



**HAL**  
open science

# Itô formulae for the stochastic heat equation via the theories of regularity structure and rough paths

Carlo Bellingeri

► **To cite this version:**

Carlo Bellingeri. Itô formulae for the stochastic heat equation via the theories of regularity structure and rough paths. Probability [math.PR]. Sorbonne Université, 2019. English. NNT : 2019SORUS028 . tel-02935889v2

**HAL Id: tel-02935889**

**<https://theses.hal.science/tel-02935889v2>**

Submitted on 10 Sep 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SORBONNE UNIVERSITÉ  
École Doctorale des Sciences Mathématiques de Paris Centre  
*Laboratoire de Probabilités, Statistique et Modélisation*

THÈSE DE DOCTORAT  
Discipline : **Mathématiques**

présentée par  
**Carlo BELLINGERI**

---

**Formules d'Itô pour l'équation de la chaleur  
stochastique à travers les théories des chemins  
rugueux et des structures de régularité**

---

dirigée par **Lorenzo ZAMBOTTI**

Rapporteurs :

M. **Antoine LEJAY**                      Université de Lorraine  
M. **Nicolas PERKOWSKI**              Humboldt-Universität zu Berlin

Soutenue le 12 Juillet 2019 devant le jury composé de :

M. <b>Paul GASSIAT</b>	Université Paris-Dauphine	Examinateur
M. <b>Cyril LABBÉ</b>	Université Paris-Dauphine	Examinateur
M. <b>Antoine LEJAY</b>	Université de Lorraine	Rapporteur
M. <b>Thierry LÉVY</b>	Sorbonne Université	Examinateur
M. <b>Frédéric PATRAS</b>	Université de la Côte d'Azur	Examinateur
M. <b>Lorenzo ZAMBOTTI</b>	Sorbonne Université	Directeur



*...Bénie sois-tu, puissante Matière,  
Évolution irrésistible,  
Réalité toujours naissante,  
Toi qui, faisant éclater à tout moment nos cadres,  
nous obliges à poursuivre toujours plus loin la Vérité...*

Pierre Teilhard de Chardin, *Hymne à la Matière*.



# Remerciements/Ringraziamenti

C'est avec une immense joie que j'exprime ma gratitude à tous ceux qui ont rendu possible l'écriture de ce travail.

Tout d'abord, c'est un plaisir de remercier chaleureusement mon directeur de thèse, Lorenzo Zambotti, pour sa patience, sa disponibilité et tous les encouragements qu'il m'a donnés. Merci de m'avoir montré toute la beauté des équations aux dérivées partielles stochastiques et merci de toutes les discussions (mathématiques ou non) que nous avons eues pendant ces années.

Je tiens aussi à exprimer toute ma reconnaissance à mes rapporteurs, Nicolas Perkowski et Antoine Lejay, pour avoir accepté de lire en détail le manuscrit de thèse et pour l'avoir enrichi avec leurs commentaires. De même, je suis très honoré de la présence dans le jury de Cyril Labbé, Paul Gassiat, Thierry Levy et Frédéric Patras.

Faire les mathématiques n'est pas seulement un travail solitaire, mais aussi une grande passion à communiquer. J'exprime donc ma gratitude à tous les chercheurs avec lesquels j'ai pu échanger mes idées pendant que j'écrivais cette thèse : Peter Friz, Eric Saias, Kurusch Ebrahimi-Fard, Massimiliano Gubinelli, Titus Lupu, Tom Klose, Charles Curry, Paul Melotti, Florian Bechtold, Nikolas Tapia, Nicolas Gilliers, Yvain Bruned, Tristan Haugomat. Merci pour toutes vos suggestions.

Pendant ces années parisiennes et même avant, j'ai eu la chance de rencontrer sur mon chemin des amis fidèles qui m'ont aidé dans la relecture de certaines sections. Merci infiniment donc à Andrea Bondesan, Mattia Cavicchi, Vincent Daruvar et mon frère "jumeau" de thèse Henri Elad-Altman. *Grazie di tutto.*

Une bonne partie de ce travail est née dans les couloirs du LPSM et je suis reconnaissant au personnel administratif qui m'a énormément aidé dans les tâches administratives : Florence, Serena, Valérie, Louise, Fatime et Corinne. Un grand merci aussi à Daphné pour son aide fondamentale dans la création de ma page web. Par ailleurs, je remercie également tous mes collègues du laboratoire avec lesquels la solitude de la thèse a été plus supportable grâce aux groupes de travail, les repas à la cantine, les bières ensemble et le footing : Isao, Davide, Sarah, Alice, Alexandre, Alexis, Susana, Karen, Jean-David, Olga, Eric, Pierre-Antoine, Nicolas, Ahmed, Babacar, Flaminia, Adeline, Othmane, Lucas, Rancy, Thibaud, Guillermo, Fabio, Léa, Romain, Niccolò, Franck, Clément, Michel, Alexandra, Thibaud, Yating, Assaf...

La fin de cette thèse est aussi l'occasion de remercier mes amis parisiens : Monica, Leandro, Andrea R., Andrea Z., Pietro, Ricardo, Juan-Pablo, Alexandre, Jean Michel, Emiliano... Et aussi mes amis dans le reste du monde : Francesco, Simona, Fabio, Milo, Giacomo, Don Enzo, Carlo O., Riccardo... Vous êtes d'une compagnie précieuse

et je sais que je pourrai avoir toujours confiance en vous. Je remercie aussi Karine, Christel, Françoise, Nathalie, Gilles, Olivia, Valentina, Jessica, Giovanni, Leon et Gégé pour m'avoir appris à mettre en pratique l'Évangile dans la rencontre avec les pauvres. Un grand merci, enfin, à Eva pour m'avoir supporté pendant les derniers mois de la rédaction et pour m'avoir donné le sourire même dans les journées les plus compliquées.

L'ultimo paragrafo lo voglio dedicare alla mia famiglia. Un grazie profondo ai miei genitori Emilia e Maurizio e a mia sorella Camilla. Durante questi anni all'estero, lontano da Casalnoceto, siete stati coloro che mi sono stati più vicini in tutte le mie avventure su e giù per il mondo, con interminabili chiamate, messaggi e videochiamate. Dedico questa tesi a voi.

# Résumé

Dans cette thèse nous développons une théorie générale pour prouver l'existence de plusieurs formules de Itô sur l'équation de chaleur stochastique unidimensionnelle dirigée par un bruit blanc en espace-temps. Cela revient à définir de nouvelles notions d'intégrales stochastiques sur  $u$ , la solution de cette EDPS et à obtenir pour toute fonction assez lisse  $f$  des nouvelles identités impliquant  $f(u)$  et des termes de correction non triviaux. Ces nouvelles relations sont obtenues en utilisant la théorie des structures de régularité et la théorie des chemins rugueux. Dans le premier chapitre nous obtenons une identité intégrale et une différentielle impliquant la reconstruction de certaines distributions modélisées. Ensuite, nous discutons d'une formule générale de changement de variable pour tout chemins Hölderiens dans le contexte des chemins rugueux en le rapportant à la notion d'algèbres quasi-shuffle et à la famille des chemins rugueux dits quasi-géométriques. Enfin nous appliquons les résultats généraux sur les chemins rugueux quasi-géométriques à l'évolution temporelle du processus  $u$ . En utilisant le comportement gaussien de  $u$ , nous identifions la plupart des termes impliqués dans ces équations avec certaines constructions du calcul stochastique.

**Mot-Clés :** Formule d'Itô, Équations aux dérivées partielles stochastiques, Structures de régularité, Chemins rugueux, Algèbres quasi-shuffle, Calcul stochastique.





# Abstract

In this thesis we develop a general theory to prove the existence of several Itô formulae on the one dimensional stochastic heat equation driven by additive space-time white noise. That is denoting by  $u$  the solution of this SPDE for any smooth enough function  $f$  we define some new notions of stochastic integrals defined upon  $u$ , which cannot be defined classically, in order to deduce new identities involving  $f(u)$  and some non trivial corrections. These new relations are obtained by using the theory of regularity structures and the theory of rough paths. In the first chapter we obtain a differential and an integral identity involving the reconstruction of some modelled distributions. Then we discuss a general change of variable formula over any Hölder continuous path in the context of rough paths, relating it to the notion of quasi-shuffle algebras and the family of so called quasi-geometric rough paths. Finally we apply the general results on quasi-geometric rough paths to the time evolution of  $u$ . Using the Gaussian behaviour of the process  $u$ , most of the terms involved in these equations are also identified with some classical constructions of stochastic calculus.

**Keywords:** Itô formula, Stochastic partial differential equations, Regularity structures, Rough paths, Quasi-shuffle algebras, Stochastic calculus.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Itô formula in regularity structures . . . . .	5
1.2	Rough Itô formula . . . . .	8
1.3	Itô formulae on the time evolution . . . . .	11
<b>2</b>	<b>An Itô type formula for the additive stochastic heat equation</b>	<b>13</b>
2.1	Introduction . . . . .	13
2.2	Hölder spaces and Malliavin calculus . . . . .	17
2.3	Regularity structures . . . . .	21
2.3.1	Algebraic construction . . . . .	21
2.3.2	Models on a regularity structure . . . . .	27
2.3.3	The BPHZ renormalisation and the BPHZ model . . . . .	30
2.4	Calculus on regularity structures . . . . .	40
2.4.1	Modelled distributions . . . . .	40
2.4.2	Operations with the stochastic heat equation . . . . .	42
2.5	Itô formula . . . . .	49
2.5.1	Pathwise Itô formulae . . . . .	49
2.5.2	Identification of the differential formula . . . . .	52
2.5.3	Identification of the integral formula . . . . .	60
2.6	Renormalisation constants . . . . .	66
<b>3</b>	<b>Rough change of variable formulae and quasi-geometric rough paths</b>	<b>71</b>
3.1	Introduction . . . . .	71
3.2	Rough paths of any Hölder regularity . . . . .	73
3.2.1	Algebraic preliminaries . . . . .	73
3.2.2	Geometric and branched rough paths . . . . .	79
3.3	Controlled rough paths . . . . .	83
3.3.1	General definitions . . . . .	83
3.3.2	Integration and composition . . . . .	85
3.4	Change of variable formulae for rough paths . . . . .	91
3.4.1	Kelly's change of variable formula . . . . .	91
3.4.2	Symmetrical paths . . . . .	98
3.5	Quasi-geometric rough paths . . . . .	101
3.5.1	Quasi-shuffle algebras and quasi-geometric rough paths . . . . .	101
3.5.2	Building a symmetrical bracket extensions . . . . .	107

3.6	Quasi-geometric change of variable formula . . . . .	110
3.6.1	An alternative formula . . . . .	110
3.6.2	Extension of coordinates . . . . .	114
<b>4</b>	<b>Rough Itô formulae for the one dimensional heat equation</b>	<b>117</b>
4.1	Introduction . . . . .	117
4.2	One dimensional change of variable formula . . . . .	120
4.2.1	General results . . . . .	120
4.2.2	An existence criterion for rough paths . . . . .	124
4.3	Stochastic heat equation . . . . .	128
4.3.1	Burdzy-Swanson formula . . . . .	130
4.3.2	Cheridito-Nualart formula . . . . .	135
4.3.3	Quartic variation formula . . . . .	141
4.4	Stratonovich type formula . . . . .	145
4.4.1	A general identity between symmetric integrals . . . . .	146
4.4.2	A simplified proof of the Stratonovich formula . . . . .	152

# Chapter 1

## Introduction

This thesis is devoted to express several non-classical change of variable formulae, called in short *Itô formulae*, on a specific stochastic process.

Introduced in the seminal papers [Itô42], [Itô44], the original Itô's formula was written as an interesting remark of the general *Itô integration theory*. The main achievement of this theory was to extend the analytical theory of the Paley-Wiener integration to define an integration theory against a Brownian motion  $B$  (whose trajectories have a.s. infinite total variation), when the integrand is also a random process. In order to show that this new type of integral, *the Itô integral*, looks different from a classical Stieltjes integral, it has been remarked in [Itô44] that for any  $C^2$  function  $f: \mathbb{R} \rightarrow \mathbb{R}$  one has the a.s. identity

$$f(B_t) = f(B_s) + \int_s^t f'(B_r)dB_r + \frac{1}{2} \int_s^t f''(B_r)dr. \quad (1.0.1)$$

The relation (1.0.1) became famous in the literature as the *Itô formula*. The main consequence of this identity was precisely the possibility of integrating against  $B$  as if it was a bounded variation function in standard calculus, provided that we modify the classical formulae with an extra term containing higher order derivatives. Itô integrals and Itô formulae are at the basis of the theory of *stochastic calculus* and the ideas behind their formulation have been extended to a wide class of processes, leading to one of the most important mathematical theory of the twentieth century.

The stochastic process that we will consider in this dissertation is the one dimensional stochastic heat equation with additive space-time white noise. More precisely, this process is the solution  $u: \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$  of a one dimensional heat equation with a source term  $\xi: \mathbb{R}^2 \rightarrow \mathbb{R}$

$$\begin{cases} \partial_t u = \partial_{xx} u + \xi, \\ u(0, x) = u_0(x). \end{cases} \quad (1.0.2)$$

The main feature of this partial differential equation is that the source term  $\xi$  is a *space-time white noise*, namely a Gaussian random field on  $\mathbb{R}^2$  with zero mean and for any  $(t, x), (s, y) \in \mathbb{R}^2$  the following covariance structure

$$\mathbb{E}[\xi(t, x)\xi(s, y)] = \delta(t - s)\delta(y - x), \quad (1.0.3)$$

where  $\delta$  is a one dimensional Dirac delta function. In particular, equation (1.0.2) is one of the simplest examples of *stochastic partial differential equation* (SPDE).

Originated as an infinite dimensional extension of the stochastic calculus fifty years ago, the theory of SPDEs has been an increasingly active field of research at the intersection of probability theory, statistical mechanics and functional analysis. We refer the reader to the monographs [Wal84], [DPZ14] and [CR99] for a general introduction on the topic. Following the classical results of the theory (see e.g. [DPZ14, Thm. 5.4]) there exists a global weak solution  $u$  for any initial condition  $u_0$  solving the deterministic heat equation (see [Can84]). The linear structure of equation (1.0.2) allows also to solve it explicitly. Indeed, introducing the heat kernel on  $\mathbb{R}$

$$G_t(x) = \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{x^2}{4t}\right),$$

the solution  $u$  can be *formally* expressed by means of the variations of constants formula

$$u(t, x) = \int_{\mathbb{R}} G_t(x - y)u_0(y)dy + \int_0^t \int_{\mathbb{R}} G_{t-s}(x - y)\xi(s, y)dsdy, \quad (1.0.4)$$

as if  $\xi$  were a smooth function. We stress the fact that the above relation is only formal because the space-time white noise  $\xi$  cannot be evaluated pointwise. The only way to rigorously interpret the action of  $\xi$  inside the integral can be done by defining it as an isonormal Gaussian process on the Hilbert space  $H = L^2(\mathbb{R}^2)$ , i.e.  $\xi$  is seen as a collection of centred Gaussian random variables  $\xi = \{\xi(f)\}_{f \in H}$  verifying for any  $f, g \in H$

$$\mathbb{E}[\xi(f)\xi(g)] = \langle f, g \rangle_H. \quad (1.0.5)$$

Rewriting the random variable  $\xi(f)$  as  $\int_{\mathbb{R}^2} f(s, y)\xi(ds, dy)$  and using the integrability property of  $G$ , identity (1.0.4) can be then *rigorously* written as

$$u(t, x) = \int_{\mathbb{R}} G_t(x - y)u_0(y)dy + \int_{\mathbb{R}^2} \mathbf{1}_{[0, t]}(s)G_{t-s}(x - y)\xi(ds, dy). \quad (1.0.6)$$

Since the stochastic behaviour of the solution is only encoded in the term containing  $\xi$ , it is common to set  $u_0 = 0$ . The resulting process thus becomes an explicit example of *stochastic convolution* (see [DPZ14, Sec 5.1.2]). Note that we can carry out a similar procedure to compute the unique global weak solution of equation (1.0.2) with Dirichlet or periodic boundary conditions, by simply replacing  $G$  in expression (1.0.6) with the corresponding heat kernel.

The explicit form of the process  $u$  has lead to an intensive study of its properties since the beginning of the SPDEs theory. Two main approaches have been developed. On one hand we can consider the process  $u$  as a time evolution  $u = \{u(t, \cdot)\}_{t \geq 0}$  on an infinite dimensional space such as  $L^2(\mathbb{R})$  or  $C(\mathbb{R})$ , the space of continuous function on  $\mathbb{R}$ . Under this perspective  $\xi$  becomes an infinite dimensional Brownian motion, a *cylindrical Wiener process*  $dW_t$  (see [DPZ14]), and  $u$  is seen as the solution of a stochastic differential equation with values in an infinite dimensional space

$$\begin{cases} du = (\partial_{xx}u)dt + dW_t \\ u(0, x) = u_0(x). \end{cases} \quad (1.0.7)$$

In particular the solution  $u$  is a *Markov process* whose properties can be studied by means of the standard tools of semigroup theory (see [DPZ14]). For instance The hitting properties of independent copies of  $u$  were studied in [MT02] and an explicit description of its generator in the  $L^2$  setting has been provided in [DP99]. We also recall that the corresponding Markov process associated the solution of the equation (1.0.7) with homogeneous Dirichlet boundary conditions on  $[0, 1]$  admits as unique invariant measure the law of a standard Brownian bridge (see e.g. [Zam17, Chap. 4]). This process represents then a rigorous mathematical model of *random interface* because it can be also obtained as the scaling limit of a class of interacting particle systems, as explained in [EL15]. Others results in the same spirit were obtained in [Fun83] for the solution of (1.0.7) with other boundary conditions.

In alternative to the infinite dimensional setting, the process  $u$  can be seen also as a real-valued *Gaussian random field*  $u = \{u(t, x)\}_{t \geq 0, x \in \mathbb{R}}$ . This approach dates back to the seminal lecture-notes [Wal84] and we used it to introduce the process  $u$  in (1.0.6). Even if this point of view looks simpler from a purely probabilistic point of view because it does not involve semigroup theory, it still requires an extensive use of functional analysis. Indeed looking back at the definition of  $\xi$  in (1.0.5), an alternative way of introducing it is can be made by defining  $\xi$  as a *random distribution*, i.e. a random variable taking value in  $\mathcal{D}'(\mathbb{R}^2)$ , the set of distributions on  $\mathbb{R}^2$  and in principle we might expect a very technical theory. However in the specific case of the process  $u$ , the properties of  $\xi$  allow to compute explicitly the covariance of the stochastic convolution and we deduce that the trajectories of  $u$  are “almost”  $1/4$ -Hölder continuous in time and  $1/2$ -Hölder continuous in space. An elementary explication of this regularity property can be found in [Zam17], [MT02]. From an analytical point of view, this result can be interpreted by saying that the regularising effect of the heat kernel turns the distributional trajectories of  $\xi$  into a continuous function (we recall that the deterministic part of the process is smooth for any positive time as a consequence of the classical theory). The explicit Hölder exponents appearing for  $u$  can be also justified by using a parabolic version of the Schauder estimates (see e.g. [Kry96]) on the realisations of  $\xi$ .

One of the big disadvantages in the study of  $u$  is then the absence of the classical tools of stochastic calculus on its trajectories. Indeed the time evolution  $\{u_t\}_{t \geq 0}$  is not an infinite dimensional semimartingale but for any smooth compactly supported function  $h: \mathbb{R} \rightarrow \mathbb{R}$  the process

$$\langle u_t, h \rangle := \int_{\mathbb{R}} u(t, y) h(y) dy$$

is a real-valued semimartingale. Moreover we can fix a space or a time point  $x \in \mathbb{R}$ ,  $t \geq 0$  and the resulting evolutions  $s \rightarrow u(s, x)$ ,  $y \rightarrow u(t, y)$  are continuous processes which are not semimartingales. The main reason behind this ill-posedness comes by writing  $u$  as the solution of (1.0.7) where we immediately note that this process have the same quadratic variation of the space time white noise  $dW_t$  which is naturally an *infinite* quantity. An explicit computation of the covariance of  $u$  in [Swa07] shows also that the time evolution  $s \rightarrow u(s, x)$  is not a Markov process and it has infinite quadratic



variation. Since the powerful theory of the Itô stochastic calculus cannot be applied directly to  $u$ , many fine properties of its trajectories, such as an intrinsic theory of local times associated to a Tanaka formula or the study of its excursions, cannot be studied as in the Brownian case (see e.g. [RY04]) and there are still no general theories to approach these problems. Over the years few attempts were done to define some new tools to define a new stochastic calculus on  $u$  using for instance the Malliavin calculus (see [GNT05]) or the theory of local times for Gaussian random fields (see [Zam06]) but these results were not sufficiently investigated by the scientific community. More recently in the article [EZ18] the authors provided an intrinsic notion of local times based on the theory of Dirichlet forms (see e.g. [FOT11]) to extend the family of Bessel SPDEs (see [Zam01]) up to some critical parameters.

During the past few years the whole discipline of stochastic partial differential equations was radically innovated by the emergence of the theory of *regularity structures* (see [Hai14]). The main goal of this new mathematical theory was to establish some rigorous well posedness results for *singular* SPDEs, few important classes of equations present in the statistical physics literature but whose meaning was still a challenge for the mathematical community. Many of these equations are constructed as a non-linear perturbation of the  $d$ -dimensional extension of the equation (1.0.2). For instance the well-known KPZ equation (see [KPZ86]) writes formally as the equation (1.0.2) with a quadratic perturbation

$$\partial_t h = \partial_{xx} h + (\partial_x h)^2 + \xi. \quad (1.0.8)$$

Since we expect  $h$  to share the same space regularity as  $u$ , there is no classical analytical theory to define  $(\partial_x h)^2$ , the square of the distributional term  $\partial_x h$ . In order to solve equation (1.0.8) the theory mixes several new results on the analysis of PDEs, algebra and stochastic calculus (see [BHZ19], [CH16], [BCCH17]) to achieve a rigorous definition for the term  $(\partial_x h)^2$  and the solution  $h$  (See also [Hai13] for further details). Even if this equation is by far more complex than (1.0.2) we still remark that the process  $u$  plays a fundamental role in the study of the solution because it approximates locally the random field  $h$  in the same way as polynomials approximates smooth functions. The study of the singular SPDEs has also generated other theories to study the same family of equation with other analytical tools. Even in those context the knowledge of  $u$  is still fundamental to give a rigorous meaning to these equations. We recall for instance the theory of paracontrolled distributions [GIP15].

The theory of regularity structures should not be seen as an isolated phenomenon in the probability panorama but instead the culminating point of an another theory, also at the intersection between analysis and stochastic processes, the theory of *rough paths* (see [Lyo98], [Gub04]). Basically speaking this theory extends the Riemann-Lebesgue integration theory to perform the *rough integration*, a deterministic integration theory against a generic non smooth continuous path. One of the main consequence is that in some cases it is possible to interpret the Itô stochastic integral as the path-wise realisation of a specific *rough integral*. Similarly we also obtain that the solution of some an Itô diffusion can be seen as the path-wise realisation of a so called *rough differential equation*, a deterministic evolution driven by this rough integration. The combined techniques of rough paths and regularity structures have deeply transformed

the field of stochastic calculus showing interesting applications and connections with many other branches of applied and pure mathematics (see [FH14]).

The main scope of our thesis is then to define new integration theories upon the stochastic convolution process  $u: \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ , that is the explicit solution of

$$\begin{cases} \partial_t u = \partial_{xx} u + \xi \\ u(0, x) = 0 \end{cases} \quad (1.0.9)$$

using the techniques developed in the theory of rough paths and regularity structures. Moreover in the same spirit of Itô's first researches, for any new notion of integral we will investigate the associated change of variable formulae. The choice of starting from a zero initial condition is simply done because we are interested in studying the stochastic behaviour of the solution and as shown in formula (1.0.6) in order to recover the general results on (1.0.2) it is sufficient to add a deterministic smooth function.

The thesis is divided into three different chapters, each associated with a scientific paper [Bel18] [Bel19a] [Bel19b] (submitted or in preparation) dedicated to a particular aspect of this problem. In particular the first chapter concentrates on the regularity structure approach and the remaining chapters are devoted to a deeper comprehension of the change of variable formula in the rough path context. We present here a quick overview of the main contributions contained in each chapter.

## 1.1 Itô formula in regularity structures

The first chapter will be entirely focused on the theory of regularity structures. In particular we will explain how a theory designed to resolve equations like the KPZ equation in (1.0.8) can contribute to express also a change of variable formula on  $u$  which is classically well-posed. The fundamental concept behind this connection is that for any sufficiently regular function  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  the random field  $\varphi(u)$ , which is completely well defined by itself, it is in turn the *solution* of a singular SPDE similar to (1.0.8). How can we deduce such equation? Working as if  $\xi$  were a smooth function, we obtain formally from the chain rule

$$\partial_t(\varphi(u)) = \varphi'(u)\partial_t u, \quad \partial_{xx}(\varphi(u)) = \varphi''(u)(\partial_x u)^2 + \varphi'(u)\partial_{xx} u, \quad (1.1.1)$$

Thus if  $u$  satisfies the stochastic heat equation (1.0.9), we obtain immediately that  $\varphi(u)$  satisfies the new relation

$$\begin{cases} \partial_t v = \partial_{xx} v + \varphi'(u)\xi - \varphi''(u)(\partial_x u)^2 \\ v(0, x) = \varphi(0). \end{cases} \quad (1.1.2)$$

By the variation of constants formula on the second equation, we deduce from the equation (1.1.2) a formal expression for  $\varphi(u)$

$$\begin{aligned} \varphi(u(t, x)) = & \varphi(0) + \int_0^t \int_{\mathbb{R}} G_{t-s}(x-y)\varphi'(u(s, y))\xi(s, y)dsdy \\ & - \int_0^t \int_{\mathbb{R}} G_{t-s}(x-y)\varphi''(u(s, y))(\partial_x u)^2(s, y)dsdy. \end{aligned} \quad (1.1.3)$$

Clearly the reasoning behind the chain rules in (1.1.1) is only this formal because the process  $u$  is a.s. not differentiable but this is not a simple problem of regularity of functions because the terms  $\varphi'(u)\xi$ ,  $\varphi''(u)(\partial_x u)^2$  appearing in (1.1.2) and (1.1.3), in analogy with the equation (1.0.8), do not have any rigorous meaning. The main reason for which these terms are ill posed is because they involve formally products of  $\xi$  and  $\partial_x u$ , terms which can be defined only as random variables with values in the space distribution  $\mathcal{D}'(\mathbb{R}^2)$  (we recall that  $u$  is non-differentiable in space).

Since the beginning of the theory of distributions, the problem of finding a notion of product has been a crucial one in mathematical analysis (see [Sch54]). Explaining the functional spaces where we can actually perform a product for distributions has produced many interesting theories in mathematical analysis such as the paraproduct theory (see [BCD11]) or the Colombeau algebra (see [Col84]). Thus we might expect to interpret these products using a technical analytical machinery. Actually it has been shown in [Hai14] that both  $\xi$  and  $\partial_x u$  belong a.s. to some functional spaces where there exists no continuous bilinear map to express the terms  $\varphi'(u)\xi$  and  $\varphi''(u)(\partial_x u)^2$ , showing that this operation cannot be attacked with the known analytical theories.

One of the great achievements obtained in the theory of regularity structures is then to build a coherent mathematical framework to deal with such expressions. The main idea behind the whole theory is then to express a random distribution  $v$ , in terms of a universal family of *generalised stochastic polynomials* describing locally  $v$  as polynomials describe smooth functions. Vaguely speaking this is equivalent to say that there exists a family of random distributions  $\{\Pi_x \tau(y)\}_{\tau \in T}$ , depending on a reference point  $x \in \mathbb{R}^2$ , a variable  $y \in \mathbb{R}^2$ , a fixed probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a finite set  $T$  such that for any random distribution  $v$  we can write formally

$$v(y) - v(x) = \sum_{i \in I} a_i(x) \Pi_x \tau_i(y) + r(y, x), \quad \text{a.s.} \quad (1.1.4)$$

where  $I$  is a finite subset  $I \subset T$ ,  $\{a_i\}_{i \in I}$  is a family of random fields process playing the same role as the *derivatives* and  $r$  is a random *remainder* term which is sufficiently small with respect to some metric. The equation (1.1.4) is still a formal identity because it involves the evaluation of distributions on points but it can rigorously be written with some new definitions. Following the formula (1.1.4) we can say that the random distribution  $v$  is approximated by the triplet  $V = (A, \Theta, I)$  where  $A$  and  $\Theta$  are defined respectively as

$$A(x) := (a_1(x), \dots, a_{|I|}(x)), \quad \Theta := \{\tau_i\}_{i \in I}. \quad (1.1.5)$$

One of the fundamental results expressed in [Hai14] is that if we dispose of a family of generalised stochastic polynomials  $\{\Pi_x \tau(y)\}_{\tau \in T}$  (called *model*) then there exists a family triplets  $V = (A, \Theta, I)$  (called *modelled distribution*) such that for any triplet belonging to this class we can associated a unique random distribution to it denoted by  $\mathcal{R}(V)$  (called *the reconstruction of V*) satisfying the identity (1.1.4). This result, which is referred in the theory as the *reconstruction theorem*, is a cornerstone of the theory, characterising a random distribution in terms of a local expansion of a finite number of elements. Under this perspective we can define a product between two random distributions.

