}^{\mathbf{L}^\sigma}_{\mathbf{p}, \mathbf{q}}}{b_{(i_\mu), (n^\sigma)_0^{\mu-1}}^{(n^\sigma)_\mathbf{p}, \mathbf{q}} \cdots b_{I, (n^\sigma)_0^{\mu-1}}^{(n^\sigma)_\mathbf{p}, \mathbf{q}}} (b_{i-|(n^\sigma)_\mathbf{p}, \mathbf{q}|}^{i+(n^\sigma)_0^{\mu-1}+|(n^\sigma)_\mathbf{q}^{\mu,r}-1})^{v+1} \\
&\quad \times E_{i-|(n^\sigma)_\mathbf{p}, \mathbf{q}|, i+(n^\sigma)_0^{\mu-1}+|(n^\sigma)_\mathbf{q}^{\mu,r}|} \tag{6.44}
\end{aligned}$$

is analytic near $w = 0$. Here $\mathbf{L}^\sigma \triangleq (L_{\sigma(1)}, \dots, L_{\sigma(r)})$ and respectively, we denote $\mathbf{n}^\sigma \triangleq (n_{\sigma(1)}, \dots, n_{\sigma(r)})$ and $(n^\sigma)_0^{\mu-1} \triangleq n_0 + n_{\sigma(1)} + \dots + n_{\sigma(\mu-1)}$.

Let i and μ be given fixed. The first crucial observation is that any permutation $\sigma \in \mathcal{S}_r$ can be written as

$$\sigma = \tau \circ (\zeta \otimes \theta),$$

where τ is a $(\mu-1, r+1-\mu)$ -shuffle, ζ is a permutation over $\{1, \dots, \mu-1\}$ and θ is a permutation over $\{\mu, \dots, r\}$. In this way, the summation over σ and I in (6.44) can be rewritten as

$$\begin{aligned}
&\sum_{\tau \in \mathcal{P}(\mu-1, r+1-\mu)} \sum_{\zeta \in \mathcal{S}_{\{1, \dots, \mu-1\}}} \Theta_{L_{\tau \circ \zeta(\mu-1)}, \dots, L_{\tau \circ \zeta(1)}, n_0}^i (b_i^{i+(n^{\tau \circ \zeta})_0^{\mu-1}-1})^{r-\mu} \\
&\quad \times \sum_{s=0}^{r+1-\mu} (-1)^s \sum_{\theta \in \mathcal{S}_{\{\mu, \dots, r\}}} \sum_{I \in \mathcal{W}_{\mu,r}: |\mathbf{p}|=s} \frac{\Psi_{I, (n^{\tau \circ \zeta})_0^{\mu-1}}^{\mathbf{L}^{\tau \circ \theta}}_{\mathbf{p}, \mathbf{q}}}{b_{(i_\mu), (n^{\tau \circ \zeta})_0^{\mu-1}}^{(n^{\tau \circ \theta})_\mathbf{p}, \mathbf{q}} \cdots b_{I, (n^{\tau \circ \zeta})_0^{\mu-1}}^{(n^{\tau \circ \theta})_\mathbf{p}, \mathbf{q}}} \\
&\quad \times (b_{i-|(n^{\tau \circ \theta})_\mathbf{p}, \mathbf{q}|}^{i+(n^{\tau \circ \zeta})_0^{\mu-1}+|(n^{\tau \circ \theta})_\mathbf{q}^{\mu,r}-1})^{v+1} E_{i-|(n^{\tau \circ \theta})_\mathbf{p}, \mathbf{q}|, i+(n^{\tau \circ \zeta})_0^{\mu-1}+|(n^{\tau \circ \theta})_\mathbf{q}^{\mu,r}|}. \tag{6.45}
\end{aligned}$$

It is important to note that the expression in the last two lines depends only on τ and is independent of ζ , since

$$(n^{\tau \circ \zeta})_0^{\mu-1} \triangleq n_0 + n_{\tau(\zeta(1))} + \dots + n_{\tau(\zeta(\mu-1))} = n_0 + n_{\tau(1)} + \dots + n_{\tau(\mu-1)}.$$