We consider two random distribution  $v^1, v^2$  satisfying  $v^1 = \mathcal{R}(V^1)$  and  $v^2 = \mathcal{R}(V^2)$  for two fixed modelled distributions  $V^1 = (A^1, \Theta^1, I^1)$ ,  $V^2 = (A^2, \Theta^2, I^2)$  where  $A^1, A^2, \Theta^1, \Theta^2$  are written like (1.1.5). Then if we are able to define a finite number of new generalised stochastic polynomial  $\Pi_x(\tau_i^1 \tau_j^2)(y)$  on a specific set of indices we can define an abstract notion of product between the modelled distribution  $V^1 V^2$  and the product between  $v^1$  and  $v^2$  becomes the random distribution  $\mathcal{R}(V^1 V^2)$ , that is the unique random distribution satisfying

$$\mathcal{R}(V^1 V^2)(y) - \mathcal{R}(V^1 V^2)(x) = \sum_{(i,j)} a_i^1(x) a_j^2(x) \Pi_x(\tau_i^1 \tau_j^2)(y) + \tilde{r}(y, x) \quad \text{a.s.} \quad (1.1.6)$$

where  $\tilde{r}$  is the stochastic remainder. Even if we still keep the discussion at an informal level, we stress the fact the the resulting product operation  $(v^1, v^2) \rightarrow \mathcal{R}(V^1 V^2)$  provides a measurable operation with some properties of a product but its definition requires more properties than the simple theory of distributions.

The first chapter of this thesis will be devoted to justify in detail the equation defining  $\varphi(u)$  (1.1.2) and the integral formula (1.1.3) in the context of this new notion of product between distributions sketched above. For technical reasons the starting process  $u$  will be the solution of (1.0.9) with periodic boundary condition, written as a random field  $u: \mathbb{R}_+ \times \mathbb{T} \rightarrow \mathbb{R}$ , where  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$  is the one dimensional torus. First we will introduce an explicit family of generalised stochastic polynomials which is able to describe a solution of (1.1.2). The existence of such family will be obtained by specifying and simplifying some technical results of the theory contained in [BHZ19], [CH16]. Once obtained the generalised stochastic polynomials, we will show explicitly four different modelled distributions  $\Phi'(U)$ ,  $\Phi''(U)$ ,  $\Xi$  and  $\partial_x U$  satisfying

$$\mathcal{R}(\Phi'(U)) = \varphi'(u), \quad \mathcal{R}(\Phi''(U)) = \varphi''(u), \quad \mathcal{R}(\partial_x U) = \partial_x u, \quad \mathcal{R}(\Xi) = \xi. \quad (1.1.7)$$

From the knowledge of these modelled distributions and the underlying model we will obtain that it is also possible to define the modelled distributions  $\Phi'(U)\Xi$  and  $\Phi''(U)(\partial_x U)^2$ , giving a rigorous meaning to the terms  $\varphi'(u)\xi$ ,  $\varphi''(u)(\partial_x u)^2$ .

The remarkable property of this construction appears then when we want to link this definition with the differential identity (1.1.2). Indeed for any sufficiently smooth  $\varphi$  a key result of the chapter is the *rough differential Itô formula*

$$(\partial_t - \partial_{xx})\varphi(u) = \mathcal{R}(\Phi'(U)\Xi) - \mathcal{R}(\Phi''(U)(\partial_x U)^2), \quad (1.1.8)$$

where both sides are interpreted a.s. as distributions. Therefore this new notion of product satisfies formally the same identity of standard calculus, provided that we interpret the terms in a distributional way. Writing the expression of (1.1.8) using the variation of the constants formula for the heat equation with a distributional source, we obtain an a.s. expression for  $\varphi(u)$  called *rough integral Itô formula*:

$$\varphi(u(t, x)) = P * \mathbf{1}_{[0,t]} \mathcal{R}(\Phi'(U)\Xi)(t, x) - P * \mathbf{1}_{[0,t]} \mathcal{R}(\Phi''(U)(\partial_x U)^2)(t, x), \quad (1.1.9)$$

where  $P$  is the heat Kernel on the torus, the operation  $*$  is the convolution operation on  $\mathbb{R}^2$  and  $\mathbf{1}_{[0,t]}$  is an operator between distributions extending the multiplication with

the indicator function of  $[0, t] \times \mathbb{T}$ . Similarly as before the identity (1.1.9) is a rigorous way to interpret (1.1.3).

The identities (1.1.8), (1.1.9) explain the behaviour of  $\varphi(u)$  as a solution of a partial differential equation whereas they do not study the probabilistic meaning of the random distributions  $\mathcal{R}(\Phi'(U)\Xi)$ ,  $\mathcal{R}(\Phi''(U)(\partial_x U)^2)$ . The last part of the chapter will be dedicated to study for any test function  $\psi: \mathbb{R}_+ \times \mathbb{T} \rightarrow \mathbb{R}$  the behaviour of the random variables  $\mathcal{R}(\Phi'(U)\Xi)(\psi)$ ,  $\mathcal{R}(\Phi''(U)(\partial_x U)^2)(\psi)$  using the explicit Gaussian structure of the process  $u$  and the space time white noise  $\xi$ , as introduced in work [Zam06]. In particular for any test function  $\psi: \mathbb{R}_+ \times \mathbb{T} \rightarrow \mathbb{R}$  whose support is contained in  $[0, t] \times \mathbb{T}$  compactly we will obtain the following *identification theorems*

$$\mathcal{R}(\Phi'(U)\Xi)(\psi) = \int_0^t \int_{\mathbb{T}} \varphi'(u(s, y)) \psi(s, y) dW_{s, y} . \quad (1.1.10a)$$

$$\begin{aligned} \mathcal{R}(\Phi''(U)(D_x U)^2)(\psi) &= -\frac{1}{2} \int_0^t \int_{\mathbb{T}} \psi(s, y) \varphi''(u(s, y)) C(s) dy ds \quad (1.1.10b) \\ &+ \int_{[0, t]^2 \times \mathbb{T}^2} \left[ \int_{s_2 \vee s_1}^t \int_{\mathbb{T}} \psi(s, y) \varphi''(u(s, y)) \partial_x P_{s-s_1}(y - \mathbf{y}_1) \partial_x P_{s-s_2}(y - \mathbf{y}_2) dy ds \right] dW_{s, y}^2 , \end{aligned}$$

where the integral  $dW_{s, y}$  is a Walsh integral taken with respect to the martingale measure associated to  $W$ , the brownian sheet associated to  $\xi$  (see [Wal84]);  $C: (0, T) \rightarrow \mathbb{R}$  is the function

$$C(s) := \int_{\mathbb{T}} P(s, y)^2 dy \quad (1.1.11)$$

and the integral  $dW_{s, y}^2$  is the multiple Skorohod integral of order two associated to  $W$  (see the general notion of Skorohod integral of order two in [NN10]). The content of the chapter is based upon the article [Bel18].

## 1.2 Rough Itô formula

Differently from the main topic of the thesis, in the second chapter we will not speak directly of the process  $u$ . Indeed the whole discussion will focus on the generalisation of the Itô formula (1.0.1) using the rough paths theory.

Introduced at the end of the nineties in the seminal paper [Lyo98], the theory of rough path was built to solve the following family of controlled differential equations

$$\begin{cases} dY_t = f(Y_t) dX_t \\ Y_0 = y_0 , \end{cases} \quad (1.2.1)$$

where  $Y$  and  $X$  are continuous functions  $Y: [0, T] \rightarrow \mathbb{R}^e$ ,  $X: [0, T] \rightarrow \mathbb{R}^d$  and  $f: \mathbb{R}^e \rightarrow \mathbb{R}^{e \times d}$  is smooth. The key point in equation (1.2.1) is that we want define a notion of solution even when the trajectories of the *driving path*  $X$  are “rough”, that is neither differentiable nor absolute continuous. Even if the formulation of this problem is purely deterministic, most of its results are then applied in the probabilistic setting by considering the realisation of some continuous stochastic process  $X$  as driving path (e.g.

a fractional Brownian motion). Therefore this theory can be used to develop a robust solution theory for stochastic differential equations which cannot be defined classically.

The usual way to formulate the problem (1.2.1) is in its integral form. Hence a large part of the theory is dedicated to give a rigorous meaning to integrals such as

$$\int_s^t f(X_r) dX_r \quad (1.2.2)$$

where  $X: [0, T] \rightarrow \mathbb{R}$  is a rough function with Hölder regularity  $\gamma \in (0, 1)$  and  $f: \mathbb{R} \rightarrow \mathbb{R}$  is smooth. As expressed at the beginning, this problem motivated the definition of the Itô calculus when  $X$  is a Brownian motion. However the main feature of this rough path point of view is then to provide an analytical way to define the integral (1.2.2) independently of the stochastic properties of the trajectories. A first interesting solution to this problem has been provided by a result of W. Young [You36], defining (1.2.2) as the limit

$$\int_s^t f(X_r) dX_r := \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} f(X_u)(X_v - X_u), \quad (1.2.3)$$

where  $\mathcal{P}_n$  is a generic sequence of partitions of  $[s, t]$  whose mesh-size converges to 0. One of the big obstacles of this result is that this approximation holds if and only if the parameter  $\gamma \in (1/2, 1)$ , excluding the possibility of defining a similar object in the Brownian case (whose trajectories are almost  $1/2$ -Hölder). The fundamental progress introduced in [Lyo98] and explained in a simpler formalism by [Gub04] was that even when  $\gamma < 1/2$  it is still possible to define the operation (1.2.2) through a limiting procedure. Nevertheless instead of a classical Riemann sum as in the Young case (1.2.3) the approximations of this operation depend on a larger path  $\mathbf{X}$  containing the driving path  $X$  and a finite number of extra components that behave formally as the iterated integrals of  $X$ . This new object  $\mathbf{X}$  takes the name of *rough path over  $X$* , and we can formally define the integral (1.2.2) as a *rough integral* as long as we can define a rough path  $\mathbf{X}$  over the driving path  $X$ . To stress the fact that this integral does not depend uniquely on  $X$  we will denote it by  $\int_s^t f(X_r) d\mathbf{X}_r$ .

The general definition of a rough path  $\mathbf{X}$  when the driving path is  $\gamma$ -Hölder for  $\gamma \in [1/3, 1/2)$  is at the basis of the monograph [FH14]. However when we want to treat the general case  $\gamma \in (0, 1)$  two main notions of rough path are defined in the literature to give a rigorous meaning to the rough integral: the *geometric rough paths* and the *branched rough paths*. The first family dates back to the first formulation of the theory and it imposes that the extra components of a geometric rough path (the objects which replace the iterated integral) should respect the usual integration by parts formulae rule of the iterated integral of smooth paths. One disadvantage of this definition is that we cannot recover the Itô integration theory in this family. The other class was introduced in [Gub10] and using it we can easily include the stochastic calculus as a particular case but the cost of this inclusion is the presence of more components than for a geometric rough path. A detailed comparison of these two families has been discussed [HK15] showing also some interesting connections with the theory of Hopf algebras.

The presence of two different methods to construct the rough integral leads naturally to study the problem of the change of variable formula. Particularly if  $X = (X^1, \dots, X^d): [0, T] \rightarrow \mathbb{R}^d$  is a generic  $\gamma$ -Hölder path contained in a rough path  $\mathbf{X}$  and  $\varphi: \mathbb{R}^d \rightarrow \mathbb{R}$  is a sufficiently smooth function, does the change of variable formula

$$\varphi(X_t) = \varphi(X_s) + \sum_{i=1}^d \int_s^t \partial_i \varphi(X_r) d\mathbf{X}_r^i \quad (1.2.4)$$

still hold? On one hand if  $\mathbf{X}$  is geometric, the formula (1.2.4) is true and it follows directly from the definition of a geometric rough path. On the other hand the same problem shows a more interesting behaviour when  $\mathbf{X}$  is a branched rough path. This topic has been studied in the PhD thesis of David Kelly [Kel12] and a new formula was obtained in this context by means of the notion of *bracket extension*  $\hat{\mathbf{X}}$  over a branched rough path  $\mathbf{X}$ . Loosely speaking a bracket extension is again a branched rough path containing the information given by  $\mathbf{X}$  and in addition to that it encodes also some new paths  $X^{(i_1 \dots i_n)}$  for  $i_1, \dots, i_n = 1, \dots, d$  called  $(i_1 \dots i_n)$ -order variations, which extend in an algebraic and analytical context the notion of quadratic variation of a semimartingale. By fixing a bracket extension  $\hat{\mathbf{X}}$  over a branched rough path  $\mathbf{X}$ , we obtain the general formula

$$\begin{aligned} \varphi(X_t) = \varphi(X_s) &+ \sum_{i=1}^d \int_s^t \partial_i \varphi(X_r) d\mathbf{X}_r^i \\ &+ \sum_{n=2}^N \sum_{i_1, \dots, i_n=1}^d \frac{1}{n!} \int_s^t \frac{\partial^{i_1} \dots \partial^{i_n} \varphi(X_r)}{\partial x_{i_1} \dots \partial x_{i_n}} d\hat{\mathbf{X}}_r^{(i_1 \dots i_n)}, \end{aligned} \quad (1.2.5)$$

where  $N = \lfloor 1/\gamma \rfloor$  and the other terms appearing in (1.2.5) are rough integrals with respect to the paths  $X^{(i_1 \dots i_n)}$ . We call the identity (1.2.5) *rough Itô formula* because it represents a deterministic generalisation of the stochastic identity (1.0.1). Looking at this theory from a probabilistic point of view, the formula (1.2.5) establishes an interesting starting point to study non standard change of variable formulae on general stochastic processes with continuous trajectories. Unfortunately it turns out that the notion itself of bracket extension does not seem to be adapted for applications to stochastic processes because of its intrinsic algebraic complexity.

The main improvement obtained in chapter two is then a new reformulation of Kelly's original theory in order to establish for a generic  $\gamma$ -Hölder path  $X$  a formula similar to (1.2.5) but using a simpler algebraic structure to encode the  $(i_1 \dots i_n)$ -order variations. Specifically we will introduce a new notion of rough paths called *quasi-geometric rough paths* and we will restate the formula in this new context. This new family of rough paths was recently introduced in the paper [BCEF18] in a much abstract context and it relies on the theory of quasi-shuffle algebras (see [Hof00], [HI17]). Basically speaking a quasi-geometric rough path  $\mathbb{X}$  over  $X$  encodes in a single object the same information as a bracket extension but using a different indexing of the coordinates. Indeed of starting from the paths  $X^{(i_1 \dots i_n)}$  for  $i_1, \dots, i_n = 1, \dots, d$  we will represent them in terms of multi-indices  $X^k$  where  $k = (k_1, \dots, k_d)$  and using  $\{e_i\}_{i=1}^d$

the canonical basis in  $\mathbb{N}^d$  we will adopt the identification  $X^i = X^{e_i}$ . These new paths are called  $k$ -order variations. Using this new notation for any fixed quasi geometric rough path  $\mathbb{X}$  over  $X$  the new change of variable formula reads as

$$\varphi(X_t) = \varphi(X_s) + \sum_{k \in \mathbb{N}^d: |k| \leq N} \frac{1}{k_1! \cdots k_d!} \int_s^t \frac{\partial^{k_1} \cdots \partial^{k_d} \varphi(X_r)}{\partial x_1 \cdots \partial x_d} d\mathbb{X}_r^k, \quad (1.2.6)$$

where the value  $|k|$  corresponds with the absolute value of the multi-index  $k$ .

One of the big advantages of the class of geometric rough paths is the existence of an *explicit* one to one correspondence between them and the geometric rough paths over the *extended path* of the  $k$ -order variations. Therefore the problem of constructing a quasi-geometric rough path can be reformulated as the construction of a geometric rough path over a higher dimensional path and in the stochastic case we can apply some standard criteria from the literature [FV10b] to show the existence of explicit quasi-geometric rough paths. This chapter is based upon the article [Bel19a].

### 1.3 Itô formulae on the time evolution

In the last chapter of the thesis we will apply the results on the rough Ito formula to deduce some new Itô formulae on the *time evolution* of the stochastic heat equation.

Recalling the regularity of the space and the time trajectories of the solution of (1.0.9), both the trajectories  $s \rightarrow u(s, x)$ ,  $y \rightarrow u(t, y)$  are Gaussian processes which should fit naturally in the rough path theory. However all the known work in the literature focuses only on the *spatial evolution* mostly because of this process has Hölder regularity almost  $1/2$ . For a detailed discussion of the applications of rough paths theory to this process we refer the reader to [Hai11], [HW13], [FGGR16] and the general monograph [FH14].

The process we will consider in this chapter is defined from a slight modification of the usual stochastic heat equation defined in (1.0.9) because we consider

$$\begin{cases} \partial_t u = \frac{1}{2} \partial_{xx} u + \xi \\ u(0, x) = 0. \end{cases}$$

Then for some fixed  $x \in \mathbb{R}$  we set  $X_t := u(t, x)$ . Since  $X$  is a real valued process whose Hölder regularity is almost  $1/4$ , the rough change of variable formula (1.2.6) applied to any random quasi-geometric rough path  $\mathbb{X}$  over  $X$  implies the identity

$$\begin{aligned} \varphi(X_t) = \varphi(X_s) &+ \int_s^t \varphi'(X_r) d\mathbb{X}_r^1 + \frac{1}{2!} \int_s^t \varphi''(X_r) d\mathbb{X}_r^2 \\ &+ \frac{1}{3!} \int_s^t \varphi^{(3)}(X_r) d\mathbb{X}_r^3 + \frac{1}{4!} \int_s^t \varphi^{(4)}(X_r) d\mathbb{X}_r^4, \end{aligned} \quad (1.3.1)$$

where every integral in the right hand side of (1.3.1) is the a.s. realisation of a rough integrals. In analogy with the stochastic calculus, where there are several ways to



express the same change of variable formula using different notions of stochastic integral (e.g. the Itô and Stratonovich formulation) we will show the existence of four different change of variable formulae on  $X$  by proving the existence of four random quasi-geometric rough paths over the process. Once obtained an analytical description of these identities, as we discussed in the first section, we will recover the probabilistic nature of some rough integrals by identifying them with some classical objects from stochastic calculus.

Summing up both the analytic and the probabilistic results, the resulting formulae are the following: first we have the *geometric formula*

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \varphi'(X_r) d\mathbf{X}_r^1, \quad (1.3.2)$$

obtained by choosing  $\mathbf{X}$ , the unique one-dimensional geometric rough path over  $X$  (which can be identified to a trivial quasi-geometric extension over  $X$ ). Then we construct two quasi-geometric rough paths  $\mathfrak{X}$  and  $\hat{\mathfrak{X}}$  such that the third and the fourth order rough integrals on the right hand side of (1.3.1) are identically zero and the second order ones are identified to some classical objects. Plugging them in (1.3.1) we obtain the identities

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \varphi'(X_r) d\mathfrak{X}_r^1 + \frac{\kappa}{2} \int_s^t \varphi''(X_r) dB_r, \quad (1.3.3)$$

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \varphi'(X_r) d\hat{\mathfrak{X}}_r^1 + \frac{1}{2} \int_s^t \varphi''(X_r) d\sigma_r^2, \quad (1.3.4)$$

where  $B$  is an independent Brownian motion,  $\sigma^2$  is the variance function  $\sigma_r^2 = \mathbb{E}[X_r^2]$  and the integrals in the formulae are an Itô integral and a Stieltjes one. The identities (1.3.3) and (1.3.4) are called the *Burdzy-Swanson formula* and the *Cheridito-Nualart formula* [BS10] [CN05]. Finally we prove the existence of a quasi-geometric rough path  $\mathcal{X}$  such that the fourth order integral is a Lebesgue integral

$$\begin{aligned} \varphi(X_t) = & \varphi(X_s) + \int_s^t \Phi'(X_r) d\mathcal{X}_r^1 + \frac{1}{2!} \int_s^t \Phi''(X_r) d\mathcal{X}_r^2 \\ & + \frac{1}{3!} \int_s^t \Phi^{(3)}(X_r) d\mathcal{X}_r^3 + \frac{6}{4! \pi} \int_s^t \varphi^{(4)}(X_r) dr, \end{aligned} \quad (1.3.5)$$

thereby proving the existence of a deterministic “quartic variation” for the process  $X$ . Moreover the sum of the remaining rough integrals in the right hand side of (1.3.5) can be also associated with the limit of a compensated Riemann sum related to the higher order Föllmer calculus, introduced in [CP19]. This chapter is based upon the article [Bel19b].

# Chapter 2

## An Itô type formula for the additive stochastic heat equation

### 2.1 Introduction

We consider  $\{u(t, x) : t \in [0, T], x \in \mathbb{T} = \mathbb{R}/\mathbb{Z}\}$  the solution of the additive stochastic heat equation with periodic boundary conditions and zero initial value:

$$\begin{cases} \partial_t u = \partial_{xx} u + \xi, \\ u(0, x) = 0 \\ u(t, 0) = u(t, 1) \\ \partial_x u(t, 0) = \partial_x u(t, 1) \end{cases} \quad (2.1.1)$$

where  $\xi$  is a real valued space-time white noise over  $\mathbb{R} \times \mathbb{T}$ . This equation was originally formulated to model a one dimensional string exposed to a stochastic force (see [Fun83]). From a theoretical point of view, the equation (2.1.1) represents one of the simplest examples of a stochastic PDE whose solution can be written explicitly, the so called *stochastic convolution* (see e.g. [Wal84], [DPZ14]). Writing  $\xi = \partial W / \partial t \partial x$ , where  $W$  is the Brownian sheet associated to  $\xi$ , one has

$$u(t, x) = \int_0^t \int_{\mathbb{T}} P_{t-s}(x - y) dW_{s,y}, \quad (2.1.2)$$

where the integral  $dW_{s,y}$  is a Walsh integral taken with respect to the martingale measure associated to  $W$  and  $P : (0, +\infty) \times \mathbb{T} \rightarrow \mathbb{R}$  is the fundamental solution of the heat equation with periodic boundary conditions:

$$P_t(x) = \sum_{m \in \mathbb{Z}} G_t(x + m), \quad G_t(x) = \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{x^2}{4t}\right).$$

It is well known in the literature (see e.g. [DPZ14]) that almost surely  $u$  admits a continuous modification in both variables  $t$  and  $x$  and it satisfies the equation (2.1.1) in a weak sense, that is for any smooth function  $l : \mathbb{T} \rightarrow \mathbb{R}$  one has

$$\int_{\mathbb{T}} u(t, y) l(y) dy = \int_0^t \int_{\mathbb{T}} u(s, y) l''(y) dy ds + \int_0^t \int_{\mathbb{T}} l(y) dW_{s,y}. \quad (2.1.3)$$

Looking at  $u$  as a process with values in an infinite dimensional space, the process  $u_t = u(t, \cdot)$  is also Feller diffusion taking values in  $C(\mathbb{T})$  (the space of periodic continuous functions) and  $L^2(\mathbb{T})$ . Its hitting properties were intensively studied in [MT02] by means of the Markov property, potential theory and the theory of Gaussian processes. Nevertheless some classical tools of the infinite dimensional stochastic calculus (see e.g. [DPZ14]) such as the Itô formula cannot be applied to  $u_t$ . Moreover for any fixed  $x \in \mathbb{T}$  it has been shown in [Swa07] that the process  $t \rightarrow u(t, x)$  has an a.s. infinite quadratic variation. Therefore any attempt to apply classically the powerful theory of Itô calculus seems pointless. Introduced in 2014 and explained through the famous “quartet” of articles ([Hai14] [BHZ19] [CH16] [BCCH17]), the theory of regularity structures has provided a very general framework to prove local pathwise existence and uniqueness of a wide family of stochastic PDEs driven by space-time white noise. In this chapter we will show how these new techniques allow to formulate an Itô formula for  $u$ . The formula itself will be expressed under a new form, reflecting the new perspective under which the stochastic PDEs are analysed. Indeed for any fixed smooth function  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ , we will study the quantity  $(\partial_t - \partial_{xx})\varphi(u)$ , interpreted as a space-time random distribution. This choice is heuristically motivated by the parabolic form of the equation (2.1.1) defining  $u$  and it is manageable by the regularity structures, where it is possible to manipulate random distributions. Thus we are searching for a random distribution  $g_\varphi$ , depending on higher derivatives of  $\varphi$ , such that, denoting by  $\langle \cdot, \cdot \rangle$  the duality bracket, one has a.s. the identity

$$\langle (\partial_t - \partial_{xx})\varphi(u), \psi \rangle = \langle g_\varphi, \psi \rangle, \quad (2.1.4)$$

for any test function  $\psi$ . We will refer to this formula as a *differential Itô formula*, because of the presence of a differential operator on the left hand side of (2.1.4). By uniqueness of the heat equation with the distributional source  $g_\varphi$  (see section 2.2), for every  $(t, x) \in [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$ , we can write formally

$$\varphi(u(t, x)) = \varphi(0) + \int_0^t \int_{\mathbb{T}} P_{t-s}(x - y) g_\varphi(s, y) ds dy \quad (2.1.5)$$

where for any fixed  $(t, x)$  the equality (2.1.5) hold a.s. We call a similar identity an *integral Itô formula* because of the double integral on the right hand side of (2.1.5). This resulting formula may be one possible tool to improve our comprehension of the trajectories of  $u$ , even if it is still not clear whether it will be as effective as it is for finite dimensional diffusions (see e.g. [RY04]).