Let us denote the summation over (θ, I) in (6.45) by $\mathcal{T}_{s, \tau}$. It follows that

$$\begin{aligned}
(6.45) &= \sum_{\tau \in \mathcal{P}(\mu-1, r+1-\mu)} (b_i^{i+(n^\tau)_0^{\mu-1}-1})^{r-\mu} \sum_{s=0}^{r+1-\mu} (-1)^s \mathcal{T}_{s, \tau} \\
&\quad \times \sum_{\zeta \in \mathcal{S}_{\{1, \dots, \mu-1\}}} \Theta_{L_{\tau \circ \zeta(\mu-1)}, \dots, L_{\tau \circ \zeta(1)}, n_0}^i.
\end{aligned}$$

According to the induction hypothesis $\mathbf{P}(r-1)$, the function

$$\sum_{\zeta \in \mathcal{S}_{\{1, \dots, \mu-1\}}} \Theta_{L_{\tau \circ \zeta(\mu-1)}, \dots, L_{\tau \circ \zeta(1)}, n_0}^i$$

is (as a function of w) analytic near $w = 0$ and vanishes at $w = 0$.

Now our task is reduced to showing that the function $\mathcal{T}_{s,\tau}$ is analytic near $w = 0$ for any fixed s and τ . To ease notation, we will temporarily write

$$\bar{n}_0 \triangleq (n^\tau)_0^{\mu-1}, \quad \bar{\mathbf{L}} \triangleq (L_{\tau(\mu)}, \dots, L_{\tau(r)}), \quad \bar{\mathbf{n}} \triangleq (n_{\tau(\mu)}, \dots, n_{\tau(r)}).$$

Given a permutation $\theta \in \mathcal{S}_{\{\mu, \dots, r\}}$, we denote

$$\bar{\mathbf{L}}^\theta \triangleq (\bar{L}_{\theta(\mu)}, \dots, \bar{L}_{\theta(r)}) = (L_{\tau(\theta(\mu))}, \dots, L_{\tau(\theta(r))})$$

and respectively,

$$\bar{\mathbf{n}}^\theta \triangleq (\bar{n}_{\theta(\mu)}, \dots, \bar{n}_{\theta(r)}) = (n_{\tau(\theta(\mu))}, \dots, n_{\tau(\theta(r))}).$$

Under the above simplified notation, our target $\mathcal{T}_{s,\tau}$ is defined by

$$\begin{aligned} \mathcal{T}_{s,\tau} \triangleq & \sum_{\theta \in \mathcal{S}_{\{\mu, \dots, r\}}} \sum_{I \in \mathcal{W}_{\mu,r}: |\mathbf{p}|=s} \frac{\Psi_{I, \bar{n}_0}^{(\bar{\mathbf{L}}^\theta)_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}}}{b_{(i_\mu), \bar{n}_0}^{(\bar{\mathbf{n}}^\theta)_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}} \dots b_{I, \bar{n}_0}^{(\bar{\mathbf{n}}^\theta)_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}}} \\ & \times \left(b_{i - |(\bar{\mathbf{n}}^\theta)_{\bar{\mathbf{p}}}^{\mu,r}|}^{i + \bar{n}_0 + |(\bar{\mathbf{n}}^\theta)_{\bar{\mathbf{q}}}^{\mu,r}| - 1} \right)^{v+1} E_{i - |(\bar{\mathbf{n}}^\theta)_{\bar{\mathbf{p}}}^{\mu,r}|, i + \bar{n}_0 + |(\bar{\mathbf{n}}^\theta)_{\bar{\mathbf{q}}}^{\mu,r}|} \end{aligned} \quad (6.46)$$

The next crucial observation is that the set of words $I \in \mathcal{W}_{\mu,r}$ with precisely s number of -1 's is in one-to-one correspondence with the set $\mathcal{P}(s; \mu, r)$ of $(s, r+1-\mu-s)$ -shuffles over $\{\mu, \dots, r\}$, i.e. permutations $\rho \in \mathcal{S}_{\{\mu, \dots, r\}}$ satisfying

$$\rho(\mu) < \rho(\mu+1) < \dots < \rho(\mu+s-1), \quad \rho(\mu+s) < \dots < \rho(r).$$

Indeed, given $\rho \in \mathcal{P}(s; \mu, r)$ one can define

$$I_\rho = (i_\mu, \dots, i_r) : i_{\rho(j)} = \begin{cases} -1, & j = \mu, \dots, \mu+s-1; \\ 1, & j = \mu+s, \dots, r. \end{cases}$$