In order to obtain these identities, we will follow the general philosophy of the regularity structure theory. Instead of working directly with the process  $u$ , we will consider  $\{u_\varepsilon\}_{\varepsilon>0}$  an approximating sequence of  $u$ , solving a so called “Wong-Zakai” formulation of (2.1.1) (see [HP15] for this approximation procedure on a wider class of equations)

$$\begin{cases} \partial_t u_\varepsilon = \partial_{xx} u_\varepsilon + \xi_\varepsilon, \\ u_\varepsilon(0, x) = 0 \\ u_\varepsilon(t, 0) = u_\varepsilon(t, 1) \\ \partial_x u_\varepsilon(t, 0) = \partial_x u_\varepsilon(t, 1) \end{cases} \quad (2.1.6)$$

where the random fields  $\{\xi_\varepsilon\}_{\varepsilon>0}$  are defined by extending  $\xi$  periodically on  $\mathbb{R}^2$  and convolving it with a fixed smooth, compactly supported function  $\rho: \mathbb{R}^2 \rightarrow \mathbb{R}$  such that  $\int_{\mathbb{R}^2} \rho = 1$  and  $\rho(t, x) = \rho(-t, -x)$ . That is denoting by  $*$  the convolution on  $\mathbb{R}^2$  for any  $\varepsilon > 0$  we set

$$\rho_\varepsilon = \varepsilon^{-3} \rho(\varepsilon^{-2}t, \varepsilon^{-1}x), \quad \xi_\varepsilon(t, x) = (\rho_\varepsilon * \xi)(t, x). \quad (2.1.7)$$

The inhomogeneous scaling in the mollification procedure is chosen accordingly to the parabolic nature of the equation (2.1.1). This regularisation makes  $\xi_\varepsilon$  an a.s. smooth function  $\xi_\varepsilon: [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$  and the equation (2.1.6) admits for any  $\varepsilon > 0$  an a.s. periodic strong solution (in the analytical sense)  $u_\varepsilon: [0, T] \times \mathbb{T} \rightarrow \mathbb{R}$  which is smooth in space and time. Therefore in this case  $(\partial_t - \partial_{xx})\varphi(u_\varepsilon)$  is calculated by applying the classical chain rule between  $u_\varepsilon$  and  $\varphi$ , obtaining

$$\partial_t(\varphi(u_\varepsilon)) = \varphi'(u_\varepsilon)\partial_t u_\varepsilon, \quad \partial_x(\varphi(u_\varepsilon)) = \varphi'(u_\varepsilon)\partial_x u_\varepsilon, \quad (2.1.8)$$

$$\partial_{xx}(\varphi(u_\varepsilon)) = \varphi''(u_\varepsilon)(\partial_x u_\varepsilon)^2 + \varphi'(u_\varepsilon)\partial_{xx} u_\varepsilon. \quad (2.1.9)$$

which yields:

$$(\partial_t - \partial_{xx})\varphi(u_\varepsilon) = \varphi'(u_\varepsilon)\xi_\varepsilon - \varphi''(u_\varepsilon)(\partial_x u_\varepsilon)^2. \quad (2.1.10)$$

Let us understand heuristically what happens when  $\varepsilon \rightarrow 0^+$ . Since  $\rho_\varepsilon$  is an approximation of the delta function,  $u$  is a.s. continuous and the derivative is a continuous operation between distributions, we can reasonably infer that the left hand side of (2.1.10) converges in some sense to  $(\partial_t - \partial_{xx})\varphi(u)$ . Thus the right hand side of (2.1.10) should converge too to some limit distribution. However, written under this form, it is very hard to study this right hand side because it is possible to show

$$\|\varphi'(u_\varepsilon)\xi_\varepsilon\| \xrightarrow{\mathbb{P}} +\infty, \quad \|\varphi''(u_\varepsilon)(\partial_x u_\varepsilon)^2\| \xrightarrow{\mathbb{P}} +\infty$$

with respect to some norm (see the remark 2.5.2). These two results suggest a cancellation phenomenon between two objects whose divergences compensate each other. This simple cancellation phenomenon between two diverging random quantities, which lies at the heart of the recent study of singular SPDEs, has already been noticed in the pioneering article [Zam06] (see also [GNT05], [Lan07]) and now we are able to reinterpret that result in the general context of the renormalization theory, as explained in the theory of regularity structures. By means of the notion of modelled distribution and the reconstruction theorem, we can also explain the limit as the difference of two explicit random distributions. However, these limits are only characterised by some analytical properties which cannot allow to understand immediately their probabilistic representation. Therefore the convergence is also linked with some specific *identification theorems* which describe their law. Summing up both these results we can state the main theorem of the chapter:

**Theorem 2.1.1** (Integral and differential Itô formula). *Let  $\varphi$  be a function of class  $C_b^4(\mathbb{R})$ , the space of  $C^4$  functions with all its derivatives bounded. Then for any test*

function  $\psi: \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$  with  $\text{supp}(\psi) \subset (0, T) \times \mathbb{T}$ , one has

$$\begin{aligned} \langle (\partial_t - \partial_{xx})\varphi(u), \psi \rangle = & \int_0^T \int_{\mathbb{T}} \varphi'(u_s(y)) \psi(s, y) dW_{s,y} + \frac{1}{2} \int_0^T \int_{\mathbb{T}} \psi(s, y) \varphi''(u_s(y)) C(s) dy ds \\ & - \int_{[0, T]^2 \times \mathbb{T}^2} \left[ \int_{\mathbf{s}_2 \vee \mathbf{s}_1}^T \int_{\mathbb{T}} \psi(s, y) \varphi''(u_s(y)) P'_{s-\mathbf{s}_1}(y - \mathbf{y}_1) P'_{s-\mathbf{s}_2}(y - \mathbf{y}_2) dy ds \right] dW_{\mathbf{s}, \mathbf{y}}^2. \end{aligned}$$

Moreover for any  $(t, x) \in [0, T] \times \mathbb{T}$  we have a.s.

$$\begin{aligned} \varphi(u_t(x)) = & \varphi(0) + \int_0^t \int_{\mathbb{T}} P_{t-s}(x - y) \varphi'(u_s(y)) dW_{s,y} + \frac{1}{2} \int_0^t \int_{\mathbb{T}} P_{t-s}(x - y) \varphi''(u_s(y)) C(s) dy ds \\ & - \int_{[0, t]^2 \times \mathbb{T}^2} \left[ \int_{\mathbf{s}_2 \vee \mathbf{s}_1}^t \int_{\mathbb{T}} P_{t-s}(x - y) \varphi''(u_s(y)) P'_{s-\mathbf{s}_1}(y - \mathbf{y}_1) P'_{s-\mathbf{s}_2}(y - \mathbf{y}_2) dy ds \right] dW_{\mathbf{s}, \mathbf{y}}^2, \end{aligned}$$

where in both formulae  $P'_s(y) = \partial_x P_s(x)$ , the integral  $dW_{\mathbf{s}, \mathbf{y}}^2$  is the multiple Skorohod integral of order two integrating the variables  $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2)$  and  $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$ ,  $u_s(y) = u(s, y)$  and  $C: (0, T) \rightarrow \mathbb{R}$  is the integrable function  $C(s) := \|P_s(\cdot)\|_{L^2(\mathbb{T})}^2$ .

*Remark 2.1.2.* It is natural to ask whether the same techniques could be applied to a generic convex function  $\varphi$ , in order to establish a Tanaka formula for  $u$ . In case  $\varphi$  is not a regular function, the formalism of regularity structures does not work anymore (see the section 2.4). However even if we try to generalise the Theorem 2.1.1 using only the Malliavin calculus, the presence of a multiple Skorohod integral of order 2 in both formulae would require a priori that the random variable  $\varphi''(u(s, y))$  ought to be twice differentiable (in the Malliavin sense). Hence the condition  $\varphi \in C_b^4(\mathbb{R})$  in the statement appears to be optimal. Finally any Tanaka formula would require a robust theory of local times associated to  $u$ , yet this notion is very ambiguous in the literature: On the one hand using some general results on Gaussian variables (such as [GH80]) we can prove the existence of a local time for any  $x \in \mathbb{T}$  of the process  $t \rightarrow u(t, x)$  with respect to its occupation measure on  $[0, T]$ . On the other hand an alternative notion of a local time for  $u$  has been developed by means of distributions on the Wiener space in [GNT05].

We discuss now the organization of the chapter: in the section 2.3 and 2.4 we will apply the general theorems of the regularity structures theory to build the analytical and algebraic tools to study the problem: all the constructions are mostly self contained. In some cases we will also recall some previous results obtained in [HP15] and [BHZ19], [CH16]. Then in the section 2.5 we will combine all these tools to obtain firstly two formulae involving only objects built in the previous sections (we will refer to them as *pathwise Itô formulae*) and later the identifications theorems.

We finally remark that some of the techniques presented here could potentially be used to establish an Itô formula for a non-linear perturbation of (2.1.1), the so called

generalised KPZ equation:

$$\begin{cases} \partial_t u = \partial_{xx} u + g(u)(\partial_x u)^2 + h(u)(\partial_x u) + k(u) + f(u)\xi, \\ u(0, x) = u_0(x) \\ u(t, 0) = u(t, 1) \\ \partial_x u(t, 0) = \partial_x u(t, 1) \end{cases} \quad (2.1.11)$$

where  $g, f, h, k$  are smooth functions and  $u_0 \in C(\mathbb{T})$  is a generic initial condition. (We refer the reader to [Bru15], [Hai16]). Establishing such a formula in this generalized setting shall be subject to further investigations. Other possible directions of research may also take into account the Itô formula for the solutions of other stochastic PDEs with Dirichlet boundary conditions (see [GH17]) and, using the reformulation in the regularity structures context of differential equations driven by fractional Brownian motion (see [BFG<sup>+</sup>17]), we could recover some classical results in the literature of fractional processes (see e.g. [FR02], [RV93]).

## 2.2 Hölder spaces and Malliavin calculus

We recall here some preliminary notions and notations we will use throughout the chapter. For any space-time variable  $z = (t, x) \in \mathbb{R}^2$ , in order to preserve the different role of time and space in the parabolic equation (2.1.1) we define, with an abused notation, its parabolic norm as

$$\|z\| := \sqrt{|t|} + |x|.$$

Moreover for any multi-index  $k = (k_1, k_2)$  the parabolic degree of  $k$  is given by  $|k| := 2k_1 + k_2$  and we adopt the multinomial notation for monomials  $z^k = t^{k_1} x^{k_2}$  and derivatives  $\partial^k = \partial_t^{k_1} \partial_x^{k_2}$  (the derivative  $\partial_{xx}^2 f$  will be denoted in some cases by  $\partial_{xx} f$  to shorten the notation). Accordingly to the definition of  $\rho_\varepsilon$  in (2.1.7), the parabolic rescaling of any function  $\eta: \mathbb{R}^2 \rightarrow \mathbb{R}$  of parameter  $\lambda > 0$  and centred at  $z = (t, x)$  is given by

$$\eta_z^\lambda(\bar{z}) := \lambda^{-3} \eta\left(\frac{\bar{t} - t}{\lambda^2}, \frac{\bar{x} - x}{\lambda}\right), \quad \bar{z} = (\bar{t}, \bar{x}).$$

For any non integer  $\alpha \in \mathbb{R}$ , a function  $f: \mathbb{R}^2 \mapsto \mathbb{R}$  belongs to the  $\alpha$  Hölder space  $\mathcal{C}^\alpha$  when one of these conditions is verified:

- If  $0 < \alpha < 1$ ,  $f$  is continuous and for any compact set  $\mathcal{K} \subset \mathbb{R}^2$

$$\|f\|_{\mathcal{C}^\alpha(\mathcal{K})} := \sup_{z \in \mathcal{K}} |f(z)| + \sup_{\substack{z, w \in \mathcal{K} \\ z \neq w}} \frac{|f(z) - f(w)|}{\|z - w\|^\alpha} < \infty.$$

- If  $\alpha > 1$ ,  $f$  has  $[\alpha]$  continuous derivative in space and  $[\alpha/2]$  continuous derivative in time, where  $[\cdot]$  is the integer part of a real number. Moreover for any compact set  $\mathcal{K} \subset \mathbb{R}^2$

$$\|f\|_{\mathcal{C}^\alpha(\mathcal{K})} := \sup_{z \in \mathcal{K}} \sup_{|k| \leq [\alpha]} |\partial^k f(z)| + \sup_{\substack{z, w \in \mathcal{K} \\ z \neq w}} \sup_{|k| = [\alpha]} \frac{|\partial^k f(z) - \partial^k f(w)|}{\|z - w\|^{\alpha - [\alpha]}} < \infty.$$

- If  $\alpha < 0$ , denoting  $r = -[\alpha] + 1$ ,  $f \in \mathcal{S}'(\mathbb{R}^2)$ , the set of tempered distribution on  $\mathbb{R}^2$ , and the dual of the function of order  $C^r$  on  $\mathbb{R}^2$ . Moreover for every compact set  $\mathcal{K} \subset \mathbb{R}^2$

$$\|f\|_{\mathcal{C}^\alpha(\mathcal{K})} := \sup_{z \in \mathcal{K}} \sup_{\eta \in \mathcal{B}_r} \sup_{\lambda \in (0,1]} \frac{|\langle f, \eta_z^\lambda \rangle|}{\lambda^\alpha} < \infty,$$

where  $\mathcal{B}_r$  is the set of all test functions  $\eta$  supported on  $\{z \in \mathbb{R}^2: \|z\| \leq 1\}$  such that all the directional derivatives up to order  $r$  are bounded in the sup norm.

The spaces  $\mathcal{C}^\alpha$  and the respective localised version  $\mathcal{C}^\alpha(D)$ , defined on a open set  $D \subset \mathbb{R}^2$  are naturally a family of Fréchet spaces. Moreover for any  $\alpha > 0$  and compact set  $\mathcal{K} \subset \mathbb{R}^2$ , defining  $\mathcal{C}^\alpha(\mathcal{K})$  by restriction of  $f$  on  $\mathcal{K}$ , we obtain a Banach space using the quantity  $\|f\|_{\mathcal{C}^\alpha(\mathcal{K})}$ . The elements  $f \in \mathcal{C}^\alpha(\mathbb{R} \times \mathbb{T})$  are interpreted as elements of  $\mathcal{C}^\alpha$  whose space variable lives in  $\mathbb{T}$ . Most of the classical analytical operations apply to the  $\mathcal{C}^\alpha$  spaces as follows:

- *Derivation* if  $f \in \mathcal{C}^\alpha$  and  $k$  is a multi-index then the map  $f \rightarrow \partial^k f$  is a continuous map from  $\mathcal{C}^\alpha$  to  $\mathcal{C}^\beta$  where  $\beta = \alpha - |k|$ .
- *Schauder estimates* (see [Sim97]) if  $P$  is the Heat kernel on some domain, then the space-time convolution with  $P$ ,  $f \rightarrow P * f$  is a well defined map for every  $f$  supported on positive times and it sends continuously  $\mathcal{C}^\alpha$  in  $\mathcal{C}^{\alpha+2}$  for every real non integer  $\alpha$ .
- *Product* (see [Hai14, Prop. 4.14]) for any real non integer  $\beta$  the map  $(f, g) \rightarrow f \cdot g$  defined over smooth functions extends *continuously* to a bilinear map  $B: \mathcal{C}^\alpha \times \mathcal{C}^\beta \rightarrow \mathcal{C}^{\alpha \wedge \beta}$  if and only if  $\alpha + \beta > 0$ .

The Hölder spaces and the operations defined on them provide a natural setting to formulate the deterministic PDE

$$\begin{cases} \partial_t v - \partial_{xx} v = g, \\ v(0, x) = v_0(x) \\ v(t, 0) = v(t, 1) \\ \partial_x v(t, 0) = \partial_x v(t, 1), \end{cases} \quad (2.2.1)$$

where  $g \in \mathcal{C}^\beta(\mathbb{R} \times \mathbb{T})$  and  $v_0 \in \mathcal{C}(\mathbb{T})$ . For any  $\beta > 0$ , classical results on PDE theory (see e.g. [Kry96]) imply that there exists a unique strong solution  $v \in \mathcal{C}^{\beta+2}([0, T] \times \mathbb{T})$  of (2.2.1) which is given explicitly by the so called *variation of the constant formula*

$$v(t, x) = \int_{\mathbb{T}} P_t(x - y) v_0(y) dy + (P * \mathbf{1}_{[0,t]} g)(t, x), \quad (2.2.2)$$

where for any  $t > 0$ ,  $\mathbf{1}_{[0,t]}$  is the indicator function of the interval  $[0, t] \times \mathbb{T}$ . Furthermore if we consider  $\beta \in (-2, 0)$  non-integer, the equation (2.2.1) admits again a unique solution  $v \in \mathcal{C}^{\beta+2}([0, T] \times \mathbb{T})$  satisfying (2.2.1) but only in a distributional sense. This solution can be expressed again by the formula (2.2.2) by interpreting  $\mathbf{1}_{[0,t]}$  as a continuous linear map  $\mathbf{1}_{[0,t]}: \mathcal{C}^\beta(\mathbb{R} \times \mathbb{T}) \rightarrow \mathcal{C}^\beta(\mathbb{R} \times \mathbb{T})$  such that  $(\mathbf{1}_{[0,t]} g)(\psi) = g(\psi)$  for any smooth test function  $\psi$  satisfying  $\text{supp}(\psi) \subset [0, t] \times \mathbb{T}$  and  $(\mathbf{1}_{[0,t]} g)(\psi) = 0$  if

$\text{supp}(\psi) \cap [0, t] \times \mathbb{T} = \emptyset$  (see [HP15, Lem. 6.1]). In particular it is possible to show (see [Hai14, Prop. 6.9] and [GH17, Prop. 2.15]) that for any test function  $\psi$

$$(\mathbf{1}_{[0,t]}g)(\psi) = \lim_{N \rightarrow +\infty} g(\varphi_N \psi), \quad (2.2.3)$$

where  $\varphi_N$  is a fixed sequence of smooth cutoff functions converging a.e. to  $\mathbf{1}_{(0,t) \times \mathbb{T}}$ . Thus the solution of the equation (2.2.1) is given by the same formula (2.2.2) if  $g \in \mathcal{C}^\beta(\mathbb{R} \times \mathbb{T})$ . The following procedure can be adapted straightforwardly a linear map  $\mathbf{1}_{[s,t]}: \mathcal{C}^\beta(\mathbb{R} \times \mathbb{T}) \rightarrow \mathcal{C}^\beta(\mathbb{R} \times \mathbb{T})$  for any interval  $[s, t] \subset \mathbb{R}$  eventually unbounded.

The equation (2.1.1) can be expressed in the context of the spaces  $\mathcal{C}^\alpha$ . Indeed for every  $\kappa > 0$  interpreting  $\xi$  as the derivative in space and time of the Brownian sheet  $W$ , there exists a modification of  $\xi$  belonging to  $\mathcal{C}^{-3/2-\kappa}(\mathbb{R} \times \mathbb{T})$  and defining  $\xi_\varepsilon$  as in (2.1.7) one has, as  $\varepsilon \rightarrow 0^+$ ,  $\xi_\varepsilon \rightarrow \xi$  in probability for the topology of  $\mathcal{C}^{-3/2-\kappa}(\mathbb{R} \times \mathbb{T})$  (see [Hai14, Lem. 10.2]). This theorem has some very deep consequences. Indeed choosing  $\kappa < 1/2$  and  $v_0 = 0$ , we can apply the deterministic results of (2.2.1) with every a.s. realisation of  $\xi$  and by uniqueness of the solution (2.1.1) we obtain the pathwise representation

$$u(t, x) = (P * \mathbf{1}_{[0,t]} \xi)(t, x), \quad u_\varepsilon(t, x) = (P * \mathbf{1}_{[0,t]} \xi_\varepsilon)(t, x). \quad (2.2.4)$$

Hence applying the Schauder estimates, we deduce immediately that every a.s. realisation of  $u$  belongs to  $\mathcal{C}^{1/2-\kappa}([0, T] \times \mathbb{T})$ . Therefore if we want to have a pathwise notion to an object like “ $u\xi$ ” the sum of the Hölder regularity of each factor will be  $-1 - 2\kappa$  and there is no classical way to understand this product starting from the product of two smooth approximations. The same reasoning applies also for the formal object “ $(\partial_x u)^2$ ”. We recall finally that for every distribution  $u \in \mathcal{S}'(\mathbb{R} \times \mathbb{T})$  we can define its *periodic lifting*  $\tilde{u} \in \mathcal{S}'(\mathbb{R}^2)$  defined for every test function  $\psi: \mathbb{R}^2 \rightarrow \mathbb{R}$

$$\tilde{u}(\psi) = u\left(\sum_{m \in \mathbb{Z}} \psi(\cdot, \cdot + m)\right). \quad (2.2.5)$$

Thanks to this operation we obtain that  $\tilde{u}$  coincides with the classical lift to  $\mathbb{R}^2$  of a periodic function when  $u$  is a function and we have

$$\tilde{u}(t, x) = (G * \widetilde{\mathbf{1}_{[0,t]} \xi})(t, x), \quad \tilde{u}_\varepsilon(t, x) = (G * \widetilde{\mathbf{1}_{[0,t]} \xi_\varepsilon})(t, x). \quad (2.2.6)$$

From a probabilistic point of view,  $\xi$  is an isonormal Gaussian process on  $H = L^2(\mathbb{R} \times \mathbb{T})$  defined on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . That is we can associate with any  $f \in H$  a real Gaussian random variable  $\xi(f)$  such that for any couple  $f, g \in H$  one has

$$\mathbb{E}[\xi(f)\xi(g)] = \int_{\mathbb{R}} \int_{\mathbb{T}} f(s, x)g(s, x)ds dx.$$

We denote by  $I_n: H^{\otimes n} \rightarrow L^2(\Omega)$ ,  $n \geq 1$  the multiple stochastic Wiener integral with respect to  $\xi$  (see [Nua95]).  $I_n$  is an isometry between the symmetric elements of  $H^{\otimes n}$  equipped with the norm  $\sqrt{n!} \|\cdot\|_{H^{\otimes n}}$  and the Wiener chaos of order  $n$ , the closed linear subspace of  $L^2(\Omega)$  generated by  $\{H_n(\xi(h)): \|h\|_H = 1\}$  where  $H_n$  is the  $n$ -th Hermite



polynomial. Thus we have the natural identifications  $\xi(f) = I_1(f) = \int_{\mathbb{R}} \int_{\mathbb{T}} f(s, y) dW_{s,y}$ . Let us introduce some elements of the Malliavin calculus with respect to  $\xi$  (see [Nua95] for a thorough introduction on this subject). We consider  $\mathcal{S} \subset L^2(\Omega)$  the set of random variables  $F$  of the form

$$F = h(\xi(f_1), \dots, \xi(f_n)),$$

where  $h: \mathbb{R}^n \rightarrow \mathbb{R}$  is a Schwartz test function and  $f_1, \dots, f_n \in H$ . The Malliavin derivative with respect to  $\xi$  (see [Nua95, Def. 1.2.1]) is the  $H$ -valued random variables  $\nabla F = \{\nabla_t F: t \in \mathbb{R} \times \mathbb{T}\}$  defined by

$$\nabla_t F = \sum_{i=1}^n \frac{\partial h}{\partial x_i}(\xi(f_1), \dots, \xi(f_n)) f_i(t).$$

Iterating the procedure and adopting the usual convention  $\nabla^0 = \text{id}$ , for any  $k \geq 0$  one can define the  $k$ -th Malliavin derivative  $\nabla^k F = \{\nabla_{t_1 \dots t_k}^k F: t_1, \dots, t_k \in \mathbb{R} \times \mathbb{T}\}$ , which are  $H^{\otimes k}$ -valued random variables. Moreover starting from a separable Hilbert  $V$  and considering the random variables  $G \in \mathcal{S}_V$  of the form

$$G = \sum_{i=1}^n F_i v_i \quad F_i \in \mathcal{S}, \quad v_i \in V,$$

we can also define the  $H^{\otimes k} \otimes V$  random variable  $\nabla^k G$ . In all of these cases the operator  $\nabla^k$  is closable and its domain contains for any  $1 \leq p < \infty$  the space  $\mathbb{D}^{k,p}(V)$ , the closure of  $\mathcal{S}_V$  with respect to the norm  $\|\cdot\|_{k,p,V}$  defined by

$$\|F\|_{k,p,V}^p := \mathbb{E}[\|F\|_V^p] + \sum_{l=1}^k \mathbb{E}[\|\nabla^l F\|_{H^{\otimes l} \otimes V}^p]. \quad (2.2.7)$$

(the space  $\mathbb{D}^{k,p}(\mathbb{R})$  is denoted by  $\mathbb{D}^{k,p}$ ). Trivially all variables belonging to some finite Wiener chaos are infinitely differentiable. We denote by  $\delta: \text{Dom}(\delta) \subset L^2(\Omega; H) \rightarrow L^2(\Omega)$  the adjoint operator of  $\nabla$  defined by duality as

$$\mathbb{E}[\delta(u)F] = \mathbb{E}[\langle u, \nabla F \rangle_H]$$

for any  $u \in \text{Dom}(\delta)$ ,  $F \in \mathbb{D}^{1,2}$ . The operator  $\delta$  is known in the literature as the Skorohod integral and for any  $u \in \text{Dom}(\delta)$  we will write again  $\delta(u)$  with the symbol  $\int_{\mathbb{R}} \int_{\mathbb{T}} u(s, y) dW_{s,y}$  because  $\delta$  is a proper extension of the stochastic integration over a class of non adapted integrands. Using the same procedure we define  $\delta^k: \text{Dom}(\delta^k) \subset L^2(\Omega; H^{\otimes k}) \rightarrow L^2(\Omega)$ , the adjoint of  $\nabla^k$ . Similarly to before we call the operator  $\delta^k$  the multiple Skorohod integral of order  $k$  and we denote it by the notation

$$\int_{(\mathbb{R} \times \mathbb{T})^k} u((t_1, x_1), \dots, (t_k, x_k)) dW_{t,x}^k.$$

Let us recall the main properties of  $\delta^k$ .

- *Extension of the Wiener integral* For any  $h \in H^{\otimes k}$ , we have  $\delta^k(h) = I_k(h)$ .

- *Product Formula* (see [NN10, Lem. 2.1]) Let  $u \in \text{Dom}(\delta^k)$  be a symmetric function in the variables  $t_1, \dots, t_k$  and  $F \in \mathbb{D}^{k,2}$ . If for any couple of positive integers  $j, r$  such that  $0 \leq j + r \leq k$  one has  $\langle \nabla^r F, \delta^j u \rangle_{H^{\otimes r}} \in L^2(\Omega; H^{\otimes(k-r-j)})$  then  $\langle \nabla^r F, u \rangle_{H^{\otimes r}} \in \text{Dom}(\delta^{q-r})$  and we have

$$F\delta^k(u) = \sum_{r=0}^k \binom{k}{r} \delta^{k-r}(\langle \nabla^r F, u \rangle_{H^{\otimes r}}). \quad (2.2.8)$$

- *Continuity property* (see [NN10, Pag. 8]) We have the inclusion  $\mathbb{D}^{k,2}(H^{\otimes k}) \subset \text{Dom}(\delta^k)$  and the map  $\delta^2: \mathbb{D}^{k,2}(H^{\otimes k}) \rightarrow L^2(\Omega)$  is continuous. In other words there exists a constant  $C > 0$  such that for any  $u \in \mathbb{D}^{k,2}(H^{\otimes k})$  one has

$$\|\delta^k(u)\|_{L^2(\Omega)} \leq C\|u\|_{\mathbb{D}^{k,2}(H^{\otimes k})}. \quad (2.2.9)$$

Extending periodically  $\xi$  and the Brownian sheet  $W$  to  $\mathbb{R}^2$ , we can transfer the Walsh integral as well as the Skorohod integral to stochastic processes  $H: \Omega \times \mathbb{R}^2 \rightarrow \mathbb{R}$  through the definition:

$$\int_{\mathbb{R}^2} H(s, y) d\widetilde{W}_{s,y} := \int_{\mathbb{R} \times \mathbb{T}} \sum_{m \in \mathbb{Z}} H(s, y + m) dW_{s,y}, \quad (2.2.10)$$

as long as the right hand side above is well defined. Similar definitions hold for the multiple Skorohod integral of order  $k$ , *mutatis mutandis*. Using this notation we express

$$\xi_\varepsilon(t, x) = \int_{\mathbb{R}^2} \rho_\varepsilon(t - s, x - y) d\widetilde{W}_{s,y}, \quad u(t, x) = \int_0^t \int_{\mathbb{R}} G(t - s, x - y) d\widetilde{W}_{s,y}.$$

## 2.3 Regularity structures

In this part we will recall some general concepts of the theory of regularity structures to show the existence of an explicit *regularity structure* and a *model*. These objects will permit to define some analytical operations on  $u$ . For a quick introduction to the whole theory we refer the reader to [Hai15]

### 2.3.1 Algebraic construction

The starting point of the theory is the notion of a *regularity structure*  $(A, T, G)$ , a triple of the following elements:

- A discrete lower bounded real subset  $A$  containing 0.
- A graded vector space  $T = \bigoplus_{\alpha \in A} T_\alpha$  such that each space  $T_\alpha$  is a Banach space with norm  $\|\cdot\|_\alpha$  and  $T_0$  is generated by a single element  $\mathbf{1}$ .
- A group  $G$  of linear operators on  $T$  such that for each  $\alpha \in A$ ,  $a$  in  $T_\alpha$  and  $\Gamma$  in  $G$ , one has  $\Gamma\mathbf{1} = \mathbf{1}$  and

$$\Gamma a - a \in \bigoplus_{\beta < \alpha} T_\beta. \quad (2.3.1)$$

The element of a regularity structure are interpreted as a set of generalised polynomial whose elements are capable to perform a ‘‘Taylor expansion’’ of distributions. Recalling the equations (2.1.10) and (2.1.6), our aim is then to build a regularity structure  $\mathcal{T}$  whose elements are capable to describe for any  $\varepsilon > 0$  the systems of equations

$$\begin{cases} \partial_t u_\varepsilon = \partial_{xx} u_\varepsilon + \xi_\varepsilon \\ \partial_t v_\varepsilon = \partial_{xx} v_\varepsilon + \varphi'(u_\varepsilon)\xi_\varepsilon - \varphi''(u_\varepsilon)(\partial_x u_\varepsilon)^2. \end{cases} \quad (2.3.2)$$

Let us give a first description of  $\mathcal{T}$  in terms of abstract symbols. We start by considering the real polynomials on two indeterminates. For any multi-index  $k \in \mathbb{N}^2$ ,  $k = (k_1, k_2)$  we will write  $\mathbf{X}^k$  as a shorthand for the monomial  $X_1^{k_1} X_2^{k_2}$  while the unit will be denoted by  $\mathbf{1}$ . In this way we will be able to describe smooth functions. At the same time, we introduce an additional abstract symbol  $\Xi$  to represent the space-time white noise  $\xi$  and for any symbol  $\sigma$  and  $k \in \mathbb{N}^2$  we introduce a family of symbols  $\mathcal{I}_k(\sigma)$  ( $\mathcal{I}_{(0,0)}(\sigma)$  is denoted by  $\mathcal{I}(\sigma)$ ) to represent formally the convolution of the  $k$ -th derivative of the heat kernel with the function associated to the symbol  $\sigma$ . Since  $\mathcal{I}_k(\mathbf{X}^m)$  should be identified with a smooth function, we simply put it to 0 to avoid repetitions. Finally for any two symbols  $\tau_1, \tau_2$  we consider also the symbol  $\tau_1 \tau_2$ , obtained by juxtaposition of  $\tau_1$  and  $\tau_2$  and identified up to some classical identifications, namely the juxtaposition with  $\mathbf{1}$  does not change the symbol,  $\mathbf{X}^l \mathbf{X}^k = \mathbf{X}^{l+m}$  and the iterated juxtaposition of the same symbol is denoted by an integer power. Adding all these formal rules, we denote by  $F$  the set of all possible formal expressions satisfying

- $\{\mathbf{X}^k\}_{k \in \mathbb{N}^2} \cup \{\Xi\} \subset F$ .
- For any  $\tau_1, \tau_2 \in F$ ,  $\tau_1 \tau_2 \in F$ .
- For any  $\sigma \in F$  and  $m \in \mathbb{N}^2$ ,  $\mathcal{I}_m(\sigma) \in F$ .