In this way, $\bar{\mathbf{p}} = (\rho(\mu), \dots, \rho(\mu+s-1))$ and $\bar{\mathbf{q}} = (\rho(\mu+s), \dots, \rho(r))$. Using this correspondence, one can write

$$\begin{aligned} \mathcal{T}_{s,\tau} = & \sum_{\theta \in \mathcal{S}_{\{\mu, \dots, r\}}} \sum_{\rho \in \mathcal{P}(s; \mu, r)} \frac{\Psi_{I_\rho, \bar{n}_0}^{(\bar{\mathbf{L}}^{\theta \circ \rho})_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}}}{b_{(i_\mu), \bar{n}_0}^{(\bar{\mathbf{n}}^{\theta \circ \rho})_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}} \dots b_{I_\rho, \bar{n}_0}^{(\bar{\mathbf{n}}^{\theta \circ \rho})_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}}} \\ & \times \left(b_{i - |(\bar{\mathbf{n}}^{\theta \circ \rho})_{\bar{\mathbf{p}}}^{\mu,r}|}^{i + \bar{n}_0 + |(\bar{\mathbf{n}}^{\theta \circ \rho})_{\bar{\mathbf{q}}}^{\mu,r}| - 1} \right)^{v+1} E_{i - |(\bar{\mathbf{n}}^{\theta \circ \rho})_{\bar{\mathbf{p}}}^{\mu,r}|, i + \bar{n}_0 + |(\bar{\mathbf{n}}^{\theta \circ \rho})_{\bar{\mathbf{q}}}^{\mu,r}|}. \end{aligned} \quad (6.47)$$

Here $\bar{\mathbf{p}} \triangleq (\mu, \dots, \mu+s-1)$ and $\bar{\mathbf{q}} \triangleq (\mu+s, \dots, r)$ are the two canonical subwords which are both independent of ρ . Recall from Remark 6.25 that the function $\Psi_{I_\rho, \bar{n}_0}^{(\bar{\mathbf{L}}^{\theta \circ \rho})_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}}$ depends on I_ρ only through the numbers of ∓ 1 's in it. In particular, one has

$$\Psi_{I_\rho, \bar{n}_0}^{(\bar{\mathbf{L}}^{\theta \circ \rho})_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}} = \Psi_{\bar{I}, \bar{n}_0}^{(\bar{\mathbf{L}}^{\theta \circ \rho})_{\bar{\mathbf{p}}, \bar{\mathbf{q}}}^{\mu,r}}$$

for all $\rho \in \mathcal{P}(s; \mu, r)$ where

$$\bar{I} = \underbrace{(-1, \dots, -1)}_s, \underbrace{(1, \dots, 1)}_{r+1-\mu-s}.$$

With this observation in mind, the point is that one can now exchange the summations in (6.47), and after doing so, since the θ -summation is a full symmetrisation one could just replace $\theta \circ \rho$ by θ (ignoring the role of ρ). In other words, one has

$$\begin{aligned} \mathcal{T}_{s,\tau} &= \sum_{\rho \in \mathcal{P}(s;\mu,r)} \sum_{\theta \in \mathcal{S}_{\{\mu, \dots, r\}}} \frac{\Psi_{\bar{I}, \bar{n}_0}^{(\bar{I}^\theta)_{\mathbf{p}, \hat{\mathbf{q}}}^{\mu, r}}}{b_{(i_\mu), \bar{n}_0}^{(\bar{\mathbf{n}}^\theta)_{\mathbf{p}, \hat{\mathbf{q}}}^{\mu, r}} \dots b_{I_\rho, \bar{n}_0}^{(\bar{\mathbf{n}}^\theta)_{\mathbf{p}, \hat{\mathbf{q}}}^{\mu, r}}} \\ &\quad \times \left(b_{i-|\bar{\mathbf{n}}^\theta|_{\mathbf{p}}}^{i+\bar{n}_0+|\bar{\mathbf{n}}^\theta|_{\hat{\mathbf{q}}}-1} \right)^{v+1} E_{i-|\bar{\mathbf{n}}^\theta|_{\mathbf{p}}, i+\bar{n}_0+|\bar{\mathbf{n}}^\theta|_{\hat{\mathbf{q}}}}^{\mu, r} \\ &= \sum_{\theta \in \mathcal{S}_{\{\mu, \dots, r\}}} \Psi_{\bar{I}, \bar{n}_0}^{(\bar{I}^\theta)_{\mathbf{p}, \hat{\mathbf{q}}}^{\mu, r}} \cdot \left(b_{i-|\bar{\mathbf{n}}^\theta|_{\mathbf{p}}}^{i+\bar{n}_0+|\bar{\mathbf{n}}^\theta|_{\hat{\mathbf{q}}}-1} \right)^{v+1} \\ &\quad \times \left[\sum_{\rho \in \mathcal{P}(s;\mu,r)} \frac{1}{b_{(i_\mu), \bar{n}_0}^{(\bar{\mathbf{n}}^\theta)_{\mathbf{p}, \hat{\mathbf{q}}}^{\mu, r}} \dots b_{I_\rho, \bar{n}_0}^{(\bar{\mathbf{n}}^\theta)_{\mathbf{p}, \hat{\mathbf{q}}}^{\mu, r}}} \right] E_{i-|\bar{\mathbf{n}}^\theta|_{\mathbf{p}}, i+\bar{n}_0+|\bar{\mathbf{n}}^\theta|_{\hat{\mathbf{q}}}}^{\mu, r}. \end{aligned}$$