We write  $\mathcal{F}$  for the free vector space generated by  $F$ . Similarly to polynomials, we define a *homogeneity* map  $|\cdot|: F \rightarrow \mathbb{R}$  whose values have the properties of the degree of polynomials but in the context of the Hölder spaces. In particular we set recursively

- $|\mathbf{X}^k| := 2k_1 + k_2$  the parabolic degree (we imagine  $X_1$  as a time variable for the parabolic degree);
- $|\Xi| := -3/2 - \kappa$  for some fixed parameter  $\kappa > 0$ ;
- $|\mathcal{I}_k(\tau)| := |\tau| + 2 - 2k_1 - k_2$ ,  $|\tau\tau'| := |\tau| + |\tau'|$  for any  $\tau, \tau' \in F$ .

Starting from the linear space  $\mathcal{F}$ , we introduce a subset of  $F$  where we choose all reasonable products that we will need in our calculations. We write  $\mathcal{I}_1(\Xi)$  as shorthand of  $\mathcal{I}_{(0,1)}(\Xi)$ .

**Definition 2.3.1.** We define the sets of symbols  $T, U, U' \subset F$  as the smallest triple of sets satisfying:

- $\{\Xi\} \subset T$ ,  $\{\mathcal{I}(\Xi)\} \cup \{\mathbf{X}^k\}_{k \in \mathbb{N}^2} \subset U$ ,  $\{\mathcal{I}_1(\Xi)\} \cup \{\mathbf{X}^k\}_{k \in \mathbb{N}^2} \subset U'$ ;
- for every  $k \geq 0$  and any finite family of elements  $\tau_1, \dots, \tau_k \in U$  and any couple of elements  $\sigma_1, \sigma_2 \in U'$  then  $\{\tau, \tau\Xi, \tau\sigma_1, \tau\sigma_1\sigma_2\} \subset T$  and  $\tau \in U$ , where  $\tau = \tau_1 \dots \tau_k$ .

We denote also by  $\mathcal{T}$  and  $\mathcal{U}$  respectively the free vector space upon  $T$  and  $U$ .

The definition of  $T$  has an equivalent description in terms of symbols. Defining  $V = \{\mathcal{I}(\Xi)^m \mathbf{X}^l : m \in \mathbb{N}, l \in \mathbb{N}^2\}$  and for any  $\sigma \in \{\Xi, \mathcal{I}_1(\Xi), \mathcal{I}_1(\Xi)^2\}$   $V_\sigma := \sigma V$  the set of all symbols of the form  $\sigma$  times an element of  $V$ , it is straightforward to show the identities

$$U = V, \quad T = V_\Xi \sqcup V_{\mathcal{I}_1(\Xi)^2} \sqcup V_{\mathcal{I}_1(\Xi)} \sqcup V. \quad (2.3.3)$$

Therefore, denoting by  $\mathcal{V}_\sigma$  the free vector space generated upon  $V_\sigma$ , we have the decomposition  $\mathcal{T} = \mathcal{V}_\Xi \oplus \mathcal{V}_{\mathcal{I}_1(\Xi)^2} \oplus \mathcal{V}_{\mathcal{I}_1(\Xi)} \oplus \mathcal{U}$ . Let us give the construction of the structure group associated to  $\mathcal{T}$ . For any  $h \in \mathbb{R}^3$ ,  $h = (h_1, h_2, h_3)$  we define the function  $\Gamma_h : \mathcal{T} \rightarrow \mathcal{T}$  as the unique linear map such that

$$\Gamma_h(\sigma \mathcal{I}(\Xi)^m \mathbf{X}^l) := \sigma [(X_1 + h_1 \mathbf{1})^{l_1} (X_2 + h_2 \mathbf{1})^{l_2} (\mathcal{I}(\Xi) + h_3 \mathbf{1})^m], \quad (2.3.4)$$

for any  $\sigma \in \{\Xi, \mathcal{I}_1(\Xi), \mathcal{I}_1(\Xi)^2, \mathbf{1}\}$ ,  $m \in \mathbb{N}$ ,  $l \in \mathbb{N}^2$ . Using this explicit definition it is straightforward to show

$$\Gamma_h \Gamma_k = \Gamma_{h+k} \quad (2.3.5)$$

for any  $h, k \in \mathbb{R}^3$  and the map  $h \rightarrow \Gamma_h$  is injective. We will denote by  $\mathcal{G}$  the group  $\{\Gamma_h : h \in \mathbb{R}^3\}$ .

**Proposition 2.3.2.** *For any  $\kappa < 1/2$ , the triple  $(\mathcal{A}, \mathcal{T}, \mathcal{G})$  where  $\mathcal{A} = \{|\tau| : \tau \in T\}$  is a regularity structure.*

*Proof.* To prove that  $\mathcal{A}$  is a discrete lower bounded set, we show that for any  $\beta \in \mathbb{R}$  the set  $I := \{\tau \in T : |\tau| \leq \beta\}$  is finite. For any  $\tau \in I$  by means of the identity (2.3.3) there exist two indices  $m \in \mathbb{N}$ ,  $n \in \mathbb{N}^2$  and  $\sigma \in \{\Xi, \mathcal{I}_1(\Xi), \mathcal{I}_1(\Xi)^2\}$  such that  $\tau = \sigma \mathcal{I}(\Xi)^m \mathbf{X}^n$ . From  $|\tau| \leq \beta$  we deduce

$$n_1 + 2n_2 + (1/2 - \kappa)m \leq \beta - |\sigma|. \quad (2.3.6)$$

Imposing  $\kappa < 1/2$ , the left hand side of the inequality is strictly bigger or equal than 0 and the set  $I$  is bounded. This finiteness result implies also the identity  $\mathcal{T} = \bigoplus_{\gamma \in \mathcal{A}} \mathcal{T}_\gamma$  where  $\mathcal{T}_\gamma = \{\tau \in T : |\tau| = \gamma\}$ . Moreover there is no need to specify a norm on  $\mathcal{T}_\gamma$ , since it is finite dimensional. Finally the property (2.3.1) comes directly from Newton's binomial formula and the positive homogeneity of the symbol  $\mathcal{I}(\Xi)$ .  $\square$

*Remark 2.3.3.* The Definition 2.3.1 is a simplification of the vector space introduced in [Hai16, Pag. 7] with fewer symbols. The triple  $(\mathcal{A}, \mathcal{T}, \mathcal{G})$  is also intimately linked with  $(\mathcal{A}^{HP}, \mathcal{T}^{HP}, \mathcal{G}^{HP})$ , the regularity structure defined in [HP15, Pag. 13-14]. More precisely we consider  $U^{HP}$  the smallest set of symbols of  $F$  such that  $\{\mathbf{X}^k\}_{k \in \mathbb{N}^2} \subset U$  and satisfying the properties

$$\tau \in U^{HP} \Rightarrow \mathcal{I}(\tau), \mathcal{I}(\Xi \tau) \in U^{HP}; \quad \tau, \tau' \in U^{HP} \Rightarrow \tau \tau' \in U^{HP}$$

Introducing the set  $T_\Xi^{HP} = \{\Xi v \in F : v \in U^{HP}\}$  and  $\mathcal{T}_\Xi^{HP}, \mathcal{U}^{HP}$  the corresponding free vector spaces defined on these sets, the space  $\mathcal{T}^{HP}$  is defined by  $\mathcal{T}^{HP} = \mathcal{T}_\Xi^{HP} \oplus \mathcal{U}^{HP}$ .

Looking at  $\mathcal{T} \cap \mathcal{T}^{HP} = \mathcal{V}_\Xi \oplus \mathcal{U}$ , it is also possible to show that the action of the group  $\mathcal{G}^{HP}$  coincides with that of  $\mathcal{G}$  on these subspaces and from the explicit definition of  $\Gamma$  one has  $\Gamma(\mathcal{V}_\Xi) \subset \mathcal{V}_\Xi$  and  $\Gamma(\mathcal{U}) \subset \mathcal{U}$  for any  $\Gamma \in \mathcal{G}$  (see the Remark 2.3.3). Hence the subspaces  $\mathcal{V}_\Xi$  and  $\mathcal{U}$  are respectively a sector of regularity  $-3/2 - \kappa$  and a function-like sector of both  $\mathcal{T}^{HP}$  and  $\mathcal{T}$  (see [Hai14, Def. 2.5]). Due to these identifications, we can transfer some results of [HP15] to our context.

*Remark 2.3.4.* As a matter of fact we can restrict our considerations once and for all to a subspace of  $\mathcal{T}$  generated by all symbols with homogeneity less than some parameter  $\zeta > 0$ . By convention we denote by  $|\cdot|_\beta$  the euclidean norm on  $\mathcal{T}_\beta$  (The euclidean norm is coherent with [HP15] but there is no “canonical” choice since  $\mathcal{T}_\beta$  is finite dimensional). For any  $\beta \in A$ , we will denote by  $\mathcal{Q}_\beta$  and  $\mathcal{Q}_{<\beta}$  the projection operator respectively on  $\mathcal{T}_\beta$  and  $\bigotimes_{\alpha < \beta} \mathcal{T}_\alpha$ .

*Remark 2.3.5.* A consequence of the definition (2.3.4) implies that for any  $h \in \mathbb{R}^3$  one has  $\Gamma_h \tau \tau' = \Gamma_h \tau \Gamma_h \tau'$  for every symbol  $\tau, \tau' \in T$  such that also their product  $\tau \tau'$  belongs to  $T$ . We remark that the explicit expression of  $\Gamma_h$  in (2.3.4) can be easily rewritten as

$$\Gamma_h = (id \otimes \tilde{h})\Delta, \quad (2.3.7)$$

where  $\tilde{h}$  is a real character over  $\mathcal{U}$  (we can easily identify  $\mathcal{U}$  with the space of polynomials in three coordinates) and  $\Delta: \mathcal{T} \rightarrow \mathcal{T} \otimes \mathcal{U}$  is the only unique linear map such that for every  $i = 1, 2$ ,  $\sigma \in \{\Xi, \mathcal{I}_1(\Xi), \mathcal{I}_1(\Xi)^2\}$  and all  $\tau, \tau' \in T$  such that  $\tau \tau' \in T$  one has

$$\begin{aligned} \Delta X_i &= X_i \otimes \mathbf{1} + \mathbf{1} \otimes X_i, & \Delta \mathbf{1} &= \mathbf{1} \otimes \mathbf{1}, & \Delta \sigma &= \sigma \otimes \mathbf{1} \\ \Delta \mathcal{I}(\Xi) &= \mathcal{I}(\Xi) \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{I}(\Xi), & \Delta \tau \tau' &= \Delta \tau \Delta \tau'. \end{aligned} \quad (2.3.8)$$

Thus the group  $\mathcal{G}$  coincides with the classical way to define the structure group  $G$  of a regularity structure. Moreover comparing the relations (2.3.8) with the explicit definitions given in [Hai14, Sec. 8.1] and [HP15, Pag. 14] we obtain that the group  $\mathcal{G}$  is simply the restriction of these general constructions when they are applied on the set of symbols defining  $\mathcal{T}$ .

In order to apply some general results obtained in [BHZ19] and [CH16], we show how to express the regularity structure  $\mathcal{T}$  by means of the formalism of *trees*. Let us recall some basic notations. We start by considering labelled, rooted trees  $\tau$  (LR tree), that is  $\tau$  is a combinatorial tree (finite connected simple graph with a non-empty set of nodes  $N_\tau$  and a set of edges  $E_\tau$  without cycles and not planar) where we fixed a specific node  $\rho_\tau \in N_\tau$  called the root of  $\tau$ . The trees we consider are also *labelled* i.e. there exists a finite set of labels  $\mathcal{L}$  and a function  $\mathbf{t}: E_\tau \rightarrow \mathcal{L}$ . These trees are the building blocks of a more general family of trees. We define a *decorated tree* as a triple  $\tau_\mathbf{c}^\mathbf{n} = (\tau, \mathbf{n}, \mathbf{c})$  where  $\tau$  is a LR rooted tree and  $\mathbf{n}: N_\tau \rightarrow \mathbb{N}^2$ ,  $\mathbf{c}: E_\tau \rightarrow \mathbb{N}^2$  are two fixed functions. The set of decorated tree is denoted by  $\mathfrak{T}$ . Similarly to what we did for the set of symbols  $F$  we fix a scaling function  $\mathfrak{s}: \mathcal{L} \sqcup \mathbb{N}^2 \rightarrow \mathbb{R}$  and we define a homogeneity map  $|\cdot|_\mathfrak{s}: \mathfrak{T} \rightarrow \mathbb{R}$  as

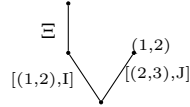
$$|\tau_\mathbf{c}^\mathbf{n}|_\mathfrak{s} := \sum_{e \in E_\tau} \mathfrak{s}(\mathbf{t}(e)) - \mathfrak{s}(\mathbf{c}(e)) + \sum_{x \in N_\tau} \mathfrak{s}(\mathbf{n}(x)), \quad (2.3.9)$$

for any  $\tau_{\mathfrak{e}}^{\mathfrak{n}} \in \mathfrak{T}$ . Moreover for any two elements  $\tau_{\mathfrak{e}}^{\mathfrak{n}}, \sigma_{\mathfrak{e}}^{\mathfrak{n}'} \in \mathfrak{T}$  we can also define a product tree  $\tau_{\mathfrak{e}}^{\mathfrak{n}} \sigma_{\mathfrak{e}}^{\mathfrak{n}'}$  by simply considering  $\tau \sigma$ , the tree obtained by joining the roots of  $\tau_{\mathfrak{e}}^{\mathfrak{n}}$  and  $\sigma_{\mathfrak{e}}^{\mathfrak{n}'}$  and imposing  $\mathfrak{n}(\rho_{\iota(\tau\sigma)}) = \mathfrak{n}(\rho_{\iota(\tau)}) + \mathfrak{n}'(\rho_{\iota(\sigma)})$ . Finally for any  $m \in \mathbb{N}^2$ ,  $l \in \mathcal{L}$  we define an application  $\mathcal{E}_m^l: \mathfrak{T} \rightarrow \mathfrak{T}$  (called grafting operation) as follows: for any  $\sigma \in \mathfrak{T}$ ,  $\mathcal{E}_m^l \sigma \in \mathfrak{T}$  is the tree with zero decoration on the root obtained by adding one more edge decorated by  $(l, m)$  to the root of  $\sigma$ . The set  $\mathfrak{T}$  can be constructed recursively starting from the root trees  $\{\bullet_k\}_{k \in \mathbb{N}^2}$  and applying iteratively the grafting operations and the multiplication.

Looking at the equation (2.3.2), we choose in this case a set of labels with three elements  $\mathcal{L} = \{\Xi, I, J\}$  ( $\Xi$  should represent the noise and  $I, J$  are respectively the heat kernel in the first and the second equation). Moreover, the function  $\mathfrak{s}$  is defined by

$$\mathfrak{s}(k_1, k_2) = 2k_1 + k_2, \quad \mathfrak{s}(\Xi) = -\frac{3}{2} - \kappa, \quad \mathfrak{s}(I) = \mathfrak{s}(J) = 2, \quad (2.3.10)$$

where  $\kappa > 0$  is a fixed parameter (this choice of  $\mathfrak{s}$  is done to imitate the parabolic degree and the Schauder estimates). we can easily draw a decorated tree  $\tau_{\mathfrak{e}}^{\mathfrak{n}} \in \mathfrak{T}$  by simply putting its root at the bottom and decorating the nodes and the edges with the non zero values of  $\mathfrak{n}, \mathfrak{e}$ . For example when we write the tree



we have three nodes with a zero  $\mathfrak{n}$  decoration and the edge labelled with  $\Xi$  has zero  $\mathfrak{e}$  decoration.

After applying these choices, similarly to what we have done to define  $T$ , we can build a regularity structure from  $\mathfrak{T}$  by choosing a suitable subset of trees  $\mathfrak{T}' \subset \mathfrak{T}$ . This operation is then formalised in the context of decorated trees by the notion of a “rule” which describes what type of edges are allowed next to every label (see [BHZ19, Def. 5.7]). More precisely, denoting by  $E$  the set of all finite multisets of  $\mathcal{L} \times \mathbb{N}^2$ , a rule is a function  $R: \mathcal{L} \rightarrow P(E) \setminus \{\emptyset\}$  where  $P(E)$  is the power set of  $E$ . Let us explain in our case what rule we choose in this context.

**Definition 2.3.6.** Writing  $I$  and  $I_1$  as a shorthand for  $(I, (0, 0)), (I, (0, 1)) \in \mathcal{L} \times \mathbb{N}^2$  we define

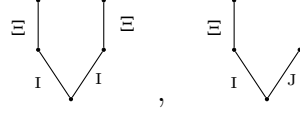
$$\begin{aligned} R(\Xi) &= \{()\}, & R(I) &= \{(), \Xi\}, \\ R(J) &= \{(), \underbrace{(I \cdots I)}_{k \text{ times}}, \underbrace{(I \cdots I, I_1)}_{k \text{ times}}, \underbrace{(I \cdots I, I_1, I_1)}_{k \text{ times}}, \underbrace{(I \cdots I, \Xi)}_{k \text{ times}}, k \in \mathbb{N}\}, \end{aligned} \quad (2.3.11)$$

where the brackets  $\{\}$  describe a subset of  $E$  and the brackets  $()$  a multiset of  $\mathcal{L} \times \mathbb{N}^2$  (the symbol  $()$  denotes the empty multiset).

Once we established a rule, we can consider the set of all decorated trees which *strongly conforms to the rule*  $R$  (see [BHZ19, Def. 5.8]), denoted by  $\mathfrak{T}(R)$ , that is  $\tau_{\mathfrak{e}}^{\mathfrak{n}} \in \mathfrak{T}(R)$  if the following properties are satisfied

- Looking at the edges attached at the root  $\rho_\tau$ , they can be expressed as  $R(l)$  for some  $l \in \mathcal{L}$ ;
- for any node  $x \in N_\tau \setminus \{\rho_\tau\}$ , all the edges attached at  $x$  can be written as  $R(\mathfrak{t}(e))$ , where  $e$  is the unique edge linking  $x$  to its parent.

For example we use the shorthand notation  $J = [(0, 0), J]$  and we consider the two trees



the tree on the left hand side strongly conforms to the rule  $R$  but the tree on the right one does not because the multiset  $(I, J)$  is not in the image of  $R$ . From the Definition 2.3.6 it is straightforward to see that all possible decorations of the trees in  $\mathfrak{T}(R)$  are of three types. We will abbreviate them with the shorthand notations

$$\Xi \begin{array}{c} | \\ \vdots \\ | \\ = \\ | \end{array}, \quad I \begin{array}{c} | \\ = \\ | \end{array}, \quad I_1 \begin{array}{c} | \\ = \\ | \\ = \\ | \end{array}.$$

Thanks to this operation we can extract a specific subset of trees, but if we want to build a regularity structure it is necessary to check that the rule  $R$  has two last fundamental properties. Firstly  $R$  must be *subcritical* (see [BHZ19, Def. 5.14]), that is there must exist a function  $\text{reg}: \mathcal{L} \rightarrow \mathbb{R}$  such that when it is extended to  $E$  as follows

$$\text{reg}(l, k) := \text{reg}(l) - \mathfrak{s}(k), \quad \text{reg}(N) := \sum_{(l,k) \in M} \text{reg}(l, k),$$

for any  $(l, k) \in \mathcal{L} \times \mathbb{N}^2$  and  $M \in E$ , then we have

$$\text{reg}(l) < \mathfrak{s}(l) + \inf_{M \in R(l)} \text{reg}(M),$$

For any  $l \in \mathcal{L}$ . Moreover  $R$  must also be *normal* (see [BHZ19, Def. 5.8]), that is  $R(l) = \{()\}$  for every  $l \in \mathcal{L}$  such that  $\mathfrak{s}(l) < 0$  and for any couple of multisets  $M, N \in E$  of  $E$  such that  $N \in R(l)$  for some  $l \in \mathcal{L}$  and  $M \subset N$ , then  $M \in R(l)$ . Both properties are relatively easy to check in this specific case. Indeed the rule  $R$  is normal by construction and we can verify the subcriticality using the function  $\text{reg}: \mathcal{L} \rightarrow \mathbb{R}$  defined by

$$\text{reg}(\Xi) = \frac{3}{2} - 2\kappa, \quad \text{reg}(I) = \text{reg}(J) = \frac{1}{2} - 3\kappa,$$

as long as  $\kappa$  is sufficiently small. These two properties allow us to apply the results [BHZ19, Prop. 5.21], [BHZ19, Prop. 5.39] and the definition [BHZ19, Def. 6.22] to prove the following result

**Proposition 2.3.7.** *There exists a rule  $R'$  such that  $R(l) \subset R'(l)$  for every  $l \in \mathcal{L}$  and a group  $\mathcal{G}'$  such that the triple  $(\mathcal{A}', \mathcal{T}', \mathcal{G}')$  is a regularity structure, where  $\mathcal{T}'$  is the free vector space generated from  $\mathfrak{T}(R')$  and  $\mathcal{A} = \{|\tau|_{\mathfrak{s}}: \tau \in \mathfrak{T}(R')\}$ .*

Even if we could give an explicit description of the group  $\mathcal{G}'$  and  $\mathcal{T}'$  given in the Proposition 2.3.7, for our purposes it is sufficient to establish a relation between the regularity structure  $\mathcal{T}$  and  $\mathcal{T}'$ . From the explicit definition of  $F$  and  $\mathfrak{T}$  it is possible to define recursively an injective map  $\iota: F \rightarrow \mathfrak{T}$  as follows:

- for any  $m \in \mathbb{N}^2$  we set

$$\iota(\Xi) := \begin{array}{c} \vdots \\ \downarrow \\ \bullet \end{array}, \quad \iota(\mathbf{X}^m) := \bullet_m$$

- For any symbol  $\sigma$  such that  $\iota(\sigma)$  is defined, then  $\iota(\mathcal{I}_k(\sigma)) := \mathcal{E}_k^I(\sigma)$ .
- For any couple of symbols  $\sigma, \sigma'$  such that  $\iota(\sigma)$  and  $\iota(\sigma')$  are well defined we set  $\iota(\sigma\sigma') = \iota(\sigma)\iota(\sigma')$ .

We present two examples of the action of  $\iota$ :

$$\iota(\Xi^2 \mathcal{I}(\Xi)) = \begin{array}{c} \vdots \quad \vdots \quad \vdots \\ \diagdown \quad \downarrow \quad \diagup \\ \bullet \end{array}, \quad \iota(\mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi \mathbf{X}^{(3,4)})) = \begin{array}{c} \vdots \quad \vdots \quad \vdots \\ \diagdown \quad \downarrow \quad \diagup \\ \bullet \end{array}^{(3,4)}.$$

Restricting the map  $\iota$  on  $T$  and extending it by linearity we have the following inclusion

**Proposition 2.3.8.** *The regularity structure  $(\mathcal{A}, \mathcal{T}, \mathcal{G})$  is contained in  $(\mathcal{A}', \mathcal{T}', \mathcal{G}')$  in the sense of the inclusion explained in [Hai14, Sec. 2.1].*

*Proof.* The theorem is a strict consequence of the choices done to define  $\mathcal{T}'$ . Firstly by definition of  $R$ , every decorated tree  $\iota(\tau)$  for some  $\tau \in T$  strongly conforms to the rule  $R$ , therefore it will conform to the rule  $R'$ . Moreover by construction of  $\mathfrak{s}$  in (2.3.10) we have  $|\iota(\sigma)|_{\mathfrak{s}} = |\sigma|$  for any  $\sigma \in F$ . Thus  $\mathcal{A} \subset \mathcal{A}'$ . Finally, when we consider the groups  $\mathcal{G}, \mathcal{G}'$ , it has been showed in [BHZ19, Pag. 89] that the group  $\mathcal{G}'$  acts on  $\iota(\mathcal{T})$  in the same way as the operator  $\Delta$  explained in the Remark 2.3.5. Therefore we obtain the inclusion of the regularity structures.  $\square$

*Remark 2.3.9.* The function  $\iota$  is clearly an injective map function from  $F$  to  $\mathfrak{T}$  but there are many trees of  $\mathfrak{T}$  which do not belong to  $\iota(F)$ . In particular all the trees whose edges are only labelled with  $I$  are not contained, because we identified all the symbols  $\mathcal{I}_k(\mathbf{X}^m)$  to zero. Moreover none of trees labelled with  $J$  belong to  $\iota(F)$ . The presence of two different labels  $I, J$  to denote the heat kernel is actually done to isolate all the tree we really need for our calculation and if we identified  $I = J$  in the Definition 2.3.6 we would obtain the rule for generalised KPZ equation as explained in [BHZ19, Sec. 5.4]. In what follows we will identify both symbols and decorated trees, without writing explicitly the map  $\iota$ .

### 2.3.2 Models on a regularity structure

The algebraic structure comes also with a *model* associated to it. In order to recall this notion and to simplify the whole exposition, we fix a parameter  $\zeta \geq 2$  and with an abuse of notation we will identify all along the chapter  $\mathcal{T}$  (respectively its canonical basis  $T$ ) with the finite dimensional vector space  $\mathcal{Q}_{<\zeta} \mathcal{T}$  (resp. the finite set  $\{\tau \in T: |\tau| < \zeta\}$ ). The same applies also for the sets  $V_{\Xi}, V_{\mathcal{I}_1(\Xi)^2}, V_{\mathcal{I}_1(\Xi)}$ .



**Definition 2.3.10.** A model on  $(\mathcal{A}, \mathcal{T}, \mathcal{G})$  consists of a pair  $(\Pi, \Gamma)$  given by:

- A map  $\Gamma: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathcal{G}$  such that  $\Gamma_{zz} = id$  and  $\Gamma_{zv}\Gamma_{vw} = \Gamma_{zw}$  for any  $z, v, w \in \mathbb{R}^2$ .
- A collection  $\Pi = \{\Pi_z\}_{z \in \mathbb{R}^2}$  of linear maps  $\Pi_z: \mathcal{T} \mapsto \mathcal{S}'(\mathbb{R}^2)$  such that  $\Pi_z = \Pi_v\Gamma_{vz}$  for any  $z, v \in \mathbb{R}^2$ .

Furthermore, for every compact set  $\mathcal{K} \subset \mathbb{R}^2$ , one has

$$\|\Pi\|_{\mathcal{K}} := \sup \left\{ \frac{|(\Pi_z\tau)(\eta_z^\lambda)|}{\lambda^{|\tau|}} : z \in \mathcal{K}, \lambda \in (0, 1], \tau \in T, \eta \in \mathcal{B}_2 \right\} < \infty, \quad (2.3.12)$$

$$\|\Gamma\|_{\mathcal{K}} := \sup \left\{ \frac{|\Gamma_{zw}(\tau)|_\beta}{\|z - w\|^{|\tau| - \beta}} : z \neq w \in \mathcal{K}, \|z - w\| \leq 1, \tau \in T, \beta < |\tau| \right\} < \infty, \quad (2.3.13)$$

where the set of test functions  $\mathcal{B}_2$  was already introduced in the Section 2.2.

This notion plays a fundamental role in the whole theory, because it associates to any  $\tau \in T$  an explicit distribution  $\Pi_z\tau$  belonging in some way to  $\mathcal{C}^{|\tau|}$ . In order to compare two different models defined on the same structure, we endow  $\mathcal{M}$ , the set of all models on  $(\mathcal{A}, \mathcal{T}, \mathcal{G})$ , with the topology associated to the corresponding system of semi-distances induced by the conditions (2.3.12) and (2.3.13):

$$\|(\Pi, \Gamma), (\bar{\Pi}, \bar{\Gamma})\|_{\mathcal{M}(\mathcal{K})} := \|\Pi - \bar{\Pi}\|_{\mathcal{K}} + \|\Gamma - \bar{\Gamma}\|_{\mathcal{K}}. \quad (2.3.14)$$

Since we want to study the processes on a finite time horizon, it is sufficient to verify the conditions (2.3.12) (2.3.13) on a fixed compact set  $\mathcal{K}$  containing  $[0, T] \times [0, 1]$  and we will avoid any reference of it in the notation. In this way  $(\mathcal{M}, \|\cdot\|_{\mathcal{M}})$  becomes a complete metric space ( $\mathcal{M}$  is not a Banach space because the sum of models is not necessarily a model!). In particular if a sequence  $(\Pi^n, \Gamma^n)$  converge to  $(\Pi, \Gamma)$ , then  $\Pi_z^n\tau$  converges to  $\Pi_z\tau$  in the sense of tempered distributions for any  $z, \tau$ . To define correctly a model over a symbol of the form  $\mathcal{I}(\sigma)$ , we need a technical lemmas related to a suitable decompositions of  $G$ , the heat kernel on  $\mathbb{R}$  interpreted as a function  $G: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ .

**Lemma 2.3.11** (First decomposition). (see [Hai14, Lemma 5.5]) *There exists a couple of functions  $K: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ ,  $R: \mathbb{R}^2 \rightarrow \mathbb{R}$  such that  $G(z) = K(z) + R(z)$  in such a way that  $R$  is  $C^\infty(\mathbb{R}^2)$  and  $K$  satisfies:*

- $K$  is a smooth function on  $\mathbb{R}^2 \setminus \{0\}$ , supported on the set  $\{(t, x) \in \mathbb{R}^2: x^2 + |t| \leq 1\}$  and equal to  $G$  on  $\{(t, x) \in \mathbb{R}_+ \times \mathbb{R}: x^2 + t < 1/2, t > 0\}$ .
- $K(t, x) = 0$  for  $t \leq 0, x \neq 0$  and  $K(t, -x) = K(t, x)$ .
- For every polynomial  $Q: \mathbb{R}^2 \mapsto \mathbb{R}$  of parabolic degree less than  $\zeta$ , one has

$$\int_{\mathbb{R}^2} K(t, x)Q(t, x) dx dt = 0. \quad (2.3.15)$$

*Remark 2.3.12.* Thanks to these lemmas, it is possible to *localise* on a compact support the regularising action of the heat kernel. Indeed it is also possible to show (see [Hai14, Lem 5.19]) that the map  $v \rightarrow K * v$  sends continuously  $\mathcal{C}^\alpha$  in  $\mathcal{C}^{\alpha+2}$  for any non integer  $\alpha \in \mathbb{R}$  and any distribution  $v$  not necessarily compactly supported.

In what follows for any given realisation of  $\tilde{\xi}_\varepsilon$ , the periodic lifting of  $\xi_\varepsilon$ , we will provide the construction of  $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  a sequence of models associated to  $\xi^\varepsilon$  and converging to a model  $(\hat{\Pi}, \hat{\Gamma})$  related to  $\xi$ . As a further simplification, we parametrise all possible models  $(\Pi, \Gamma)$  on  $(\mathcal{A}, \mathcal{T}, \Gamma)$  with a couple  $(\mathbf{\Pi}, f)$  where  $\mathbf{\Pi}: \mathcal{T} \rightarrow \mathcal{S}'(\mathbb{R}^2)$  and  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ . Indeed it is straightforward to check that for any given couple  $(\mathbf{\Pi}, f)$  the operators

$$\Pi_z := \mathbf{\Pi} \Gamma_{f(z)}, \quad \Gamma_{zz'} := \Gamma_{f(z')-f(z)}. \quad (2.3.16)$$

satisfy trivially the algebraic relationships in the Definition 2.3.10, because of the identity (2.3.5). Since any realisation of  $\xi_\varepsilon$  is smooth, we firstly build a model upon any deterministic smooth function  $\xi: \mathbb{R}^2 \rightarrow \mathbb{R}$  adding randomness in a second time.

**Proposition 2.3.13.** *Let  $\xi: \mathbb{R}^2 \rightarrow \mathbb{R}$  be a smooth periodic function. If the map  $\mathbf{\Pi}$  satisfies the conditions*

$$\mathbf{\Pi} \mathbf{1} = 1, \quad \mathbf{\Pi} \mathbf{X}^k \tau = z^k \mathbf{\Pi} \tau, \quad (2.3.17)$$

$$\mathbf{\Pi} \mathcal{I}_k(\sigma) = \partial^k(K * \mathbf{\Pi}(\sigma)), \quad \mathbf{\Pi} \Xi = \xi, \quad (2.3.18)$$

$$\mathbf{\Pi} \bar{\tau} \tau = \mathbf{\Pi} \bar{\tau} \mathbf{\Pi} \tau; \quad (2.3.19)$$

defined for any  $k \in \mathbb{N}^d$ ,  $\tau, \bar{\tau} \in T$  such that  $\tau \mathbf{X}^k \in T$ ,  $\mathcal{I}_k(\tau) \in T$  and  $\tau \bar{\tau} \in T$ , then there exists a unique couple  $(\mathbf{\Pi}, f)$  such that, using the identifications (2.3.16), the associated operators  $(\Pi, \Gamma)$  is a model. We call it the canonical model of  $\xi$ .

*Proof.* The hypotheses on  $\xi$  and the conditions (2.3.17) (2.3.18) implies straightforwardly that  $\mathbf{\Pi} \tau$  is a smooth function for any  $\tau \in T$  which is not a product of symbols. Therefore in this case the point-wise product on the right hand side of the equation (2.3.19) is well defined and by linearity the operator  $\mathbf{\Pi}$  exists and it is unique. In order to choose  $f$  by means of (2.3.5), we compute explicitly

$$\begin{aligned} \Pi_z(\sigma \mathcal{I}(\Xi)^m \mathbf{X}^k)(\bar{z}) &= \mathbf{\Pi} \Gamma_{f(z)}(\sigma \mathcal{I}(\Xi)^m \mathbf{X}^k) \\ &= \mathbf{\Pi}(\sigma)(\bar{z})(\bar{z} + (f(z))_{1,2})^k [(K * \xi)(\bar{z}) + (f(z))_3]^m. \end{aligned} \quad (2.3.20)$$

for any  $z, \bar{z} \in \mathbb{R}^2$ ,  $\sigma \in \{\mathcal{I}_1(\Xi), \mathcal{I}_1(\Xi)^2, \Xi, \mathbf{1}\}$  and  $k, m$  as before. Imposing the condition

$$f(z)_i = -z_i, \quad i = 1, 2 \quad f(z)_3 = -(K * \mathbf{\Pi} \Xi)(z). \quad (2.3.21)$$

we obtain immediately the bound  $\Pi_z \tau(\bar{z}) \leq C \|\bar{z} - z\|^{|\tau|}$  for some constant  $C > 0$  depending on  $\xi$  and uniformly on  $\tau \in T$ . Thus the condition (2.3.12) is satisfied. On the other hand, using the definition (2.3.21) and the explicit formulae (2.3.16) (2.3.4) it is trivial to verify the property (2.3.13) when  $\tau \in \{\mathcal{I}(\Xi), X_1, X_2\}$ . Applying the multiplicative property of  $\Gamma$  (see the Remark 2.3.5), we conclude.  $\square$

*Remark 2.3.14.* The existence of a canonical model is a general result already proved in [Hai14, Prop 8.27] but we repeat a simplified version of that proof to take in account the slightly different notation of this article. Looking at the definition of  $f$  in (2.3.21), we remark that  $f$  depends on  $\mathbf{\Pi}$  but in order to define it we do not need the multiplicative property of  $\mathbf{\Pi}$  (2.3.19), nor the smoothness of  $\xi$ . Therefore for any map  $\mathbf{\Pi}: \mathcal{T} \rightarrow \mathcal{S}'(\mathbb{R}^2)$ , choosing  $f$  as in (2.3.21), we can uniquely associate to it the couple  $\mathcal{L}(\mathbf{\Pi}) := (\Pi, \Gamma)$  where  $\Pi$  and  $\Gamma$  are given by (2.3.16).  $\mathcal{L}(\mathbf{\Pi})$  is not necessarily a model but if  $\mathbf{\Pi}\Xi = \xi \in \mathcal{C}^{-3/2-\kappa}$  for some  $0 < \kappa < 1/2$  and  $\mathbf{\Pi}$  satisfies (2.3.17), (2.3.18) (these conditions can be always formulated in case when  $\mathbf{\Pi}$  takes values on distributions and if they hold we call the map  $\mathbf{\Pi}$  *admissible*), then the proof of the Proposition 2.3.13 implies also that the operators  $\Gamma_{zz'}$  given by  $\mathcal{L}(\mathbf{\Pi})$  will always satisfy the property (2.3.13). The choice of a kernel  $K$  satisfying (2.3.15) is due in order to be compatible with our previous assumption on the symbols  $\mathcal{I}_k(\mathbf{X}^m)$ .

*Remark 2.3.15.* If  $\xi$  is also periodic in the space variable it is straightforward to prove that the canonical model  $(\Pi, \Gamma)$  associated to  $\xi$  satisfies also

$$\Pi_{(t,x+m)\tau}(t', x' + m) = \Pi_{(t,x)\tau}(t', x'), \quad \Gamma_{(t,x+m)(t',x'+m)\tau} = \Gamma_{(t,x)(t',x')\tau} \quad (2.3.22)$$

for any couple of space-time points  $z = (t, x)$ ,  $z' = (t', x')$ ,  $m \in \mathbb{Z}$ ,  $\tau \in T$ . Thus the canonical model is also *adapted* to the action of translation on  $\mathbb{R}$  (for this definition see [Hai14, Definition 3.33]). Roughly speaking this property allows to apply the notion of models also for distributions periodic in space.

*Remark 2.3.16.* Recalling the inclusion of the regularity structure  $\mathcal{T}$  in  $\mathcal{T}'$  as explained in the Proposition 2.3.8, we can immediately extend the Definition 2.3.10 to define a model  $(\Pi', \Gamma')$  over the regularity structure  $(\mathcal{A}', \mathcal{T}', \mathcal{G}')$ . In particular for any smooth function  $\xi: \mathbb{R}^2 \rightarrow \mathbb{R}$  we can define again a canonical model in this context starting from an explicit function  $\tilde{\mathbf{\Pi}}: \mathfrak{T}(R') \rightarrow \mathcal{C}^\infty$ . Using the grafting operation the application  $\tilde{\mathbf{\Pi}}$  is defined recursively for any  $k, m \in \mathbb{N}^2$ ,  $\tau, \tau' \in \mathfrak{T}$

$$\tilde{\mathbf{\Pi}}(\bullet_k)(z) := z^k, \quad \tilde{\mathbf{\Pi}}(\mathcal{E}_m^J(\tau))(z) = \tilde{\mathbf{\Pi}}(\mathcal{E}_m^I(\tau))(z) = \partial_k^m(K * \tilde{\mathbf{\Pi}}(\tau))(z),$$

$$\tilde{\mathbf{\Pi}}(\mathcal{E}_m^\Xi(\bullet_k))(z) := \partial_x^m(\xi(z)z^k), \quad \tilde{\mathbf{\Pi}}(\tau\tau')(z) := \tilde{\mathbf{\Pi}}(\tau)(z)\tilde{\mathbf{\Pi}}(\tau')(z).$$

These conditions are sufficient to define  $\tilde{\mathbf{\Pi}}$  without knowing in detail  $R'$  and the existence of a model is provided by [BHZ19, Prop. 6.12]. By construction when we restrict  $\tilde{\mathbf{\Pi}}$  on  $T$  we obtain the properties (2.3.17) (2.3.18) (2.3.19).

### 2.3.3 The BPHZ renormalisation and the BPHZ model

For any  $\varepsilon > 0$  we denote by  $\mathbf{\Pi}^\varepsilon$  and  $\mathcal{L}(\mathbf{\Pi}^\varepsilon) := (\Pi^\varepsilon, \Gamma^\varepsilon)$  the canonical model obtained by applying the Proposition 2.3.13 with a fixed a.s. realisation of  $\tilde{\xi}_\varepsilon$ . Since  $\xi_\varepsilon$  converges to  $\xi$  a.s. in the sense of distributions as recalled in the section 2.2, we would like to define a model by studying the convergence of the sequence  $(\Pi^\varepsilon, \Gamma^\varepsilon)$  as  $\varepsilon \rightarrow 0$ . Unfortunately, it is well known from [HP15] that the sequence  $\mathbf{\Pi}^\varepsilon(\mathcal{I}(\Xi)\Xi)$  does not converge as a distribution, implying that  $\mathcal{L}(\mathbf{\Pi}^\varepsilon)$  does not converge. A natural way to get rid of this

ill-posedness and to prove a general convergence result is the main content of [BHZ19] and [CH16]. The main consequence of these general results will be the existence of an explicit sequence of applications  $F_\varepsilon: \mathcal{M} \rightarrow \mathcal{M}$  such that the sequence  $F_\varepsilon(\mathcal{L}(\mathbf{\Pi}^\varepsilon)) := (\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  converges in probability to some random model. The model  $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  and the limiting model are referred in the literature as the *BPHZ renormalisation* and the *BPHZ model*.

In order to satisfy the bounds (2.3.13) for  $\hat{\Gamma}^\varepsilon$  uniformly on  $\varepsilon > 0$ , it is reasonable to write  $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  as  $\mathcal{L}(\hat{\mathbf{\Pi}}^\varepsilon)$ , for some admissible map  $\hat{\mathbf{\Pi}}^\varepsilon: \mathcal{T} \rightarrow \mathcal{S}'(\mathbb{R}^2)$  (see the Remark 2.3.14). This property can be obtained by defining a sequence of linear maps  $\{A_\varepsilon\}_{\varepsilon>0}: \mathcal{T} \rightarrow \mathcal{T}$  satisfying

$$\begin{aligned} A_\varepsilon \mathbf{1} &= \mathbf{1}, & A_\varepsilon \mathcal{I}_k(\tau) &= \mathcal{I}_k(A_\varepsilon \tau), \\ A_\varepsilon \mathbf{X}^k \tau &= \mathbf{X}^k A_\varepsilon \tau, & A_\varepsilon \Xi &= \Xi, \end{aligned} \tag{2.3.23}$$

for any  $k \in \mathbb{N}^d$  and  $\tau \in T$  such that  $\tau \mathbf{X}^k \in T$ ,  $\mathcal{I}_k(\tau) \in T$ . Indeed combining the properties of (2.3.23) with the explicit definition of  $\mathbf{\Pi}^\varepsilon$  in the proposition 2.3.13, the application  $\mathbf{\Pi}^\varepsilon A_\varepsilon$  is automatically admissible (by analogy we call  $\{A_\varepsilon\}$  an *admissible renormalisation scheme*) and we can define the couple  $\mathcal{L}(\mathbf{\Pi}^\varepsilon A_\varepsilon)$ . However the conditions (2.3.23) are not sufficient to prove that  $\mathcal{L}(\mathbf{\Pi}^\varepsilon A_\varepsilon)$  is again a model. As a matter of fact, writing the elements of  $\mathcal{T}$  as trees and embedding  $\mathcal{T}$  in  $\mathcal{T}'$  (see the Proposition 2.3.8), the BPHZ renormalisation is obtained from an explicit admissible renormalisation scheme  $\{\widetilde{M}_\varepsilon\}_{\varepsilon>0}: \mathcal{T} \rightarrow \mathcal{T}$  such that imposing  $\hat{\mathbf{\Pi}}^\varepsilon := \mathbf{\Pi}^\varepsilon \widetilde{M}_\varepsilon$  the couple  $\mathcal{L}(\hat{\mathbf{\Pi}}^\varepsilon)$  is again a model. By construction  $\widetilde{M}_\varepsilon$  is a linear map  $\widetilde{M}_\varepsilon: \mathcal{T}' \rightarrow \mathcal{T}'$  but an important consequence of the Theorem 2.3.17 will imply  $\widetilde{M}_\varepsilon(\mathcal{T}) \subset \mathcal{T}$ . Let us recall briefly the definition of  $\widetilde{M}_\varepsilon$  in terms of decorated trees as explained in [BHZ19, Sec. 6] and [Bru18, Sec.4] starting from  $\mathcal{T}'$  and its basis  $\mathfrak{T}(R')$  (see the Proposition 2.3.7).

Denoting by  $\bullet$  the combinatorial operation of the disjoint union of graphs, we consider  $\hat{T}_-$ , the set of all graphs  $\sigma$  such that

$$\sigma = \tau_1 \bullet \cdots \bullet \tau_n$$

for some  $n \geq 1$  and  $\{\tau_i\}_{i=1, \dots, n} \in \mathfrak{T}(R')$  union the empty graph  $\emptyset$ . The elements of  $\hat{T}_-$  are called *forests* and we denote by  $\hat{\mathcal{T}}_-$  the free vector space generated over  $\hat{T}_-$ , which is clearly a commutative algebra with  $\bullet$  whose unity is given by  $\emptyset$ . For any decorated tree  $\tau_\varepsilon^n$  we say that a forest  $\gamma \in \hat{T}_-$  is a *subforest* of  $\tau_\varepsilon^n$  ( $\gamma \subset \tau_\varepsilon^n$ ) if  $\gamma$  is an arbitrary subgraph of  $\tau_\varepsilon^n$  with no isolated vertices. For instance let us consider

$$\gamma_1 = \begin{array}{c} \cdot \\ \diagdown \quad \diagup \\ \cdot \end{array}, \quad \gamma_2 = \begin{array}{c} \cdot \\ \vdots \\ \cdot \end{array}, \quad \gamma_3 = \begin{array}{c} \cdot \\ \vdots \\ \cdot \quad \cdot \end{array}.$$

In this case  $\gamma_2$  and  $\gamma_3$  are both elements of  $\hat{T}_-$  but only  $\gamma_3 \subset \gamma_1$  because  $\gamma_2$  has an isolated vertex. The empty forest  $\emptyset$  is always a subforest. A decorated tree  $\tau_\varepsilon^n$  and a subforest  $\gamma = \sigma_1 \bullet \cdots \bullet \sigma_n$  such that  $\gamma \subset \tau_\varepsilon^n$  are used to define the contraction tree  $\mathcal{K}_\gamma \tau_\varepsilon^n = (\mathcal{K}_\gamma \tau, \mathcal{K}_\gamma \mathbf{n}, \mathcal{K}_\gamma \mathbf{e})$ , where

- $\mathcal{K}_\gamma \tau$  is the tree obtained from  $\tau$  replacing each  $\sigma_i$  with a node.

- Denoting by  $\bullet_1, \dots, \bullet_n$  each node associated to the contraction of the tree  $\sigma_i$ , the function  $\mathcal{K}_\gamma \mathbf{n}$  is equal to  $\mathbf{n}$  on every non contracted node of  $\mathcal{K}_\gamma \tau$  and for every  $i$ ,  $\mathbf{n}(\bullet_i) = \sum_{y \in N_{\sigma_i}} \mathbf{n}(y)$ .
- $\mathcal{K}_\gamma \mathbf{e}: E_{\mathcal{K}_\gamma} \rightarrow \mathbb{N}^2$  is equal to  $\mathbf{e}$  on every non contracted edge of  $\mathcal{K}_\gamma \tau$ .

In the previous example we have  $\mathcal{K}_{\gamma_3} \gamma_1 = \bullet$ .

Once we give  $\hat{\mathcal{T}}_-$ , we define  $\mathcal{T}_- := \hat{\mathcal{T}}_- / \mathcal{J}$  as the quotient algebra of  $\hat{\mathcal{T}}_-$  with respect to  $\mathcal{J}$ , the ideal of  $\hat{\mathcal{T}}_-$  generated by the set

$$J := \{ \tau_\epsilon^n \in \mathfrak{T}(R') : |\tau_\epsilon^n|_s \geq 0 \} \subset \hat{\mathcal{T}}_- .$$

The map  $\widetilde{M}_\epsilon$  is then defined for any  $\tau_\epsilon^n \in \mathfrak{T}(R')$  as

$$\widetilde{M}_\epsilon \tau_\epsilon^n := (h_\epsilon \otimes id) \Delta_- \tau_\epsilon^n . \quad (2.3.24)$$

We will describe the objects  $\Delta_-$  and  $h_\epsilon$  separately. First  $\Delta_- : \mathcal{T} \rightarrow \mathcal{T}_- \otimes \mathcal{T}$  is a linear map which is explicitly given for any  $\tau_\epsilon^n \in \mathfrak{T}(R')$  by the formula

$$\Delta_- \tau_\epsilon^n := \sum_{\gamma \subset \tau} \sum_{\mathbf{e}_\gamma, \mathbf{n}_\gamma \leq \mathbf{n}} \frac{1}{\mathbf{e}_\gamma!} \binom{\mathbf{n}}{\mathbf{n}_\gamma} p(\gamma, \mathbf{n}_\gamma + \pi \mathbf{e}_\gamma, \mathbf{e}_\gamma) \otimes (\mathcal{K}_\gamma \tau, \mathcal{K}_\gamma (\mathbf{n} - \mathbf{n}_\gamma), \mathcal{K}_\gamma \mathbf{e} + \mathbf{e}_\gamma) . \quad (2.3.25)$$

Let us explain the meaning of the formula (2.3.25). The first sum outside is done over all subforests  $\gamma \subset \tau$  and for any subforest  $\gamma$ , denoting by  $N_\gamma$  and  $\partial(\gamma, \tau)$  respectively the set of the nodes of  $\gamma$  and the edges in  $E_\tau$  that are adjacent to  $N_\gamma$ , the second sum is done over all functions  $\mathbf{n}_\gamma : N_\gamma \rightarrow \mathbb{N}^2$  and  $\mathbf{e}_\gamma : \partial(\gamma, \tau) \rightarrow \mathbb{N}^2$  such that for any  $x \in N_\gamma$   $\mathbf{n}_\gamma(x) \leq \mathbf{n}(x)$  with respect to the lexicographic order. Furthermore the operation  $p : \hat{\mathcal{T}}_- \rightarrow \mathcal{T}_-$  is the projection on the quotient and for any  $\mathbf{e}_\gamma : \partial(\gamma, \tau) \rightarrow \mathbb{N}^2$  the function  $\pi \mathbf{e}_\gamma : N_\gamma \rightarrow \mathbb{N}^2$  is given by

$$\pi \mathbf{e}_\gamma(x) := \sum_{e \in \partial(\gamma, \tau) : x \in e} \mathbf{e}_\gamma(e) .$$

The remaining combinatorial coefficients are finally interpreted in a multinomial sense, that is for any function  $l : S \rightarrow \mathbb{N}^2$  where  $S$  is a finite set we have

$$l! := \prod_{y \in S} (l(y))_1! (l(y))_2!$$

and similarly for the binomial coefficients. In principle the summations over  $\mathbf{n}_\gamma$  and  $\mathbf{e}_\gamma$  are done over an infinite set of values but the projection  $p$  makes the sum finite.

On the other hand the map  $h_\epsilon$  has the explicit form

$$h_\epsilon := g_\epsilon(\mathbf{\Pi}) \hat{A}_- . \quad (2.3.26)$$

The first object in (2.3.26) is given by a map  $\hat{A}_- : \mathcal{T}_- \rightarrow \hat{\mathcal{T}}_-$ . Its name is *twisted-antipode* and it is characterised as the only homomorphism (so then  $\hat{A}_-(\emptyset) = \emptyset$ ) such that for any tree  $\tau_\epsilon^n \neq \emptyset$ , denoting by  $M^\bullet$  the forest product, one has the identity

$$\hat{A}_- \tau_\epsilon^n = -M^\bullet(\hat{A}_- \otimes id)(\Delta_- \tau_\epsilon^n - \tau_\epsilon^n \otimes \mathbf{1}) . \quad (2.3.27)$$

Finally the last object  $g_\varepsilon(\mathbf{\Pi}): \widehat{\mathcal{T}}_- \rightarrow \mathbb{R}$  is the only real character on the algebra  $\widehat{\mathcal{T}}_-$  such that for any tree  $\tau_\varepsilon^n \in \mathfrak{T}(R')$

$$g_\varepsilon(\mathbf{\Pi})(\tau_\varepsilon^n) := \mathbb{E}\widetilde{\mathbf{\Pi}}^\varepsilon(\tau_\varepsilon^n)(0), \quad (2.3.28)$$

where  $\widetilde{\mathbf{\Pi}}^\varepsilon$  is the extension of  $\mathbf{\Pi}^\varepsilon$  over all the decorated trees as explained in the Remark 2.3.16. We combine all these definitions to obtain the explicit form of the application  $\widetilde{M}_\varepsilon$ .