where we changed the order of summation back to reach the last equality.

It remains to show that for every fixed θ , the summation

$$\mathcal{T}_{s,\tau;\theta} \triangleq \sum_{\rho \in \mathcal{P}(s;\mu,r)} \frac{1}{b_{(i_\mu), \bar{n}_0}^{(\bar{\mathbf{n}}^\theta)_{\mathbf{p}, \hat{\mathbf{q}}}^{\mu, r}} \dots b_{I_\rho, \bar{n}_0}^{(\bar{\mathbf{n}}^\theta)_{\mathbf{p}, \hat{\mathbf{q}}}^{\mu, r}}}$$

defines an analytic function near $w = 0$. The third crucial observation is that $\mathcal{T}_{s,\tau;\theta}$ is precisely in the form of Lemma 6.21. To see this point, let us shift the indices $\{\mu, \dots, r\}$ back to the standard form $\{1, \dots, R\}$ ($R \triangleq r + 1 - \mu$). Namely, we denote

$$\hat{\mathbf{n}} \triangleq (\hat{n}_1, \dots, \hat{n}_R), \quad \hat{n}_i \triangleq \bar{n}_{\theta(\mu+i-1)}; \quad \hat{\mathbf{p}} \triangleq (1, \dots, s), \quad \hat{\mathbf{q}} \triangleq (s+1, \dots, R).$$

In addition, any shuffle $\rho \in \mathcal{P}(s; \mu, r)$ corresponds to some $\hat{\rho} \in \mathcal{P}(s, R-s)$ in the obvious way (and also at the level of words $I \leftrightarrow \hat{I}_\rho$). As a result, one can rewrite

$$\mathcal{T}_{s,\tau;\theta} = \sum_{\hat{\rho} \in \mathcal{P}(s, R-s)} \frac{1}{b_{(\hat{i}_1), \bar{n}_0}^{\hat{\mathbf{n}}_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{1, R}} \dots b_{\hat{I}_\rho, \bar{n}_0}^{\hat{\mathbf{n}}_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{1, R}}}$$

We define the sequence of numbers $\{c_j\}_{1 \leq j \leq R+1}$ in the following way:

$$\begin{array}{cccccccc} b_{i-\sum_{j=1}^{s-1} \hat{n}_j-1} & \dots & b_{i-\hat{n}_1-1} & b_{i-\hat{n}_1} & b_i^{i+\bar{n}_0-1} & b_{i+\bar{n}_0}^{i+\bar{n}_0+\hat{n}_{s+1}-1} & \dots & b_{i+\bar{n}_0+\sum_{j=s+1}^R \hat{n}_j-1} \\ c_1 & \dots & c_{s-1} & c_s & c_{s+1} & c_{s+2} & \dots & c_{R+1}. \end{array}$$

Note that there is a one-to-one correspondence between $\mathcal{P}(s, R-s)$ and $\mathcal{T}(s, r-s)$ (cf. Lemma 6.21 for its definition) given by

$$\begin{aligned} \hat{\rho} &\mapsto \eta : (\eta(1), \eta(2), \dots, \eta(s), \eta(s+1), \eta(s+2), \dots, \eta(R+1)) \\ &\triangleq ((\hat{\rho}(s) + 1, \hat{\rho}(s-1) + 1, \dots, \hat{\rho}(1) + 1, 1, \hat{\rho}(s+1) + 1, \dots, \hat{\rho}(R) + 1)). \end{aligned}$$