**Theorem 2.3.17.** *By fixing  $\kappa > 0$  sufficiently small and restricting the map  $\widetilde{M}_\varepsilon$  defined in (2.3.24) on  $\mathcal{T}$ , we have  $\widetilde{M}_\varepsilon = M_\varepsilon$ , where  $M_\varepsilon: \mathcal{T} \rightarrow \mathcal{T}$  is the unique linear map satisfying  $M_\varepsilon = \text{id}$  on  $\mathcal{V}_{\mathcal{I}_1(\Xi)} \oplus \mathcal{U}$  and for any integer  $m$  and  $k \in \mathbb{N}^2$*

$$\begin{aligned} M_\varepsilon(\Xi \mathcal{I}(\Xi)^m \mathbf{X}^k) &= \Xi \mathcal{I}(\Xi)^m \mathbf{X}^k - (m C_\varepsilon^1 \mathcal{I}(\Xi)^{m-1} \mathbf{X}^k) \mathbf{1}_{m \geq 1}, \\ M_\varepsilon(\mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi)^m \mathbf{X}^k) &= \mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi)^m \mathbf{X}^k - C_\varepsilon^2 \mathcal{I}(\Xi)^m \mathbf{X}^k, \end{aligned} \quad (2.3.29)$$

where the constants  $C_\varepsilon^1$  and  $C_\varepsilon^2$  are given by

$$C_\varepsilon^1 := \mathbb{E}[\mathbf{\Pi}^\varepsilon(\Xi \mathcal{I}(\Xi))(0)] = \int_{\mathbb{R}^2} \rho_\varepsilon(z) (K * \rho_\varepsilon)(z) dz, \quad (2.3.30)$$

$$C_\varepsilon^2 := \mathbb{E}[\mathbf{\Pi}^\varepsilon(\mathcal{I}_1(\Xi)^2)(0)] = \int_{\mathbb{R}^2} (K_x * \rho_\varepsilon)^2(z) dz. \quad (2.3.31)$$

*Proof.* Thanks to the result [BHZ19, Theorem 6.17], for any  $k \in \mathbb{N}^d$ ,  $\tau \in T$  such that  $\tau \mathbf{X}^k \in T$ ,  $\mathcal{I}_k(\tau) \in T$  the map  $\widetilde{M}_\varepsilon$  always satisfies

$$\widetilde{M}_\varepsilon \mathbf{1} = \mathbf{1}, \quad \widetilde{M}_\varepsilon \mathcal{I}_k(\tau) = \mathcal{I}_k(\widetilde{M}_\varepsilon \tau), \quad \widetilde{M}_\varepsilon \mathbf{X}^k \tau = \mathbf{X}^k \widetilde{M}_\varepsilon \tau.$$

Therefore to prove the theorem it is sufficient to show for any  $m$  the identities

$$\begin{aligned} \widetilde{M}_\varepsilon(\mathcal{I}_1(\Xi) \mathcal{I}(\Xi)^m) &= \mathcal{I}_1(\Xi) \mathcal{I}(\Xi)^m, \quad \widetilde{M}_\varepsilon(\mathcal{I}(\Xi)^m) = \mathcal{I}(\Xi)^m \\ \widetilde{M}_\varepsilon(\Xi \mathcal{I}(\Xi)^m) &= \Xi \mathcal{I}(\Xi)^m - (m C_\varepsilon^1 \mathcal{I}(\Xi)^{m-1}) \mathbf{1}_{m \geq 1}, \\ \widetilde{M}_\varepsilon(\mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi)^m) &= \mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi)^m - C_\varepsilon^2 \mathcal{I}(\Xi)^m. \end{aligned} \quad (2.3.32)$$

Denoting by  $W$  the set of symbols

$$W := \{\mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi)^m, \mathcal{I}_1(\Xi) \mathcal{I}(\Xi)^m, \Xi \mathcal{I}(\Xi)^m, \mathcal{I}(\Xi)^m: m \in \mathbb{N}\}.$$

we have to calculate the operator  $\Delta_-$  and  $h_\varepsilon$  over the elements of  $W$ . In order to do that we need to know for any  $w \in W$  what are the subforests  $\gamma \subset w$  and in principle we should know the explicit form of the forests in  $\widehat{\mathcal{T}}_-$  and the rule  $R'$  which defines  $\widehat{\mathcal{T}}_-$ . However we remark for any  $w \in W$  that every subgraph  $\gamma$  included in  $w$  with no isolated vertices can be expressed as a disjoint union of trees belonging to  $\mathfrak{T}(R)$ . Thus the knowledge of  $R'$  is unnecessary and in the definition of  $\Delta_-$  in (2.3.25) we can restrict the first sum over this set. Secondly we fix  $\kappa > 0$  sufficiently small such that

the only trees of  $\mathfrak{T}(R)$  with strictly negative homogeneity that are included in  $W$  are the following

$$\left\{ \begin{array}{c} \downarrow \\ \downarrow \downarrow \\ \downarrow \downarrow \downarrow \\ \downarrow \downarrow \downarrow \downarrow \\ \downarrow \downarrow \downarrow \downarrow \downarrow \end{array} \right\}.$$

Denoting by  $\tau_m = \Xi\mathcal{I}(\Xi)^m$ , we calculate the quantity  $\widetilde{M}_\varepsilon(\tau_m)$  in case  $m = 0, 1$  explaining all the passages. Firstly we can apply the simplified version of  $\Delta_-$  in (2.3.25) and the recursive definition of  $\widehat{A}_-$  in (2.3.27) to obtain immediately

$$\Delta_- \downarrow = \emptyset \otimes \downarrow + \downarrow \otimes \mathbf{1}, \quad \widehat{A}_- \downarrow = -\downarrow,$$

$$\Delta_- \downarrow \downarrow = \emptyset \otimes \downarrow \downarrow + \downarrow \otimes \downarrow + \downarrow_{(0,1)} \otimes \downarrow + \downarrow \otimes \downarrow + \downarrow_{(0,1)} \otimes \downarrow + \downarrow \otimes \downarrow + \downarrow_{(0,1)} \otimes \downarrow + \downarrow \otimes \downarrow \otimes \mathbf{1},$$

$$\widehat{A}_- \downarrow \downarrow = -\downarrow \downarrow + \downarrow - \widehat{A}_- (\downarrow_{(0,1)}) \downarrow + \downarrow \downarrow - \widehat{A}_- (\downarrow_{(0,1)}) \downarrow - \downarrow \downarrow - \widehat{A}_- (\downarrow_{(0,1)}) \widehat{A}_- (\downarrow_{(0,1)}) \downarrow.$$

(The extra decoration with  $(0, 1)$  comes from the operation  $\pi$  in the definition of  $\Delta_-$ ). Using the general definition of  $\Delta_-$  in (2.3.25) and the recursive identity (2.3.27) the calculation in case of the symbol  $\Xi\mathbf{X}^{(0,1)}$  are given by

$$\Delta_- \downarrow_{(0,1)} = \emptyset \otimes \downarrow_{(0,1)} + \downarrow \otimes \bullet_{(0,1)} + \downarrow_{(0,1)} \otimes \mathbf{1}, \quad \widehat{A}_- (\downarrow_{(0,1)}) = -\downarrow_{(0,1)} + \downarrow \bullet_{(0,1)},$$

To complete the calculation of  $\widetilde{M}_\varepsilon$ , we need to apply  $g_\varepsilon(\mathbf{\Pi})$  on the images of the pseudo antipode. By definition of  $\widetilde{\mathbf{\Pi}}^\varepsilon$  one has

$$\begin{aligned} g_\varepsilon(\mathbf{\Pi})(\downarrow) &= \mathbb{E} \left[ \int_{\mathbb{R}^2} \rho_\varepsilon(-z_1) d\widetilde{W}_{z_1} \right] = 0, \quad g_\varepsilon(\mathbf{\Pi})(\downarrow_{(0,1)}) = 0, \\ g_\varepsilon(\mathbf{\Pi})(\downarrow \downarrow) &= \mathbb{E} \left[ \int_{\mathbb{R}^2} \rho_\varepsilon(-z_1) d\widetilde{W}_{z_1} \int_{\mathbb{R}^2} K * \rho_\varepsilon(-z_2) d\widetilde{W}_{z_2} \right] = C_\varepsilon^1. \end{aligned} \quad (2.3.33)$$

Hence we conclude firstly

$$h_\varepsilon(\downarrow) = h_\varepsilon(\downarrow_{(0,1)}) = 0, \quad h_\varepsilon(\downarrow \downarrow) = -C_\varepsilon^1. \quad (2.3.34)$$

Plugging the formulae (2.3.34) in the sums of  $\Delta_-$  we obtain the right identities of (2.3.32) for  $\widetilde{M}_\varepsilon\tau_m$   $m = 0, 1$ . Let us pass to the calculation of  $\widetilde{M}_\varepsilon\tau_m$   $m = 2, 3$ . Writing  $\Delta_- \tau_m$  and  $\widehat{A}_- \tau_m$ , a deep consequence of (2.3.34) and (2.3.33) is then all the subforests containing the trees  $\Xi$  or  $\Xi\mathbf{X}^{(0,1)}$  between the connected components will become zero after applying  $h_\varepsilon$  or  $g_\varepsilon(\mathbf{\Pi})$  thereby not giving any contribution for  $\widetilde{M}_\varepsilon$ . Denoting by  $(\dots)$  all these terms we have

$$\begin{aligned} \Delta_- \downarrow \downarrow \downarrow &= \emptyset \otimes \downarrow \downarrow \downarrow + 2 \downarrow \downarrow \otimes \downarrow + 2 \downarrow \downarrow_{(0,1)} \otimes \downarrow + \downarrow \downarrow \downarrow \otimes \mathbf{1} + (\dots), \\ \Delta_- \downarrow \downarrow \downarrow \downarrow &= \emptyset \otimes \downarrow \downarrow \downarrow \downarrow + 3 \downarrow \downarrow \downarrow \otimes \downarrow \downarrow + 6 \downarrow \downarrow_{(0,1)} \otimes \downarrow \downarrow + 2 \downarrow \downarrow \downarrow \otimes \downarrow + \downarrow \downarrow \downarrow \downarrow \otimes \mathbf{1} + (\dots), \\ \widehat{A}_- \downarrow \downarrow \downarrow &= -\downarrow \downarrow \downarrow - 2\widehat{A}_- (\downarrow \downarrow) \downarrow - 2\widehat{A}_- (\downarrow_{(0,1)}) \downarrow + (\dots), \end{aligned}$$

$$\widehat{A}_- \left( \begin{array}{c} \vdots \\ \vdots \\ \vdots \\ \vee \\ \vdots \\ \vdots \end{array} \right) = - \left( \begin{array}{c} \vdots \\ \vdots \\ \vee \\ \vdots \\ \vdots \end{array} \right) - 3\widehat{A}_- \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) + 6\widehat{A}_- \left( \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right) \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) - 2\widehat{A}_- \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) \left( \begin{array}{c} \vdots \\ \vdots \end{array} \right) + (\dots).$$

Similarly we also have

$$\Delta_- \left( \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right) = \emptyset \otimes \left( \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right) + \left( \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right) \otimes \mathbf{1} + (\dots), \quad \widehat{A}_- \left( \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right) = - \left( \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right) + (\dots).$$

Therefore the calculation of  $\widetilde{M}_\varepsilon \tau_m$  is obtained once we know the constants

$$g_\varepsilon(\mathbf{\Pi}) \left( \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right), g_\varepsilon(\mathbf{\Pi}) \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right), g_\varepsilon(\mathbf{\Pi}) \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right), g_\varepsilon(\mathbf{\Pi}) \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right).$$

The first two constants from the left are zero by definition of  $\mathbf{\Pi}^\varepsilon$  and because we are taking the expectations over a product of an *odd* number of centred Gaussian variables. On the other hand using the shorthand notation  $K_\varepsilon = K * \rho_\varepsilon$  we have

$$g_\varepsilon(\mathbf{\Pi}) \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) = \mathbb{E} \left[ \int_{\mathbb{R}^2} K_\varepsilon(-z_1) d\widetilde{W}_{z_1} \right]^2 = \int_{\mathbb{R}^2} (K_\varepsilon(z))^2 dz$$

and applying the Wick's formula for the product of four Gaussian random variables we obtain

$$\begin{aligned} g_\varepsilon(\mathbf{\Pi}) \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) &= \\ &= \mathbb{E} \left[ \int_{\mathbb{R}^2} K_\varepsilon(-z_1) d\widetilde{W}_{z_1} \int_{\mathbb{R}^2} K_\varepsilon(-z_2) d\widetilde{W}_{z_2} \int_{\mathbb{R}^2} K_\varepsilon(-z_3) d\widetilde{W}_{z_3} \int_{\mathbb{R}^2} \rho_\varepsilon(-z_4) d\widetilde{W}_{z_4} \right] \\ &= 3\mathbb{E} \left[ \int_{\mathbb{R}^2} K_\varepsilon(-z_1) d\widetilde{W}_{z_1} \int_{\mathbb{R}^2} K_\varepsilon(-z_2) d\widetilde{W}_{z_2} \right] \mathbb{E} \left[ \int_{\mathbb{R}^2} K_\varepsilon(-z_3) d\widetilde{W}_{z_3} \int_{\mathbb{R}^2} \rho_\varepsilon(-z_4) d\widetilde{W}_{z_4} \right] \\ &= 3 \int_{\mathbb{R}^2} (K_\varepsilon(z))^2 dz \int_{\mathbb{R}^2} (K_\varepsilon(z)) \rho_\varepsilon(z) dz = 3 g_\varepsilon(\mathbf{\Pi}) \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) C_\varepsilon^1. \end{aligned}$$

By replacing the values of  $g_\varepsilon(\mathbf{\Pi})$  in the above calculations of  $\widehat{A}_- \tau_m$  one has

$$h_\varepsilon \left( \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right) = h_\varepsilon \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) = h_\varepsilon \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) = 0. \quad (2.3.35)$$

Moreover the values of  $\widetilde{M}_\varepsilon \tau_m$  coincide with (2.3.32) for  $m \leq 3$ . Looking at  $\widetilde{M}_\varepsilon \tau_m$  if  $m > 3$  and  $\widetilde{M}_\varepsilon(\mathcal{I}(\Xi)^m)$ , the terms in the left factor of the sum  $\Delta_- \tau_m$  and  $\mathcal{I}(\Xi)^m$  will contain an arbitrary subforest obtained respectively from the forest product of these trees

$$\left\{ \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array}, \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array}, \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array}, \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array}, \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right\}, \quad \left\{ \begin{array}{c} \vdots \\ \vee_{(0,1)} \\ \vdots \end{array} \right\}.$$

and denoting by  $(\dots)$  the forest in the left factor which becomes zero after applying the values of  $h_\varepsilon$  in (2.3.34) (2.3.35) we obtain

$$\Delta_- \tau_m = \emptyset \otimes \tau_m + m \left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right) \otimes \underbrace{\left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right)}_{m-1} + (\dots), \quad \Delta_- \underbrace{\left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right)}_m = \emptyset \otimes \underbrace{\left( \begin{array}{c} \vdots \\ \vee \\ \vdots \end{array} \right)}_m + (\dots).$$



(The factor  $m$  appears because the tree associated to  $\Xi\mathcal{I}(\Xi)$  appears  $m$  times inside  $\tau_m$ ) Therefore we prove the first part of the equations (2.3.32). We pass to the terms of the form  $\sigma_m = \mathcal{I}_1(\Xi)^2\mathcal{I}(\Xi)^m$  and  $\eta_k = \mathcal{I}_1(\Xi)\mathcal{I}(\Xi)^k$  for  $m = 0$  and  $k \leq 1$ . Adopting the same notation as before to denote the terms that do not give any effective contribution in calculations after applying  $h_\varepsilon$  we have

$$\begin{aligned}\Delta_- \dot{\downarrow} &= \emptyset \otimes \dot{\downarrow} + \dot{\downarrow} \otimes \mathbf{1} + (\dots), & \Delta_- \dot{\downarrow}\dot{\downarrow} &= \emptyset \otimes \dot{\downarrow}\dot{\downarrow} + \dot{\downarrow} \otimes \dot{\downarrow} + \dot{\downarrow}\dot{\downarrow} \otimes \mathbf{1} + (\dots), \\ \Delta_- \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} &= \emptyset \otimes \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} + 2\dot{\downarrow} \otimes \dot{\downarrow}\dot{\downarrow} + \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \otimes \mathbf{1} + (\dots), \\ \widehat{A}_- \dot{\downarrow} &= -\dot{\downarrow} + (\dots), & \widehat{A}_- \dot{\downarrow}\dot{\downarrow} &= -\dot{\downarrow}\dot{\downarrow} + \dot{\downarrow} \otimes \dot{\downarrow} + (\dots), & \widehat{A}_- \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} &= -\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} + 2\dot{\downarrow} \otimes \dot{\downarrow} + (\dots).\end{aligned}$$

Applying the map  $\widetilde{\Pi}^\varepsilon$  we obtain also

$$\begin{aligned}g_\varepsilon(\mathbf{\Pi}) \left( \dot{\downarrow} \right) &= 0, & g_\varepsilon(\mathbf{\Pi}) \left( \dot{\downarrow}\dot{\downarrow} \right) &= \mathbb{E} \left[ \int_{\mathbb{R}^2} \partial_x K_\varepsilon(-z_1) d\widetilde{W}_{z_1} \right]^2 = C_\varepsilon^2, \\ g_\varepsilon(\mathbf{\Pi}) \left( \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \right) &= \mathbb{E} \left[ \int_{\mathbb{R}^2} K_\varepsilon(-z_1) d\widetilde{W}_{z_1} \int_{\mathbb{R}^2} \partial_x K_\varepsilon(-z_2) d\widetilde{W}_{z_2} \right] \\ &= \int_{\mathbb{R}^2} K_\varepsilon(z) \partial_x K_\varepsilon(z) dz = 0,\end{aligned}\tag{2.3.36}$$

where the first and the last identity of (2.3.36) are obtained because we take the expectation of a centred Gaussian variable and the function  $x \rightarrow K_\varepsilon(t, x) \partial_x K_\varepsilon(t, x)$  is odd in  $x$  for any  $t > 0$ . Then we obtain

$$h_\varepsilon \left( \dot{\downarrow} \right) = h_\varepsilon \left( \dot{\downarrow}\dot{\downarrow} \right) = 0, \quad h_\varepsilon \left( \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \right) = -C_\varepsilon^2,\tag{2.3.37}$$

and consequently the right identities of (2.3.32) for  $\widetilde{M}_\varepsilon \sigma_m$  and  $\widetilde{M}_\varepsilon \eta_k$ . Passing to the calculation of  $\widetilde{M}_\varepsilon \sigma_m$  for  $m = 2, 3$  we have

$$\begin{aligned}\Delta_- \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} &= \emptyset \otimes \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} + \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \otimes \dot{\downarrow} + \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \otimes \dot{\downarrow} + \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \otimes \mathbf{1} + (\dots), \\ \Delta_- \dot{\downarrow}\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} &= \emptyset \otimes \dot{\downarrow}\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} + \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \otimes \dot{\downarrow}\dot{\downarrow} + 2\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \otimes \dot{\downarrow}\dot{\downarrow} + 2\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \otimes \dot{\downarrow} + \dot{\downarrow}\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \otimes \mathbf{1} + (\dots), \\ \widehat{A}_- \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} &= -\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} - \widehat{A}_- \left( \dot{\downarrow}\dot{\downarrow} \right) \dot{\downarrow} + \widehat{A}_- \left( \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \right) \dot{\downarrow} + (\dots), \\ \widehat{A}_- \dot{\downarrow}\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} &= -\dot{\downarrow}\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} - \widehat{A}_- \left( \dot{\downarrow}\dot{\downarrow} \right) \dot{\downarrow}\dot{\downarrow} - 2\widehat{A}_- \left( \dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \right) \dot{\downarrow} - 2\widehat{A}_- \left( \dot{\downarrow}\dot{\downarrow}\dot{\downarrow}\dot{\downarrow} \right) \dot{\downarrow} + (\dots).\end{aligned}$$

Moreover using again the same algebraic notations, the Wick's formula and the definition of  $\widetilde{\Pi}^\varepsilon$  we have

$$\Delta_- \dot{\downarrow}\dot{\downarrow}_{(0,1)} = \emptyset \otimes \dot{\downarrow}\dot{\downarrow}_{(0,1)} + \dot{\downarrow}\dot{\downarrow}_{(0,1)} \otimes \mathbf{1} + (\dots), \quad \widehat{A}_- \left( \dot{\downarrow}\dot{\downarrow}_{(0,1)} \right) = -\dot{\downarrow}\dot{\downarrow}_{(0,1)} + (\dots),$$



Then the formula (2.3.39) holds as long as for any  $h \in \mathbb{R}^3$  and  $\varepsilon > 0$  one has

$$M_\varepsilon \Gamma_h = \Gamma_h M_\varepsilon . \quad (2.3.41)$$

Let us verify the identity (2.3.41) for any  $\tau \in V_\Xi \sqcup V_{\mathcal{I}_1(\Xi)^2} \sqcup V_{\mathcal{I}_1(\Xi)} \sqcup U$ . In case  $\tau \in V_{\mathcal{I}_1(\Xi)} \sqcup U$ , this identity holds trivially because  $\Gamma_h$  leaves invariant the subspace  $\mathcal{V}_{\mathcal{I}_1(\Xi)} \sqcup \mathcal{U}$  and  $M_\varepsilon$  is the identity when it is restricted to this subspace. On the other hand if  $\tau \in V_\Xi \sqcup V_{\mathcal{I}_1(\Xi)^2}$ , the multiplicative property of  $\Gamma_h$  (see the Remark 2.3.5) and the behaviour of  $M_\varepsilon$  on the polynomials in (2.3.29) reduces to verify (2.3.41) over the symbols  $\mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi)^m$  and  $\Xi \mathcal{I}(\Xi)^m$  for any  $m \geq 1$ . Writing  $h = (h_1, h_2, h_3)$  we obtain

$$\begin{aligned} M_\varepsilon \Gamma_h (\mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi)^m) &= (\mathcal{I}_1(\Xi)^2 - C_\varepsilon^2) \sum_{n=0}^m \binom{m}{n} \mathcal{I}(\Xi)^n h_3^{m-n} = \Gamma_h M_\varepsilon (\mathcal{I}_1(\Xi)^2 \mathcal{I}(\Xi)^m) , \\ M_\varepsilon \Gamma_h (\Xi \mathcal{I}(\Xi)^m) &= \sum_{n=0}^m \binom{m}{n} (\Xi \mathcal{I}(\Xi)^n - n C_\varepsilon^1 \mathcal{I}(\Xi)^{n-1}) h_3^{m-n} \\ &= \sum_{n=0}^m \binom{m}{n} \Xi \mathcal{I}(\Xi)^n h_3^{m-n} - m C_\varepsilon^1 \sum_{n'=0}^{m-1} \binom{m-1}{n'} \mathcal{I}(\Xi)^{n'} h_3^{m-1-n'} \\ &= \Xi (\mathcal{I}(\Xi) + h_3 \mathbf{1})^m - m C_\varepsilon^1 (\mathcal{I}(\Xi) + h_3 \mathbf{1})^{m-1} = \Gamma_h M_\varepsilon (\Xi \mathcal{I}(\Xi)^m) . \end{aligned}$$

Thus yielding the result. The identity (2.3.39) implies immediately the properties in the identity (2.3.22). Therefore  $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  is adapted to the action of translations.  $\square$

We study the convergence of  $\mathcal{L}(\hat{\Pi}^\varepsilon)$  in the space of models. Embedding the regularity structure  $\mathcal{T}$  into  $\mathcal{T}'$  as explained in the Proposition 2.3.8, it is possible to prove the convergence of  $\mathcal{L}(\hat{\Pi}^\varepsilon)$  using on  $T$  the general criterion exposed in [CH16, Thm. 2.15]. To apply this statement we introduce some notation. Representing all the elements of  $\tau \in T$  as decorated trees, we denote by  $E_\Xi(\tau)$  the set of edges labelled by  $\Xi$ . By construction every element  $e \in E_\Xi(\tau)$  is written uniquely as  $e = \{e_\Xi, e^\Xi\}$ , where  $e^\Xi$  is one of the terminal nodes of  $\tau$ . This decomposition allows to define the sets

$$N_\Xi(\tau) := \{e_\Xi : e \in E_\Xi(\tau)\} , \quad N^\Xi(\tau) := \{e^\Xi : e \in E_\Xi(\tau)\} , \quad N(\tau) := N_\tau \setminus N^\Xi(\tau) .$$

Moreover, expressing  $\tau$  as  $\tau_\mathfrak{n}$  for some decoration  $\mathfrak{n}, \mathfrak{e}$  we write  $\tau_\mathfrak{e}^0$  to denote the decorated tree whose decoration  $\mathfrak{n}$  is replaced by zero in every node. Let us express the convergence theorem in this context.

**Theorem 2.3.19.** *There exists a random model  $(\hat{\Pi}, \hat{\Gamma})$  such that*

$$(\hat{\Pi}_\varepsilon, \hat{\Gamma}_\varepsilon) \xrightarrow{\mathbb{P}} (\hat{\Pi}, \hat{\Gamma}) \quad (2.3.42)$$

*with respect to the metric  $\|\cdot\|_{\mathcal{M}}$ . We call  $(\hat{\Pi}, \hat{\Gamma})$  the BPHZ model.*

*Proof.* This theorem is a direct consequence of [CH16, Thm. 2.15]. Expressing the hypotheses of this theorem in our context, we obtain the thesis after checking the following property: for any  $\tau_\mathfrak{e}^\mathfrak{n} \in T$  and every subtree  $\sigma_\mathfrak{e}^\mathfrak{n}$  included in  $\tau_\mathfrak{e}^\mathfrak{n}$  such that  $\#\{N(\sigma)\} \geq 2$  one has

1) For any non-empty subset  $A \subset E_{\Xi}(\tau)$  such that  $\#\{A\} + \#\{N_{\Xi}(\sigma)\}$  is even, one has

$$|\sigma_{\mathfrak{c}}^0|_{\mathfrak{s}} + \sum_{e \in A} \mathfrak{s}(\mathfrak{t}(e)) + 3\#\{A\} > 0, \quad (2.3.43)$$

2) For any  $e \in E_{\Xi}$ ,  $|\sigma_{\mathfrak{c}}^0|_{\mathfrak{s}} - \mathfrak{s}(\mathfrak{t}(e)) > 0$  and  $|\sigma_{\mathfrak{c}}^0|_{\mathfrak{s}} > -3/2$ .

Since  $\mathfrak{s}(\Xi) = -3/2 - \kappa$ , it is sufficient to prove the condition  $|\sigma_{\mathfrak{c}}^0|_{\mathfrak{s}} > -\frac{3}{2}$  for any  $\sigma_{\mathfrak{c}}^n$  and the inequality (2.3.43) becomes

$$|\sigma_{\mathfrak{c}}^0|_{\mathfrak{s}} + \left(\frac{3}{2} - \kappa\right) \#\{A\} > 0,$$

for any non-empty subset  $A \subset E_{\Xi}(\tau)$ . It follows easily from the structure of  $T$  that the trees  $\sigma_{\mathfrak{c}}^n$  satisfying  $\#\{N(\sigma)\} \geq 2$  such that  $|\sigma_{\mathfrak{c}}^0|_{\mathfrak{s}}$  is minimal are given by

$$\sigma_{\mathfrak{c}}^0 = \begin{array}{c} \dot{\vee} \\ \vee \end{array} \text{ or } \sigma_{\mathfrak{c}}^0 = \begin{array}{c} \dot{\vee} \\ \dot{\vee} \\ \vee \end{array},$$

and  $|\sigma_{\mathfrak{c}}^0|_{\mathfrak{s}} = -1 - 2\kappa$  in both cases. Therefore we can chose the parameter  $\kappa > 0$  sufficiently small to satisfy 2) and 1).  $\square$

*Remark 2.3.20.* The BPHZ model  $(\hat{\Pi}, \hat{\Gamma})$  obtained from the Theorem 2.3.19 is an example of a random model with a.s. values on distribution and it will be the main object to formulate a new type of Itô formula for  $u$ . Recalling the inclusion of the space  $\mathcal{V}_{\Xi} \oplus \mathcal{U}$  into the regularity structure  $\mathcal{T}_{HP}$  defined in [HP15] (see the Remark 2.3.3), we can easily check that the renormalisation map  $M_{\varepsilon}$  defined in (2.3.29) and the model  $(\hat{\Pi}_{\varepsilon}, \hat{\Gamma}_{\varepsilon})$  restricted to the sector  $\mathcal{V}_{\Xi} \oplus \mathcal{U}$  coincide exactly with the renormalisation procedure developed in [HP15, Thm. 4.5] to define what in this context is called the *Itô model*. By uniqueness of the limit on this sector we can apply directly this result and we obtain immediately  $\hat{\Pi}_z \Xi = \tilde{\xi}$ , the periodic lifting of  $\xi$  and for every  $\tau \in U$ ,  $z = (t, x) \in \mathbb{R}^2$  and every smooth test function  $\psi$  such that for any  $s < t$   $\psi(s, y) = 0$  we have immediately

$$\hat{\Pi}_z(\tau \Xi)(\psi) = \int_t^{\infty} \int_{\mathbb{R}} \hat{\Pi}_z \tau(s, y) \psi(s, y) d\tilde{W}_{s,y} \quad (2.3.44)$$

We stress that (2.3.44) holds only when the test function is supported in the future. Otherwise the right hand side integrand will not be an adapted integrand and we cannot interpret  $\hat{\Pi}_z \Xi \tau(\psi)$  neither as a Skorohod integral. An explicit formula to describe the law of  $\hat{\Pi}_z \tau$  in its full generality has been developed in [CH16, Prop 4.22]. Moreover we recall that in our case we have  $\hat{\Gamma}_{zz'} = \Gamma_{\hat{f}(z') - \hat{f}(z)}$  where  $\hat{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  is given by

$$\hat{f}(z)_i = -z_i, \quad i = 1, 2 \quad \hat{f}(z)_3 = -(K * \tilde{\xi})(z).$$

The model  $(\hat{\Pi}, \hat{\Gamma})$  is also adapted to the action of translation on  $\mathbb{R}$ , as a consequence of the Proposition 2.3.18 on the converging sequence  $(\hat{\Pi}^{\varepsilon}, \hat{\Gamma}^{\varepsilon})$ .

## 2.4 Calculus on regularity structures

In this section we will show how the models  $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  and  $(\hat{\Pi}, \hat{\Gamma})$  can be used to describe respectively  $u_\varepsilon$  and  $u$  and, more generally, what kind of analytical operations we can define on a the regularity structure  $\mathcal{T}$ .

### 2.4.1 Modelled distributions

The main function of a regularity structure and a model upon that is to provide a coherent framework to approximate random distributions in a way similar polynomials approximate smooth functions via Taylor's formula. Since for any function  $f: \mathbb{R} \rightarrow \mathbb{R}$  it is possible to describe the condition  $f \in \mathcal{C}^\gamma$  in terms of  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^{|\gamma|}$ , the vector of its derivatives, for any fixed model we introduce an equivalent version of this space in our general context.

**Definition 2.4.1.** For any given parameters  $\gamma > 0$ ,  $\eta \in (-2, \gamma)$  and  $(\Pi, \Gamma)$  a model upon  $(\mathcal{A}, \mathcal{T}, \mathcal{G})$ , we define  $\mathcal{D}^{\gamma, \eta}$  as the set of all function  $U: \mathbb{R}^2 \rightarrow \mathcal{Q}_{< \gamma} \mathcal{T}$  such that for every compact set  $\mathcal{K} \subset \mathbb{R}^2$ , one has

$$|U|_{\gamma, \eta} := \sup_{z \in \mathcal{K}} \sup_{\alpha < \gamma} \frac{|U(z)|_\alpha}{|t|^{(\frac{\eta-\alpha}{2}) \wedge 0}} + \sup_{(z, z') \in \mathcal{K}^{(2)}} \sup_{\alpha < \gamma} \frac{|U(z) - \Gamma_{zz'} U(z')|_\alpha}{(|t| \wedge |t'|)^{\frac{\eta-\gamma}{2}} |z - z'|^{\gamma-\alpha}} < +\infty, \quad (2.4.1)$$

where  $\mathcal{K}^{(2)}$  denotes the set of points  $(z, z') \in \mathcal{K}^2$  such that  $|z - z'| \leq 1/2\sqrt{|t| \wedge |t'|}$ . The elements of  $\mathcal{D}^{\gamma, \eta}$  are called *modelled distributions*.

*Remark 2.4.2.* The definition of the set  $\mathcal{D}^{\gamma, \eta}$  does depend depend in a crucial way on the underlying model  $(\Pi, \Gamma)$ . To remark this dependency we will adopt for the same set the alternative notation  $\mathcal{D}^{\gamma, \eta}(\Pi)$ . Similarly we recall that the quantities  $\|U\|_{\gamma, \eta}$  depend on the compact set  $\mathcal{K}$  but we avoid to put the symbol  $\mathcal{K}$  in the notation because for our finite time horizon setting, we will henceforth prove the results on a fixed compact set  $\mathcal{K} \subset \mathbb{R}^2$  containing  $[0, T] \times [0, 1]$ . The presence of an extra parameter  $\eta$  allows more freedom than the classical  $\mathcal{C}^\gamma$  spaces. In this way the coordinates of  $U$  are to blow at rate  $\eta$  near the set  $P = \{(t, x) \in \mathbb{R}^2: t = 0\}$  and the condition  $\eta > -2$  is put to keep this singularity integrable. By definition of  $\mathcal{D}^{\gamma, \eta}$ , for any value  $\gamma \geq \gamma' > 0$  and  $U \in \mathcal{D}^{\gamma, \eta}$  the projection  $\mathcal{Q}_{< \gamma'} U \in \mathcal{D}^{\gamma', \eta}$ .

For any given model the couple  $(\mathcal{D}^{\gamma, \eta}(\Pi), |\cdot|_{\gamma, \eta})$  is clearly a Banach space. Since we will consider modelled distributions belonging to different models, for any couple of models  $(\Pi, \Gamma)$  and  $(\Pi', \Gamma')$  and modelled distributions  $U \in \mathcal{D}^{\gamma, \eta}(\Pi)$ ,  $U' \in \mathcal{D}^{\gamma, \eta}(\Pi')$  we define the quantity

$$\|U, U'\|_{\gamma, \eta} := \sup_{z, w, \alpha} \frac{|U(z) - U'(z) - \Gamma_{zw} U(w) + \Gamma'_{zw} U'(w)|_\alpha}{(|t| \wedge |t'|)^{\frac{\eta-\gamma}{2}} |z - z'|^{\gamma-\alpha}} + \sup_{z, \alpha} \frac{|U'(z) - U(z)|_\alpha}{|t|^{(\frac{\eta-\alpha}{2}) \wedge 0}},$$

where the parameters  $z, w, \alpha$  belong to the same sets as the quantity (2.4.1). This function together with the norm  $\|\cdot\|_{\mathcal{M}}$  on models endows the fibred space

$$\mathcal{M} \times \mathcal{D}^{\gamma, \eta} := \{((\Pi, \Gamma), U): (\Pi, \Gamma) \in \mathcal{M}, U \in \mathcal{D}^{\gamma, \eta}(\Pi)\}$$

of a complete metric structure (e.g. we can use the distance  $\|\cdot, \cdot\|_{\gamma, \eta} + \|\cdot, \cdot\|_{\mathcal{M}}$ ). Combining the knowledge of a model  $(\Pi, \Gamma) \in \mathcal{M}$  and  $U \in \mathcal{D}^{\gamma, \eta}(\Pi)$ , it is possible to define uniquely a distribution such that the coordinates of  $U$  has the same role of the derivatives of a function in the Taylor's formula. This association takes the name of reconstruction theorem and it is one of the main theorem in the theory of regularity structures (for its proof see [Hai14, Sec. 3, Sec. 6]).

**Theorem 2.4.3** (Reconstruction theorem). *For any  $(\Pi, \Gamma) \in \mathcal{M}$  there exists a unique map  $\mathcal{R}: \mathcal{D}^{\gamma, \eta}(\Pi) \rightarrow \mathcal{S}'(\mathbb{R}^2)$ , called the reconstruction operator, satisfying the following properties:*

- (Generalised Taylor expansion) for any compact set  $\mathcal{K} \subset \mathbb{R}^2$  there exists a constant  $C > 0$  such that

$$|(\mathcal{R}U - \Pi_z U(z))(\eta_z^\lambda)| \leq C\lambda^\gamma \quad (2.4.2)$$

uniformly over  $\eta \in \mathcal{B}_2$ ,  $\lambda \in (0, 1]$  and  $z \in \mathcal{K}$ ;

- the distribution  $\mathcal{R}U \in \mathcal{C}^{\alpha_U \wedge \eta}$  where  $\alpha_U := \min\{a \in \mathcal{A}: \mathcal{Q}_a U \neq 0\}$  and in case  $\alpha \wedge \eta = 0$  we set by convention  $\mathcal{C}^0$  the space of locally bounded functions;
- (local Lipschitz property) for any fixed  $R > 0$  and all couples  $(\Pi', \Gamma'), (\Pi, \Gamma) \in \mathcal{M}$ ,  $U \in \mathcal{D}^{\gamma, \eta}(\Pi)$ ,  $U' \in \mathcal{D}^{\gamma, \eta}(\Pi')$  such that  $\|U; U'\|_{\gamma, \eta} + \|(\Pi, \Gamma); (\Pi', \Gamma')\|_{\mathcal{M}} < R$  and  $\alpha_U = \alpha_{U'} = \alpha$ , denoting by  $\mathcal{R}$  and  $\mathcal{R}'$  the respective reconstruction operators, there exists a constant  $C > 0$  depending on  $R$  such that

$$\|\mathcal{R}'U' - \mathcal{R}U\|_{\mathcal{C}^{\alpha \wedge \eta}} \leq C (\|U, U'\|_{\gamma, \eta} + \|(\Pi, \Gamma), (\Pi', \Gamma')\|_{\mathcal{M}}). \quad (2.4.3)$$

*Remark 2.4.4.* The reconstruction map has in some rare cases an explicit expression. For instance if  $\Pi_z \tau$  is a continuous function for every  $\tau \in T$  (like the model  $\mathcal{L}(\hat{\Pi}^\varepsilon)$  or  $\mathcal{L}(\hat{\Pi}^\varepsilon)$  for any  $\varepsilon > 0$ ) and  $U \in \mathcal{D}^{\gamma, \eta}(\Pi)$ , then  $\mathcal{R}U$  is a continuous function given explicitly by

$$\mathcal{R}(U)(z) = \Pi_z(U(z))(z). \quad (2.4.4)$$

Introducing the space  $\mathcal{D}_{\mathcal{U}}^{\gamma, \eta}$  of all modelled distributions taking values in  $\mathcal{U}$ , the identity (2.4.4) holds also if  $(\Pi, \Gamma)$  is a generic model and  $U \in \mathcal{D}_{\mathcal{U}}^{\gamma, \eta}(\Pi)$ , because the elements of the canonical basis of  $\mathcal{U}$  have all non negative homogeneity (see for further details in [Hai14, Sec. 3.4]). Concerning the regularity of  $\mathcal{R}U$ , the result stated in the Theorem 2.4.3 is optimal because of the presence the parameter  $\eta$  in the definition and the possible explosion of the components of  $U$ . However if we forget the behaviour at 0 it is also possible to prove  $\mathcal{R}U \in \mathcal{C}^{\beta_U}(\mathbb{R}^2 \setminus P)$  where  $\beta_U := \min\{a \in \mathcal{A} \setminus \mathbb{N}: \mathcal{Q}_a U \neq 0\}$  (see [Hai14, Sec. 6]) and the local Lipschitz property (2.4.3) holds on the same space  $\mathcal{C}^{\beta_U}(\mathbb{R}^2 \setminus P)$ . We finally conclude that for any value  $\gamma \geq \gamma' > 0$  and  $U \in \mathcal{D}^{\gamma, \eta}$  we have the identity  $\mathcal{R}\mathcal{Q}_{<\gamma'} U = \mathcal{R}U$ , therefore to define correctly the distribution  $\mathcal{R}U$  is sufficient to fix  $\gamma > 0$  such that  $\mathcal{Q}_{<\gamma} \mathcal{T}$  is generated by the set  $\{\tau \in T: |\tau| \leq 0\}$ . In what follows we will denote by  $\hat{\mathcal{R}}_\varepsilon$  (respectively  $\hat{\mathcal{R}}$ ) the reconstruction operator associated to the space  $\mathcal{D}^{\gamma, \eta}(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  (resp.  $\mathcal{D}^{\gamma, \eta}(\hat{\Pi}, \hat{\Gamma})$ ).

*Remark 2.4.5.* If a model  $(\Pi, \Gamma)$  is adapted to the action of the translations (see the equations (2.3.22)) we and the function  $U$  is periodic in the space variable on  $\mathbb{R}^2$ ,

then using the general result [Hai14, Prop. 3.38] we obtain also  $\mathcal{R}U = \tilde{u}$  for some  $u \in \mathcal{C}^{\alpha \vee \eta}(\mathbb{R} \times \mathbb{T})$ , with an abuse of notation we can identify  $\mathcal{R}U$  with  $u$ .

By fixing a compact set Using the shorthand notation  $\mathbf{1}_+ = \mathbf{1}_{(0, \infty) \times \mathbb{R}}$ , we introduce  $\mathbf{1}_+ \Xi: \mathbb{R}^2 \rightarrow \mathcal{T}$ , defined for any  $z = (t, x) \in \mathbb{R}^2$  by:

$$(\mathbf{1}_+ \Xi)(z) := \mathbf{1}_+(z) \Xi = \begin{cases} \Xi & \text{if } t > 0, \\ 0 & \text{Otherwise.} \end{cases}$$

For any fixed realisation of  $\xi$  and any choice of the parameters  $\gamma > 0$  and  $-2 < \eta < \gamma$ , the definitions of  $\hat{\Gamma}$  and  $\Gamma^\varepsilon$  implies immediately  $\mathbf{1}_+ \Xi \in \mathcal{D}^{\gamma, \eta}(\hat{\Pi}^\varepsilon)$  for all  $\varepsilon > 0$  and  $\mathbf{1}_+ \Xi \in \mathcal{D}^{\gamma, \eta}(\hat{\Pi})$ . The reconstruction of  $\mathbf{1}_+ \Xi$  in both cases can be explicitly calculated.

**Proposition 2.4.6.** *for any  $z \in [0, T] \times \mathbb{T}$  one has*

$$\hat{\mathcal{R}}_\varepsilon(\mathbf{1}_+ \Xi)(z) = \mathbf{1}_+(z) \xi_\varepsilon(z), \quad \hat{\mathcal{R}}(\mathbf{1}_+ \Xi) = \mathbf{1}_{[0, \infty)} \xi. \quad (2.4.5)$$

where the second identity holds a.s. as distributions.

*Proof.* As we recalled in the Remark 2.4.4, to prove the first part (2.4.5) we can apply directly the identity (2.4.4) obtaining the result trivially. Using the Theorem 2.3.19, related to the convergence of models and the local Lipschitz continuity of the reconstruction map, the distribution  $\hat{\mathcal{R}}_\varepsilon(\mathbf{1}_+ \Xi)$  converges in probability to  $\hat{\mathcal{R}}(\mathbf{1}_{[0, +\infty)} \Xi)$  with respect to the topology of  $\mathcal{C}^{-3/2 - \kappa}(\mathbb{R} \times \mathbb{T})$ . Since  $\xi_\varepsilon$  converges in probability to  $\xi$  with respect to the topology  $\mathcal{C}^{-3/2 - \kappa}(\mathbb{R} \times \mathbb{T})$  (see [Hai14, Lem 10.2]) and the operator  $\mathbf{1}_{[0, +\infty)}$  (introduced in the section 2.2) extends continuously the multiplication with the indicator  $\mathbf{1}_+$ , then  $\mathbf{1}_+(z) \xi_\varepsilon$  converges in probability to  $\mathbf{1}_{[0, +\infty)} \xi$  with respect to the same topology. We conclude by uniqueness of the limit.  $\square$

*Remark 2.4.7.* Denoting by  $\mathcal{R}_\varepsilon$  the reconstruction operator with respect to the canonical model  $(\Pi^\varepsilon, \Gamma^\varepsilon)$  we have also  $\mathbf{1}_+ \Xi \in \mathcal{D}^{\gamma, \eta}(\Pi^\varepsilon)$  and  $\mathcal{R}_\varepsilon \mathbf{1}_+ \Xi = \hat{\mathcal{R}}_\varepsilon \mathbf{1}_+ \Xi$ , because  $\hat{\Pi}_z^\varepsilon \Xi = \Pi_z^\varepsilon \Xi$  for any  $\varepsilon > 0$ . Using the same argument above one has  $\mathcal{R}_\varepsilon \mathbf{1}_+ \Xi$  converges in probability to  $\mathbf{1}_{[0, +\infty)} \xi$  as before. Nevertheless the sequence  $(\Pi^\varepsilon, \Gamma^\varepsilon)$  does not converge and we cannot interpret  $\mathbf{1}_{[0, +\infty)} \xi$  as the reconstruction of some modelled distribution, unless we study the model  $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$ .

## 2.4.2 Operations with the stochastic heat equation

Although modelled distributions look very unusual, the reconstruction theorem associates to them a distribution, which is a classical analytical object. Under this identification it is possible to lift up some operations on the  $\mathcal{C}^\gamma$  spaces directly at the level of the modelled distributions as it was explained in detail in [Hai14, Sec. 4, 5, 6]. Moreover this “lifting” procedure is also continuous with respects to the intrinsic topology of the modelled distributions. In what follows we will briefly recall them to put them in relation with the stochastic heat equation.

## Convolution

The first operation to define is the convolution with  $G$ , the heat kernel on  $\mathbb{R}$ . In other terms, we analyse under which conditions we can associate continuously to any  $((\Pi, \Gamma), V) \in \mathcal{M} \times \mathcal{D}^{\gamma, \eta}$  one modelled distribution  $\mathcal{P}(V) \in \mathcal{D}^{\bar{\gamma}, \bar{\eta}}(\Pi)$  such that

$$\mathcal{R}(\mathcal{P}(V)) = G * \mathcal{R}V. \quad (2.4.6)$$

For our purposes we are not interested to describe this operation in its full generality. Indeed recalling the formulae (2.2.4) and (2.4.5) it is sufficient to define  $\mathcal{P}$  only in the case of the modelled distribution  $V = \mathbf{1}_+ \Xi$  to have an expression of  $u_\varepsilon$  and  $u$ , the solution of (2.1.6) and (2.1.1), as the reconstruction of some modelled distributions. In this case we can restate the convolution with  $G$  with the convolution with two other kernels thanks to this technical lemma (its proof is a direct consequence of [Hai14, Lemma 7.7]).

**Lemma 2.4.8** (Second decomposition). *For any fixed  $T > 0$ , there exists a function  $\tilde{R}: \mathbb{R}^2 \rightarrow \mathbb{R}$  such that*

- For every distribution  $v \in \mathcal{C}^\beta(\mathbb{R} \times \mathbb{T})$  with  $\beta > -2$  non integer and supported on  $[0, +\infty)$  one has

$$(G * \tilde{v})(z) = (K * \tilde{v})(z) + (\tilde{R} * \tilde{v})(z), \quad (2.4.7)$$

where  $K$  is the function introduced in the Lemma 2.3.11,  $z \in (-\infty, T + 1] \times \mathbb{R}$  and  $\tilde{v}$  is the periodic lifting of  $v$ .

- $\tilde{R}$  is smooth,  $\tilde{R}(t, x) = 0$  for  $t \leq 0$  and it is compactly supported.

Thanks to this decomposition it is sufficient to write  $\mathcal{P} = \mathfrak{K} + \mathfrak{R}$ , for some operators  $\mathfrak{K}$  and  $\mathfrak{R}$  satisfying

$$\mathcal{R}(\mathfrak{K}(V)) = K * \mathcal{R}V, \quad \mathcal{R}(\mathfrak{R}(V)) = \tilde{R} * \mathcal{R}V. \quad (2.4.8)$$

Considering the case of  $\mathfrak{R}$ , we remark that for any distribution  $v$  supported on positive times the distribution  $(\tilde{R} * v)$  will always be a smooth function on  $\mathbb{R}^2$  by hypothesis on  $\tilde{R}$ . Thus for any fixed couple  $((\Pi, \Gamma), V) \in \mathcal{M} \times \mathcal{D}^{\gamma, \eta}$  such that  $\mathcal{R}V$  is supported on  $\mathbb{R}_+ \times \mathbb{R}$  and the model  $(\Pi, \Gamma)$  is of the form  $\mathcal{L}(\mathbf{\Pi})$  for some admissible map  $\mathbf{\Pi}$  (see the Remark 2.3.14), the operator  $\mathfrak{R}$  can be easily defined for any  $\bar{\gamma} > 0$  as the lifting of the  $\bar{\gamma}$ -th order Taylor polynomial of  $(\tilde{R} * \mathcal{R}V)$ , that is:

$$\mathfrak{R}(V)(z) := \sum_{|k| < \bar{\gamma}} (\partial^k \tilde{R} * (\mathcal{R}V))(z) \frac{\mathbf{X}^k}{k!}. \quad (2.4.9)$$

From this definition it is straightforward to check that  $\mathfrak{R}(V) \in \mathcal{D}^{\bar{\gamma}, \bar{\eta}}(\Pi)$  for any  $\bar{\gamma} > 0$  and  $-2 < \bar{\eta} < \bar{\gamma}$  and that it satisfies the second identity of (2.4.8). Moreover the application  $\mathfrak{R}: \mathcal{M} \times \mathcal{D}^{\gamma, \eta} \rightarrow \mathcal{D}^{\bar{\gamma}, \bar{\eta}}$  is also continuous with respect to the topology of  $\mathcal{M}$ , as a consequence of [Hai14, Lem. 7.3]. This continuity property is a consequence of the compact support of  $\tilde{R}$  and it is the main reason behind the choice of a decomposition



of  $G$  different from the Lemma 2.3.11. On the other hand the kernel  $K$  is not a smooth and the definition of  $\mathfrak{K}$  depends on the model and it is given as a consequence of this general result (for its proof see the “Extension theorem” [Hai14, Thm. 5.14] and the “Multi-level Schauder estimates” [Hai14, Thm. 5.14, Prop 6.16]).

**Proposition 2.4.9.** *For any couple  $((\Pi, \Gamma), V) \in \mathcal{M} \times \mathcal{D}^{\gamma, \eta}$  where  $(\Pi, \Gamma)$  is of the form  $\mathcal{L}(\Pi)$  for some admissible map  $\Pi$  and  $\gamma > 0$ ,  $3/2 - \kappa < \eta < \gamma$  are not integers, there exists a regularity structure  $(\mathcal{A}_2, \mathcal{T}_2, \mathcal{G}_2)$  including  $(\mathcal{A}, \mathcal{T}, \mathcal{G})$ , a linear map  $\tilde{I}: \mathcal{T} \rightarrow \mathcal{T}_2$  satisfying  $\tilde{I}(\Xi) = \mathcal{I}(\Xi)$  and a model  $(\Pi^2, \Gamma^2)$  extending  $(\Pi, \Gamma)$  on  $\mathcal{T}_2$  such that, imposing  $\bar{\gamma} = \gamma + 2$ ,  $\bar{\eta} = \alpha_V \wedge \eta + 2$ , the applications  $\mathcal{N}: \mathcal{M} \times \mathcal{D}^{\gamma, \eta} \rightarrow \mathcal{D}_U^{\bar{\gamma}, \bar{\eta}}$ ,  $J: \mathcal{T} \rightarrow \mathcal{D}_U^{\bar{\gamma}, \bar{\eta}}$*

$$\mathcal{N}(V)(z) := \sum_{|k| < \gamma + 2} \left( (\partial^k K) * (\mathcal{R}V - \Pi_z V(z)) \right) (z) \frac{\mathbf{X}^k}{k!}, \quad (2.4.10)$$

$$J(z)\tau := \sum_{|k| < |\tau| + 2} \left( (\partial^k K * \Pi_z \tau) \right) (z) \frac{\mathbf{X}^k}{k!}, \quad (2.4.11)$$

are well defined and the application

$$\mathfrak{K}(V)(z) := \mathcal{Q}_{< \bar{\gamma}}(\tilde{I}(V))(z) + J(z)V(z) + \mathcal{N}(V)(z), \quad (2.4.12)$$

is a map  $\mathfrak{K}: \mathcal{M} \times \mathcal{D}^{\gamma, \eta} \rightarrow \mathcal{D}^{\bar{\gamma}, \bar{\eta}}(\Pi_2)$  satisfying the first identity of (2.4.8) without any restriction on the support of  $\mathcal{R}(V)$ . Moreover  $\mathfrak{K}$  is also continuous with respect to the topology of  $\mathcal{M}$ .

Choosing in the definition of  $\mathfrak{K}$  the same parameters  $\bar{\gamma}$  and  $\bar{\eta}$  of  $\mathfrak{K}$ , the application  $\mathcal{P} = \mathfrak{K} + \mathfrak{A}$  is a well defined map  $\mathcal{P}: \mathcal{M} \times \mathcal{D}^{\gamma, \eta} \rightarrow \mathcal{D}^{\bar{\gamma}, \bar{\eta}}$  which depends continuously on the topology of the models. We will denote by  $\hat{\mathfrak{K}}_\varepsilon, \hat{\mathfrak{A}}_\varepsilon, \hat{\mathcal{P}}_\varepsilon$  (resp.  $\hat{\mathfrak{K}}, \hat{\mathfrak{A}}, \hat{\mathcal{P}}$ ) the operators  $\mathfrak{K}, \mathfrak{A}$  and  $\mathcal{P}$  associated to the model  $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  (resp.  $(\hat{\Pi}, \hat{\Gamma})$ ). Let us we calculate  $\hat{\mathcal{P}}_\varepsilon(\mathbf{1}_+\Xi)$  and  $\hat{\mathcal{P}}(\mathbf{1}_+\Xi)$  in this case.

**Proposition 2.4.10.** *For any  $\gamma > 0$  and every  $-3/2 + \kappa < \eta < \gamma$  non integer, using the shorthand notation  $\bar{\gamma} = \gamma + 2$ , the modelled distribution  $U_\varepsilon := \hat{\mathcal{P}}_\varepsilon(\mathbf{1}_+\Xi)$  and  $U := \hat{\mathcal{P}}(\mathbf{1}_+\Xi)$  belong respectively to  $\mathcal{D}_U^{\bar{\gamma}, 1/2 - \kappa}(\hat{\Pi}^\varepsilon)$  and  $\mathcal{D}_U^{\bar{\gamma}, 1/2 - \kappa}(\hat{\Pi})$  and they are both given explicitly for any  $z = (t, x) \in [0, T] \times \mathbb{R}$  by the formulae*

$$U_\varepsilon(z) = \tilde{u}_\varepsilon(z)\mathbf{1} + \mathbf{1}_+(z) \dot{+} \sum_{0 < |k| < \bar{\gamma}} v_\varepsilon^k(z) \frac{\mathbf{X}^k}{k!}, \quad (2.4.13)$$

$$U(z) = \tilde{u}(z)\mathbf{1} + \mathbf{1}_+(z) \dot{+} \sum_{0 < |k| < \bar{\gamma}} v^k(z) \frac{\mathbf{X}^k}{k!}, \quad (2.4.14)$$

where  $v_\varepsilon^k(z) = (\partial_k \tilde{R} * \widetilde{\mathbf{1}_+\xi_\varepsilon})(z)$  and  $v^k(z) = (\partial_k \tilde{R} * \widetilde{\mathbf{1}_{[0, +\infty)}\xi})(z)$ . Moreover we have for any  $z \in [0, T] \times \mathbb{R}$ ,  $\hat{\mathcal{R}}_\varepsilon(U_\varepsilon)(z) = u_\varepsilon$  and  $\hat{\mathcal{R}}(U)(z) = u(z)$ .

*Proof.* The proposition is a direct consequence of the definition of  $\mathfrak{K}$  in the Proposition 2.4.9. In particular we have immediately  $\hat{\mathfrak{K}}_\varepsilon(\mathbf{1}_+\Xi) \in \mathcal{D}_U^{\bar{\gamma}, 1/2-\kappa}(\hat{\Pi}^\varepsilon)$  and  $\hat{\mathfrak{K}}(\mathbf{1}_+\Xi) \in \mathcal{D}_U^{\bar{\gamma}, 1/2-\kappa}(\hat{\Pi})$  because  $\alpha_{\mathbf{1}_+\Xi} = -3/2 - \kappa$ . Considering the explicit formula (2.4.12), which defines  $\mathfrak{K}$ , by definition of  $\mathbf{1}_+\Xi$  we have for any  $z \in \mathbb{R}^2$

$$\hat{\mathcal{R}}_\varepsilon(\mathbf{1}_+\Xi)(z) = \mathbf{1}_+(z)\hat{\Pi}_z^\varepsilon(\Xi)(z), \quad \hat{\mathcal{R}}(\mathbf{1}_+\Xi) = \mathbf{1}_{[0,+\infty)}\hat{\Pi}_z(\Xi).$$

Hence the function  $\mathcal{N}(\mathbf{1}_+\Xi)$  defined in (2.4.10) is constantly equal to zero in case of  $\hat{\mathfrak{K}}_\varepsilon$  and  $\hat{\mathfrak{K}}$ . Summing up the definition of  $\tilde{I}$ , the definition of  $J$  and the identity (2.4.5), we obtain

$$\begin{aligned} \hat{\mathfrak{K}}_\varepsilon(\mathbf{1}_+\Xi)(z) &= (K * \widetilde{\mathbf{1}_+\xi_\varepsilon})(z)\mathbf{1} + \mathbf{1}_+(z)\mathring{\mathbf{1}}, \\ \hat{\mathfrak{K}}(\mathbf{1}_+\Xi)(z) &= (K * \widetilde{\mathbf{1}_{[0,+\infty)}\xi})(z)\mathbf{1} + \mathbf{1}_+(z)\mathring{\mathbf{1}}. \end{aligned}$$

Applying again the identity (2.4.5) and the definition of  $\mathfrak{K}$  in (2.4.9), the formulae (2.4.13) (2.4.14) follows from the distributional identities (2.4.7) and (2.2.6). The last identities on the reconstruction follow straightforwardly from the general identity (2.4.6) and the property that the kernels  $K$  and  $R$  are 0 for negative times. Thus for any  $z = (t, x) \in [0, T] \times \mathbb{R}$  one has

$$(K * \widetilde{\mathbf{1}_{[0,+\infty)}\xi})(z) = (K * \widetilde{\mathbf{1}_{[0,t]}\xi})(z), \quad (R * \widetilde{\mathbf{1}_{[0,+\infty)}\xi})(z) = (R * \widetilde{\mathbf{1}_{[0,t]}\xi})(z)$$

and similarly with  $\xi_\varepsilon$ . Thereby obtaining the thesis.  $\square$

*Remark 2.4.11.* For any  $\varepsilon > 0$  it is also possible to consider  $\mathcal{P}_\varepsilon$ , the convolution operator associated the canonical model  $(\Pi^\varepsilon, \Gamma^\varepsilon)$ . Following the Remark 2.4.7 related to the modelled distribution  $\mathbf{1}_+\Xi$  in the case of the canonical model and the proof of the Proposition 2.4.10, we obtain also that  $\mathcal{P}_\varepsilon(\mathbf{1}_+\Xi) \in \mathcal{D}_U^{\bar{\gamma}, 1/2-\kappa}(\Pi^\varepsilon)$  and the identity  $\mathcal{P}_\varepsilon(\mathbf{1}_+\Xi)(z) = U_\varepsilon(z)$  for any  $z \in \mathbb{R}^2$ , implying  $\mathcal{R}_\varepsilon U_\varepsilon = u_\varepsilon$ . Another consequence of the Proposition 2.4.10 is then  $\bar{\gamma} > 2$ . However in order to reconstruct  $u_\varepsilon$  and  $u$  from  $U_\varepsilon$  and  $U$ , as explained in the Remark 2.4.4 we can relax this condition by writing  $U_\varepsilon$  and  $U$  as elements of  $\mathcal{D}^{\gamma', 1/2-\kappa}$  for some  $0 < \gamma' \leq \gamma$ .