It is then readily checked that

$$b_{(\hat{i}_1), \bar{n}_0}^{\hat{\mathbf{n}}_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{1,R}} \cdots b_{\hat{I}, \bar{n}_0}^{\hat{\mathbf{n}}_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{1,R}} = c_{\eta^{-1}(2)}^{\eta^{-1}(2)} c_{\eta^{-1}(3)}^{\eta^{-1}(3)} \cdots c_{\eta^{-1}(R+1)}^{\eta^{-1}(R+1)},$$

where the right hand side is defined by (6.24). According to Lemma 6.21, one concludes that

$$\sum_{\hat{\rho} \in \mathcal{P}(s, R-s)} \frac{1}{b_{(\hat{i}_1), \bar{n}_0}^{\hat{\mathbf{n}}_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{1,R}} \cdots b_{\hat{I}, \bar{n}_0}^{\hat{\mathbf{n}}_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{1,R}}} = \left(\prod_{k=1}^s c_k^s \right)^{-1} \left(\prod_{k=s+2}^{R+1} c_{s+2}^k \right)^{-1}. \quad (6.48)$$

By the definition of c_j , it is apparent that both c_k^s and c_{s+2}^k are certain consecutive sums of the b_j 's. In particular, (6.48) is of the form $\frac{1}{b_{s_1}^{t_1} \cdots b_{s_m}^{t_m}}$ for suitable $s_1 \leq t_1, \dots, s_m \leq t_m$. Since the base point (b_1^*, \dots, b_N^*) is assumed to be non-degenerate, it follows that the function $\mathcal{T}_{s, \tau; \theta}$ (as a function of w) is analytic near $w = 0$.

The proof of the induction step is now complete.

Completing the proof of Lemma 6.15

We have obtained from (6.27) and Lemma 6.22 that

$$\begin{aligned} &\hat{F}_w(H_{K_r} \partial_{e_1} \hat{\otimes}_s \cdots \hat{\otimes}_s H_{K_1} \partial_{e_1} (H_1(n_0))) \\ &= \sum_l \frac{B_l}{l!} \sum_i \sum_{\mu=1}^{r+1} \sum_{\sigma \in \mathcal{S}_r} \Theta_{K_{\sigma(\mu-1)}, \dots, K_{\sigma(1)}, n_0}^i (b_i^{i+(n^\sigma)_0^{\mu-1}-1})^{l+1-\mu} \\ &\quad \sum_{I \in \mathcal{W}_{\mu, r}} \frac{(-1)^{\varepsilon(I)} \Psi_{I, (n^\sigma)_0^{\mu-1}}^{(\mathbf{K}^\sigma)_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{\mu, r}}}{b_{(i_\mu), (n^\sigma)_0^{\mu-1}}^{(n^\sigma)_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{\mu, r}} \cdots b_{I, (n^\sigma)_0^{\mu-1}}^{(n^\sigma)_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{\mu, r}}} E_{i-|(n^\sigma)_{\hat{\mathbf{p}}, \hat{\mathbf{q}}}^{\mu, r}|, i+(n^\sigma)_0^{\mu-1}+|(n^\sigma)_{\hat{\mathbf{q}}, \hat{\mathbf{r}}}^{\mu, r}|}. \end{aligned} \quad (6.49)$$

The argument for proving the analyticity of (6.49) near $w = 0$ is identical to the analysis in Step 3 of the proof of Lemma 6.22. The key idea, the same as before, is to write any permutation $\sigma \in \mathcal{S}_r$ as $\sigma = \tau \circ (\zeta \otimes \theta)$ where $\tau \in \mathcal{P}(\mu-1, r+1-\mu)$, $\zeta \in \mathcal{S}_{\mu-1}$ and $\theta \in \mathcal{S}_{\{\mu, \dots, r\}}$. This allows one to split the σ -action in (6.22) into permutations over the $\{1, \dots, \mu-1\}$ and $\{\mu, \dots, r\}$ parts separately. By exactly the same analysis as in

Step 3 of the proof of Lemma 6.22, one finds that

$$\begin{aligned}
& \hat{F}_w(H_{K_r} \partial_{e_1} \hat{\otimes}_s \cdots \hat{\otimes}_s H_{K_1} \partial_{e_1} (H_1(n_0))) \\
&= \sum_i \sum_{\mu=1}^{r+1} \sum_{\tau \in \mathcal{P}(\mu-1, r+1-\mu)} (b_i^{i+(n^\tau)_0^{\mu-1}})^{1-\mu} \phi(b_i^{i+(n^\tau)_0^{\mu-1}}) \\
& \quad \times \left(\sum_{\zeta \in \mathcal{S}_{\mu-1}} \Theta_{K_{\tau \circ \zeta(\mu-1)}, \dots, K_{\tau \circ \zeta(1)}, n_0}^i \right) \cdot \Xi_{K_{\tau(\mu)}, \dots, K_{\tau(1)}, (n^\tau)_0^{\mu-1}}^i. \tag{6.50}
\end{aligned}$$

Here

$$\begin{aligned}
\Xi_{K_{\tau(\mu)}, \dots, K_{\tau(r)}, (n^\tau)_0^{\mu-1}}^i &\triangleq \sum_{\theta \in \mathcal{S}_{\{\mu, \dots, r\}}} \sum_{I \in \mathcal{W}_{\mu, r; |\mathbf{p}|=s}} \frac{(-1)^{\varepsilon(I)} \Psi_{I, (n^\tau)_0^{\mu-1}}^{(\mathbf{K}^{\tau \circ \theta})_{\mathbf{p}, \mathbf{q}}^{\mu, r}}}{b_{(i_\mu), (n^\tau)_0^{\mu-1}}^{(\mathbf{n}^{\tau \circ \theta})_{\mathbf{p}, \mathbf{q}}^{\mu, r}} \cdots b_{I, (n^\tau)_0^{\mu-1}}^{(\mathbf{n}^{\tau \circ \theta})_{\mathbf{p}, \mathbf{q}}^{\mu, r}}} \\
& \quad E_{i-|(\mathbf{n}^{\tau \circ \theta})_{\mathbf{p}, \mathbf{q}}^{\mu, r}|, i+(n^\tau)_0^{\mu-1}+|(\mathbf{n}^\sigma)_{\mathbf{q}}^{\mu, r}|}
\end{aligned}$$

for $1 \leq \mu \leq r$ and as a convention

$$\Xi_{K_{\tau(\mu)}, \dots, K_{\tau(r)}, (n^\tau)_0^{\mu-1}}^i \triangleq E_{i, i+n_0^r}$$

if $\mu = r + 1$. The function $\Xi_{K_{\tau(\mu)}, \dots, K_{\tau(r)}, (n^\tau)_0^{\mu-1}}^i$ is an $\mathfrak{sl}_{N+1}(\mathbb{C})$ -valued analytic function near $w = 0$ for the same reason leading to the analyticity of $\mathcal{T}_{s, \tau}$ as before (cf. (6.46)). The analyticity of (6.50) near $w = 0$ thus follows from the facts that

(i) $w = 0$ is a simple pole for the function

$$\phi(b_i^{i+(n^\tau)_0^{\mu-1}}) = \phi(b_i^* + \cdots + b_{i+(n^\tau)_0^{\mu-1}-1}^* + (n^\tau)_0^{\mu-1} w);$$

(ii) the function

$$\sum_{\zeta \in \mathcal{S}_{\mu-1}} \Theta_{K_{\tau \circ \zeta(\mu-1)}, \dots, K_{\tau \circ \zeta(1)}, n_0}^i$$

vanishes at $w = 0$.

This show that $\hat{F}_w(H_{K_r} \partial_{e_1} \hat{\otimes}_s \cdots \hat{\otimes}_s H_{K_1} \partial_{e_1} (H_1(n_0)))$ is an $\mathfrak{sl}_{N+1}(\mathbb{C})$ -valued analytic function near $w = 0$. The proof of Lemma 6.15 is now complete.

Acknowledgement

XG gratefully acknowledges the support from ARC grant DE210101352. SW gratefully acknowledges the support from Melbourne Research Scholarship.

References

- [1] F. Baudoin. *An introduction to the geometry of stochastic flows*. Imperial College Press, 2004.
- [2] H. Boedihardjo and X. Geng. $SL_2(\mathbb{R})$ -developments and signature asymptotics for planar paths with bounded variation. *Rev. Mat. Iberoamericana*, 293:720–737, 2016.
- [3] H. Boedihardjo, X. Geng, T. Lyons, and D. Yang. The signature of a rough path: uniqueness. *Adv. Math.*, 293:720–737, 2016.
- [4] H. Boedihardjo, X. Geng, and N. Souris. Path developments and tail asymptotics of signature for pure rough paths. *Adv. Math.*, 364, 2020.
- [5] H. Boedihardjo, H. Ni, and Z. Qian. Uniqueness of signature for simple curves. *J. Funct. Anal.*, 267(6):1778–1806, 2014.
- [6] T. Cass and W.F. Turner. Free probability, path developments and signature kernels as universal scaling limits. *ArXiv preprint*, 2024.
- [7] K.T. Chen. Integration of paths, geometric invariants and a generalized baker-hausdorff formula. *Ann. of Math.*, 65:163–178, 1957.
- [8] K.T. Chen. Integration of paths—a faithful representation of paths by noncommutative formal power series. *Trans. Amer. Math. Soc.*, 89:395–407, 1958.
- [9] K.T. Chen. Iterated integrals of differential forms and loop space homology. *Ann. of Math.*, 97(2):217–246, 1973.
- [10] I. Chevyrev and T. Lyons. Characteristic functions of measures on geometric rough paths. *Ann. Probab.*, 44(6):4049–4082, 2016.
- [11] C. Cuchiero, G. Gazzani, and S. Svaluto-Ferro. Signature-based models: theory and calibration. *SIAM J. Financial Math.*, 14(3):910–957, 2023.
- [12] A. Fermanian, T. Lyons, J. Morrill, and C. Salvi. New directions in the applications of rough path theory. *IEEE BITS the Information Theory Magazine*, pages 1–18, 2023.
- [13] G. Flint and T. Lyons. Pathwise approximation of sdes by coupling piecewise abelian rough paths. *ArXiv preprint*, 2015.
- [14] P.K. Friz and N. Victoir. *Multidimensional stochastic processes as rough paths*. Cambridge University Press, 2010.
- [15] X. Geng. *An introduction to the theory of rough paths*. Unpublished Online Lecture Notes, 2021.

- [16] B. Hambly and T. Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. of Math.*, 171(1):109–167, 2010.
- [17] K. Hara and M. Hino. Fractional order taylor’s series and the neo-classical inequality. *Bull. Lond. Math. Soc.*, 42:467–477, 2010.
- [18] S. Kusuoka and D. Stroock. Applications of the malliavin calculus, part iii. *J. Fac. Sci. Univ. Tokyo*, 34:391–442, 1987.
- [19] T. Lyons, M. Caruana, and T. Lévy. *Differential equations driven by rough paths*. Springer, 2007.
- [20] T. Lyons and N. Sidorova. On the radius of convergence of the logarithmic signature. *Ill. J. Math.*, 50(4):763–790, 2006.
- [21] T. Lyons and N. Victoir. Cubature on wiener space. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 460(2041):169–198, 2004.
- [22] T. Lyons and X. Weijun. Hyperbolic development and the inversion of signature. *J. Funct. Anal.*, 272(7):2933–2955, 2015.
- [23] T.J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14:215–310, 1998.
- [24] H. Ni, L. Szpruch, M. Wiese, S. Liao, and B. Xiao. Conditional sig- wasserstein gans for time series generation. *Math. Finance*, 2023.
- [25] D. Nualart. *The Malliavin calculus and related topics*. Springer-Verlag, 2006.
- [26] C. Reutenauer. *Free Lie algebras*. Clarendon Press, 1993.
- [27] C. Salvi, T. Cass, J. Foster, T. Lyons, and W. Yang. The signature kernel is the solution of a goursat pde. *SIAM J. Math. Data Sci.*, 3(3):873–899, 2021.
- [28] R.S. Strichartz. The campbell-baker-hausdorff-dynkin formula and solutions of differential equations. *J. Funct. Anal.*, 72:320–345, 1987.
- [29] B. Walker, A.D. McLeod, T. Qin, Y. Cheng, H. Li, and T. Lyons. Log neural controlled differential equations: the lie brackets make a difference. *ArXiv Preprint*, 2024.