Writing  $u_\varepsilon$  and  $u$  as the reconstruction of some modelled distribution, we obtain immediately the following convergence.

**Proposition 2.4.12.** *Let  $u_\varepsilon$  and  $u$  be the solutions respectively of the equations (2.1.6) and (2.1.1). Then as  $\varepsilon \rightarrow 0^+$*

$$\sup_{(t,x) \in [0,T] \times \mathbb{T}} |u_\varepsilon(t, x) - u(t, x)| \xrightarrow{\mathbb{P}} 0. \quad (2.4.15)$$

Moreover  $u_\varepsilon \rightarrow u$  in probability with respect to the topology of  $\mathcal{C}^{1/2-\kappa}((0, T) \times \mathbb{T})$ .

*Proof.* Thanks to the Proposition 2.4.10, the Proposition 2.4.9 and the local Lipschitz property of the reconstruction map, there exists a continuous map  $\Psi: \mathcal{M} \rightarrow \mathcal{C}^0([0, T] \times \mathbb{T})$  such that  $u_\varepsilon = \Psi((\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon))$  and  $u = \Psi((\hat{\Pi}, \hat{\Gamma}))$ . Thus the limit (2.4.15) is a direct consequence of the Theorem 2.3.19. Restricting  $u_\varepsilon$  and  $u$  on  $(0, T) \times \mathbb{T}$  and following the Remark 2.4.4 on the regularity of the reconstruction operator outside the origin, we obtain that  $\Psi$  is also a continuous map  $\Psi: \mathcal{M} \rightarrow \mathcal{C}^{1/2-\kappa}((0, T) \times \mathbb{T})$ , concluding in the same way.  $\square$

### Composition

For any  $((\Pi, \Gamma), V) \in \mathcal{M} \times \mathcal{D}_{\mathcal{U}}^{\gamma, \eta}$ , the general property of the reconstruction operator ensures us that  $\mathcal{R}V$  is a function (see the Remark 2.4.4). In particular for any function  $h: \mathbb{R} \rightarrow \mathbb{R}$  sufficiently smooth we can find a modelled distribution  $H(V)$  such that

$$\mathcal{R}(H(V)) = h \circ \mathcal{R}V. \quad (2.4.16)$$

We call this operation the *lifting of  $h$*  and we write it as a linear map  $H: \mathcal{D}_{\mathcal{U}}^{\gamma, \eta}(\Pi) \rightarrow \mathcal{D}_{\mathcal{U}}^{\gamma, \eta}(\Pi)$  (the lifting of a function  $f$  will always be denoted in capital letters  $F$ ). For any smooth function  $h$  the function  $H(V): \mathbb{R}^2 \rightarrow \mathcal{U}$  is given by

$$H(V)(z) := \mathcal{Q}_{< \gamma} \sum_{k \geq 0} \frac{h^{(k)}(v(z))}{k!} (V(z) - \mathcal{R}(V)(z)\mathbf{1})^k, \quad (2.4.17)$$

where the power  $k$  is done with respect to the product of symbols. Denoting by  $C_b^n(\mathbb{R})$  the space of  $C^n$  functions with all bounded derivatives up to the  $n$ -th order, we apply the general theory to deduce a sufficient condition to define the lifting  $H(V)$  when  $h$  is not necessarily smooth.

**Proposition 2.4.13.** *For any  $\gamma > 0$ ,  $0 \leq \eta < \gamma$ , recalling the notation  $\beta_V := \min\{a \in \mathcal{A} \setminus \mathbb{N}: \mathcal{Q}_a V \neq 0\}$  the lifting of  $h$  in (2.4.17) is well defined and it depends continuously on the topology of  $\mathcal{M} \times \mathcal{D}_{\mathcal{U}}^{\gamma, \eta}$  if  $h \in C_b^\beta(\mathbb{R})$  where  $\beta$  is the smallest integer  $\beta \geq ((\gamma/\beta_V) \vee 1) + 1$ .*

*Proof.* Following the general results [Hai14, Thm. 6.13], [HP15, Prop. 3.11], the map  $H \rightarrow H(V)$  is local Lipschitz with respect to the metric  $\|\cdot, \cdot\|_{\gamma, \eta} + \|\cdot, \cdot\|_{\mathcal{M}}$  as long as  $h$  is a  $\lambda$ -Hölder function where  $\lambda \geq ((\gamma/\beta_V) \vee 1) + 1$ . Thus we obtain the thesis.  $\square$

*Remark 2.4.14.* Applying this proposition in case of  $U_\varepsilon$  and  $U$  we obtain easily  $\beta_{U_\varepsilon} = \beta_U = |\mathcal{I}(\Xi)| = 1/2 - \kappa$ . Thus when we consider for any  $\gamma' > 0$  the projection on  $\mathcal{D}^{\gamma', 1/2 - \kappa}$  of the modelled distributions  $U_\varepsilon, U$  introduced in (2.4.13) and (2.4.14), the theorem applies for any  $h \in C_b^\beta(\mathbb{R})$  where  $\beta$  is the smallest integer  $\beta \geq ((2\gamma'/1 - 2\kappa) \vee 1) + 1$ . Since this operation depends only on the algebraic structure, we have also the same result on  $U_\varepsilon$  interpreted as a modelled distribution with respect to the canonical model  $(\Pi^\varepsilon, \Gamma^\varepsilon)$ .

### Space derivative

Thanks to its definition, the regularity structure  $\mathcal{T}$  allows us to define easily a linear map  $D_x: \mathcal{U} \rightarrow \mathcal{T}$  which behaves like a space derivative on abstract symbols. Indeed it is sufficient to characterise  $D_x$  as the unique linear map satisfying

$$\begin{aligned} D_x \mathbf{1} = 0, \quad D_x X_1 = 0, \quad D_x X_2 = \mathbf{1} \quad D_x \mathcal{I}(\Xi) = \mathcal{I}_1(\Xi), \\ D_x(\tau\sigma) = (D_x \tau)\sigma + (D_x \sigma)\tau. \end{aligned} \quad (2.4.18)$$

for any couple  $\tau, \sigma$  such that  $\sigma\tau \in U$ . Thus by composition we can define for any couple  $((\Pi, \Gamma), V) \in \mathcal{M} \times \mathcal{D}_{\mathcal{U}}^{\gamma, \eta}$  the function  $D_x V: \mathbb{R}^2 \rightarrow \mathcal{T}$ . This abstract operation which is defined at the level of  $\mathcal{U}$  can pass directly at the level of the reconstruction, thanks to the explicit structure of the models we are considering.

**Proposition 2.4.15.** *For any model  $(\Pi, \Gamma)$  of the form  $\mathcal{L}(\Pi)$  for some admissible map  $\Pi$ , the operator  $D_x$  is an abstract gradient which is compatible with  $(\Pi, \Gamma)$ , as explained in the definitions [Hai14, Def. 5.25, Def. 5.26]. Moreover for any  $V \in \mathcal{D}_U^{\gamma, \eta}$  such that  $\gamma > 1$ ,  $0 \leq \eta < \gamma$  the application  $V \rightarrow D_x V$  is an application  $D_x: \mathcal{D}_U^{\gamma, \eta}(\Pi) \rightarrow \mathcal{D}^{\gamma-1, \eta-1}(\Pi)$  depending continuously on the topology of  $\mathcal{M} \times \mathcal{D}_U^{\gamma, \eta}$  such that*

$$\mathcal{R}(D_x V) = \partial_x(\mathcal{R}V), \quad (2.4.19)$$

where the equality is interpreted in the sense of distributions.

*Proof.* By construction of the application  $D_x$  and using the multiplicative property of  $\Gamma_h$  (see the Remark 2.3.5) it is straightforward to prove recursively for any  $\beta \in \mathcal{A}$  and all  $h \in \mathbb{R}^3$  the following identities

$$D_x(\mathcal{Q}_\beta \mathcal{U}) \subset \mathcal{Q}_{\beta-1} \mathcal{U}, \quad D_x \Gamma_h = \Gamma_h D_x. \quad (2.4.20)$$

Hence  $D_x$  is an abstract gradient operator following [Hai14, Def. 5.25]. Let us fix a model  $(\Pi, \Gamma)$  of the form  $\mathcal{L}(\Pi)$  for some admissible map  $\Pi$ , then the conditions (2.3.17) (2.3.18) defining an admissible map imply that  $\Pi: \mathcal{U} \rightarrow \mathcal{S}'(\mathbb{R}^2)$  is well defined and for any  $u \in U$

$$\Pi D_x u = \partial_x \Pi u, \quad (2.4.21)$$

where the derivative  $\partial_x$  is interpreted in the sense of distributions. Summing up the properties (2.4.20) (2.4.21) and recalling the definition of  $\Pi_z$  in (2.3.16), for any  $z \in \mathbb{R}^2$  we obtain for any  $u \in U$

$$\Pi_z D_x u = \Pi \Gamma_{f(z)} D_x u = \Pi D_x \Gamma_{f(z)} u = (\partial_x \Pi) \Gamma_{f(z)} = \partial_x (\Pi \Gamma_{f(z)} u) = \partial_x \Pi_z,$$

where the equality  $(\partial_x \Pi) \Gamma_{f(z)} = \partial_x (\Pi \Gamma_{f(z)})$  holds because the operator  $\Gamma_h$  acts as a translation. Therefore  $D_x$  is compatible with  $(\Pi, \Gamma)$ , as explained [Hai14, Def. 5.26]. The remaining part of the statement follows directly from [Hai14, Prop 6.15] and the linearity of  $D_x$ .  $\square$

Applying the Proposition 2.4.19 to  $U_\varepsilon$  and  $U$ , we can write  $\partial_x u_\varepsilon$  and  $\partial_x u$  as the reconstruction of some modelled distributions.

**Corollary 2.4.16.** *For any  $\gamma' > 1$  let  $U_\varepsilon, U$  be the projection on  $\mathcal{D}^{\gamma', 1/2-\kappa}$  of the modelled distributions introduced in (2.4.13) and (2.4.14) for any fixed realisation of  $\xi$ . Then the modelled distributions  $D_x U_\varepsilon$  and  $D_x U$  belong respectively to  $\mathcal{D}^{\gamma'-1, -1/2-\kappa}$  and for any  $\varepsilon > 0$  one has*

$$\mathcal{R}_\varepsilon(D_x U) = \hat{\mathcal{R}}_\varepsilon(D_x U) = \partial_x u_\varepsilon, \quad \hat{\mathcal{R}}(D_x U) = \partial_x u \quad (2.4.22)$$

where the second identity holds on  $\mathcal{C}^{-1/2-\kappa}((0, T) \times \mathbb{T})$ .

## Product

We conclude the list of operation on modelled distributions with the notion of product between modelled distribution. Even if  $\mathcal{T}$  is not an algebra with respect to the juxtaposition product  $m$  introduced in the section 2.3, we can still consider  $m$  as a well defined bilinear map on some subspaces of  $\mathcal{T}$  such as  $m: \mathcal{U} \times \mathcal{T} \rightarrow \mathcal{T}$  or  $m: (\mathcal{V}_{\mathcal{I}_1(\Xi)} \oplus \mathcal{U}) \times (\mathcal{V}_{\mathcal{I}_1(\Xi)} \oplus \mathcal{U}) \rightarrow \mathcal{T}$ . Therefore for any couple of modelled distribution  $V_1, V_2$  and  $\gamma > 0$  we define the function  $V_1 V_2: \mathbb{R}^2 \rightarrow \mathcal{T}$  as

$$V_1 V_2(z) := \mathcal{Q}_{<\gamma}(V_1(z)V_2(z)), \quad (2.4.23)$$

as long as the point-wise product on the right hand side of (2.4.23) is well defined. The behaviour of this operation is described in [Hai14, Proposition 6.12], which we recall here.

**Proposition 2.4.17.** *Let  $(\Pi, \Gamma) \in \mathcal{M}$  and  $V_1 \in \mathcal{D}^{\gamma_1, \eta_1}(\Pi)$ ,  $V_2 \in \mathcal{D}^{\gamma_2, \eta_2}(\Pi)$  be a couple of modelled distributions such that the point-wise product is well defined. If the parameters*

$$\gamma = (\gamma_1 + \alpha_2) \wedge (\gamma_2 + \alpha_1), \quad \eta = (\eta_1 + \eta_2) \wedge (\eta_1 + \alpha_2) \wedge (\eta_2 + \alpha_1), \quad (2.4.24)$$

where for  $i = 1, 2$   $\alpha_i = \alpha_{V_i}$  satisfy the conditions  $\gamma > 0$  and  $-2 < \eta < \gamma$ , then the function  $V_1 V_2$  is a well defined element of  $\mathcal{D}^{\gamma, \eta}$ . Moreover this operation is continuous with respect to the topology of  $\mathcal{M} \times \mathcal{D}_U^{\gamma, \eta}$ .

*Remark 2.4.18.* Differently to the other operations we defined before, where we related the reconstruction operator to some classical operations on distribution, the reconstruction  $\mathcal{R}(V_1 V_2)$  cannot be defined directly as an analytical operation between  $\mathcal{R}(V_1)$  and  $\mathcal{R}(V_2)$ , even if we interpret the product as the operation between Hölder spaces explained in the section 2.2 (See in the Proposition 2.4.19 an example of a well defined product of modelled distributions such that the respective reconstruction of the factors cannot be multiplied). However in case of the canonical model  $(\Pi^\varepsilon, \Gamma^\varepsilon)$  for any  $\varepsilon > 0$  we can apply the multiplicative property of  $\Pi_z$  on symbols and the explicit form of the reconstruction operator in (2.4.4) to obtain for any couple of  $V_1 V_2 \in \mathcal{D}^{\gamma, \eta}(\Pi^\varepsilon)$  the general identity

$$\mathcal{R}_\varepsilon(V_1 V_2) = \mathcal{R}_\varepsilon V_1 \mathcal{R}_\varepsilon V_2. \quad (2.4.25)$$

We conclude the section by applying directly the Proposition 2.4.17 to verify the existence of two specific modelled distribution related to  $U_\varepsilon$  and  $U$ .

**Proposition 2.4.19.** *Let  $U_\varepsilon, U$  be the projection on  $\mathcal{D}^{\gamma', 1/2-\kappa}$  of the modelled distributions introduced in (2.4.13) and (2.4.14) for any fixed realisation of  $\xi$  and  $\gamma' > 0$ . Choosing  $\gamma' = 3/2 + 2\kappa$  for any  $\varphi \in C_b^7(\mathbb{R})$  the modelled distributions  $\Phi'(U_\varepsilon)\Xi$ ,  $\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2$  and  $\Phi'(U)\Xi$ ,  $\Phi''(U)(D_x U)^2$  are respectively well defined element of  $\mathcal{D}^{\kappa, -1-2\kappa}(\hat{\Pi}^\varepsilon)$  for any fixed  $\varepsilon > 0$  and  $\mathcal{D}^{\kappa, -1-2\kappa}(\hat{\Pi})$ . Moreover as  $\varepsilon \rightarrow 0$  we have*

$$\|\Phi'(U_\varepsilon)\Xi, \Phi'(U)\Xi\|_{\kappa, -1-2\kappa} \xrightarrow{\mathbb{P}} 0, \quad \|\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2, \Phi''(U)(D_x U)^2\|_{\kappa, -1-2\kappa} \xrightarrow{\mathbb{P}} 0. \quad (2.4.26)$$

*Proof.* Using the Proposition 2.4.13 and the Remark 2.4.14 to  $\varphi'$  and  $\varphi''$ , the modelled distributions  $\Phi''(U_\varepsilon)$   $\Phi'(U_\varepsilon)$  are well defined if  $\varphi', \varphi'' \in C_b^\beta(\mathbb{R})$  where  $\beta$  is the smallest integer such that  $\beta \geq ((2\gamma'/1 - 2\kappa) \vee 1) + 1$ . Choosing  $\gamma' = 3/2 + 2\kappa$  we have  $((2\gamma'/1 - 2\kappa) \vee 1) + 1 > 4$  and  $\beta = 5$ . Thus by hypothesis on  $\varphi$  we can lift the functions  $\varphi', \varphi''$  to modelled distributions. By construction of  $\mathcal{T}$  the point-wise product of  $\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2$ ,  $\Phi'(U_\varepsilon)\Xi$ ,  $\Phi''(U)(D_x U)^2$  and  $\Phi'(U)\Xi$  are well defined. Thus we recover the thesis by application the Proposition 2.4.19. In case of  $\Phi''(U_\varepsilon)\Xi$  and  $\Phi''(U)\Xi$  supposing that  $\Xi \in \mathcal{D}^{\gamma_2, \eta_2}$  for some  $\gamma_2 > 0$  and  $\eta_2$  sufficiently big the parameters  $\gamma$  and  $\eta$  in (2.4.24) become

$$\gamma = \gamma' - 3/2 - \kappa = \kappa, \quad \eta = 1/2 - \kappa - \alpha_\Xi = -1 - 2\kappa. \quad (2.4.27)$$

On the other hand when we consider  $\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2$  and  $\Phi''(U)(D_x U)^2$  we are doing two products. Choosing  $\gamma' > 1$  we start again from the functions  $D_x U_\varepsilon, D_x U$  given from the Corollary 2.4.16. Even in this case the parameters  $\gamma$  and  $\eta$  given by (2.4.24) for the product  $(D_x U_\varepsilon)^2, (D_x U)^2$  are given by (2.4.27) because  $\alpha_{D_x U_\varepsilon} = \alpha_{D_x U} = -1/2 - \kappa$ . Multiplying then for  $\Phi''(U)$  and  $\Phi''(U_\varepsilon)$  the parameters  $\gamma$  and  $\eta$  of this last product become

$$\gamma = (\gamma' - 3/2 - \kappa) \wedge (\gamma' - 1 - 2\kappa) = \kappa, \quad \eta = (-1 - 2\kappa) \wedge (-1/2 - 3\kappa) \quad (2.4.28)$$

which becomes equal to (2.4.27) because  $\kappa$  was fixed sufficiently small. Thus the modelled distribution are well defined and the convergence property (2.4.26) is direct consequence of the Theorem 2.3.19 and the continuous nature of the operations involved.  $\square$

*Remark 2.4.20.* Following the proof of the Proposition (2.4.19), the choice of the parameter  $\gamma'$  and  $\varphi$  in the statement could be replaced by a generic value  $\gamma' > 3/2 + \kappa$  and a function  $\varphi$  with the right number of bounded derivatives. The value  $3/2 + 2\kappa$  was simply chosen in order to find the smallest subspace where the modelled distributions  $\Phi'(U_\varepsilon)\Xi, \Phi''(U_\varepsilon)(D_x U_\varepsilon)^2, \Phi'(U)\Xi, \Phi''(U)(D_x U)^2$  are well defined.

## 2.5 Itô formula

We combine the explicit knowledge of the sequence  $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$  with the operations on the modelled distributions defined in the section 2.4 to describe the random distribution  $(\partial_t - \partial_{xx})\varphi(u)$  and  $\varphi(u)$ , when  $u$  is the solution of (2.1.1) and  $\varphi$  is a sufficiently smooth function, as explained in the introduction. The resulting formulae will be called differential and integral Itô formula, in accordance to the formal definitions given in the equations (2.1.4) and (2.1.5).

### 2.5.1 Pathwise Itô formulae

The first type of identities we show are called pathwise differential Itô Formula and pathwise integral Itô Formula. The additional adjective in their denomination is chosen because these identities involve in their terms the reconstruction of some modelled distribution, an object which is defined pathwise.

**Theorem 2.5.1** (Pathwise differential Itô Formula). *Let  $u$  be the solution of (2.1.1) and  $\varphi \in C_b^7(\mathbb{R})$ . Then we have the identity*

$$(\partial_t - \partial_{xx})(\varphi(u)) = \hat{\mathcal{R}}(\Phi'(U)\Xi) - \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2), \quad (2.5.1)$$

where the equality holds a.s. as elements of  $\mathcal{C}^{-3/2-\kappa}((0, T) \times \mathbb{T})$ .

*Proof.* The identity (2.5.1) will be obtained by rearranging the equality (2.1.10) in terms of modelled distributions and sending  $\varepsilon \rightarrow 0$ . Recalling the Proposition 2.4.10 and 2.4.19 we write  $u_\varepsilon = \hat{\mathcal{R}}_\varepsilon U_\varepsilon$  where  $U_\varepsilon$  is the projection on  $\mathcal{D}^{3/2+2\kappa, 1/2-\kappa}(\hat{\Pi}^\varepsilon)$  of the modelled distributions introduced in (2.4.13). The hypothesis on  $\varphi$  and the definition of  $U_\varepsilon$  allow to lift  $\varphi'$  and  $\varphi''$  at the level of the modelled distributions and we can rewrite the identity (2.1.10) as

$$(\partial_t - \partial_{xx})\varphi(u_\varepsilon) = (\hat{\mathcal{R}}_\varepsilon \Phi'(U_\varepsilon))(\hat{\mathcal{R}}_\varepsilon \Xi) - (\hat{\mathcal{R}}_\varepsilon \Phi''(U_\varepsilon))(\hat{\mathcal{R}}_\varepsilon D_x U_\varepsilon)^2. \quad (2.5.2)$$

On the other hand the same Proposition 2.4.19 implies that the modelled distributions  $\Phi'(U_\varepsilon)\Xi$  and  $\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2$  belong to  $\mathcal{D}^{\kappa, -1-2\kappa}(\hat{\Pi}^\varepsilon)$ . Consequently the random fields  $\hat{\mathcal{R}}_\varepsilon(\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2)$  and  $\hat{\mathcal{R}}_\varepsilon(\Phi'(U_\varepsilon)\Xi)$  are well defined functions and they are given explicitly by

$$\begin{aligned} \hat{\mathcal{R}}_\varepsilon(\Phi'(U_\varepsilon)\Xi)(z) &= \Pi_z^\varepsilon(M_\varepsilon \Phi'(U_\varepsilon)\Xi(z))(z), \\ \hat{\mathcal{R}}_\varepsilon(\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2)(z) &= \Pi_z^\varepsilon(M_\varepsilon \Phi''(U_\varepsilon)(D_x U_\varepsilon)^2(z))(z), \end{aligned} \quad (2.5.3)$$

for any  $z \in (0, T) \times \mathbb{T}$  as a consequence of the equation (2.4.4) and the Proposition 2.3.18. From these equalities we deduce an explicit relation between the functions on the left hand side of (2.5.3) and the right hand side of (2.5.2). In order to lighten the notation we write down  $\Phi'(U_\varepsilon)\Xi$ , and  $\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2$  on the canonical basis of  $\mathcal{Q}_{<\kappa}\mathcal{T}$  and  $(D_x U)^2$  on the canonical bases  $\mathcal{Q}_{<1/2+2\kappa}\mathcal{T}$  without referring explicitly to  $z \in (0, T) \times \mathbb{R}$  the indicator  $\mathbf{1}_+$  and the periodic lifting obtaining

$$\begin{aligned} \Phi'(U_\varepsilon)\Xi &= \varphi'(u_\varepsilon)\mathring{\downarrow} + \varphi''(u_\varepsilon)\mathring{\downarrow} + \varphi''(u_\varepsilon)v_\varepsilon^{(0,1)}\mathring{\downarrow}_{(0,1)} + \frac{\varphi'''(u_\varepsilon)\mathring{\downarrow}}{2} \\ &\quad + \varphi'''(u_\varepsilon)v_\varepsilon^{(0,1)}\mathring{\downarrow}_{(0,1)} + \frac{\varphi^{(4)}(u_\varepsilon)\mathring{\downarrow}}{6}, \\ (D_x U_\varepsilon)^2 &= \mathring{\downarrow} + 2v_\varepsilon^{(0,1)}\mathring{\downarrow} + (v_\varepsilon^{(0,1)})^2\mathbf{1}, \\ \Phi''(U_\varepsilon)(D_x U_\varepsilon)^2 &= \varphi''(u_\varepsilon)(D_x U_\varepsilon)^2 + \varphi'''(u_\varepsilon)\mathring{\downarrow} + \varphi'''(u_\varepsilon)v_\varepsilon^{(0,1)}\mathring{\downarrow}_{(0,1)} \\ &\quad + 2\varphi'''(u_\varepsilon)v_\varepsilon^{(0,1)}\mathring{\downarrow} + \frac{\varphi^{(4)}(u_\varepsilon)\mathring{\downarrow}}{2}. \end{aligned}$$

Then we apply of the renormalisation map  $M_\varepsilon$

$$\begin{aligned} M_\varepsilon(\Phi'(U_\varepsilon)\Xi) &= \varphi'(u_\varepsilon)\mathring{\downarrow} + \varphi''(u_\varepsilon)\left(\mathring{\downarrow} - C_\varepsilon^1\mathbf{1}\right) + \varphi''(u_\varepsilon)v_\varepsilon^{(0,1)}\mathring{\downarrow}_{(0,1)} \\ &\quad + \frac{\varphi'''(u_\varepsilon)}{2}\left(\mathring{\downarrow} - 2C_\varepsilon^1\mathring{\downarrow}\right) + \varphi'''(u_\varepsilon)v_\varepsilon^{(0,1)}\left(\mathring{\downarrow}_{(0,1)} - C_\varepsilon^1\mathbf{X}^{(0,1)}\right) \\ &\quad + \frac{\varphi^{(4)}(u_\varepsilon)}{6}\left(\mathring{\downarrow} - 3C_\varepsilon^1\mathring{\downarrow}\right), \end{aligned}$$

$$\begin{aligned}
M_\varepsilon((D_x U_\varepsilon)^2) &= \mathring{\mathbb{V}} + 2v_\varepsilon^{(0,1)} \mathring{\downarrow} + (v_\varepsilon^{(0,1)})^2 \mathbf{1} - C_\varepsilon^2 \mathbf{1} = (D_x U_\varepsilon)^2 - C_\varepsilon^2 \mathbf{1}, \\
M_\varepsilon(\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2) &= \varphi''(u_\varepsilon) \left( (D_x U_\varepsilon)^2 - C_\varepsilon^2 \mathbf{1} \right) + 2\varphi'''(u_\varepsilon) v_\varepsilon^{(0,1)} \mathring{\mathbb{V}} \\
&\quad + \varphi'''(u_\varepsilon) \left( \mathring{\mathbb{V}} - C_\varepsilon^2 \mathring{\downarrow} \right) + \varphi'''(u_\varepsilon) v_\varepsilon^{(0,1)} \left( \mathring{\mathbb{V}}_{(0,1)} - C_\varepsilon^2 \mathbf{X}^{(0,1)} \right) \\
&\quad + \frac{\varphi^{(4)}(u_\varepsilon)}{2} \left( \mathring{\mathbb{V}} - C_\varepsilon^2 \mathring{\mathbb{V}} \right).
\end{aligned}$$

To conclude the calculation we apply the operator  $\Pi_z^\varepsilon \cdot (z)$  on both sides of the above equations. As a consequence of the notion of model we have  $\Pi_z^\varepsilon \tau(z) = 0$  for every  $\tau \in T$  of the form  $\sigma_1 \sigma_2$  with  $|\sigma_1| > 0$ . Hence one has the identities

$$\hat{\mathcal{R}}_\varepsilon(\Phi'(U_\varepsilon)\Xi) = \varphi'(u_\varepsilon)\xi_\varepsilon - \varphi''(u_\varepsilon)C_\varepsilon^1 = (\hat{\mathcal{R}}_\varepsilon\Phi'(U_\varepsilon))(\hat{\mathcal{R}}_\varepsilon\Xi) - \varphi''(u_\varepsilon)C_\varepsilon^1, \quad (2.5.4)$$

$$\begin{aligned}
\hat{\mathcal{R}}_\varepsilon(\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2) &= \varphi''(u_\varepsilon) \left( \Pi_z^\varepsilon(D_x U_\varepsilon)^2(z) - C_\varepsilon^2 \right) \\
&= \hat{\mathcal{R}}_\varepsilon\Phi''(U_\varepsilon) \left( \Pi_z^\varepsilon(D_x U_\varepsilon)^2(z) - C_\varepsilon^2 \right).
\end{aligned} \quad (2.5.5)$$

Writing  $\Pi_z^\varepsilon(D_x U_\varepsilon)^2(z) = \mathcal{R}_\varepsilon((D_x U_\varepsilon)^2)$ , the multiplicative property of  $\mathcal{R}_\varepsilon$  in (2.4.25) and the identity (2.4.22) imply that the equality (2.5.5) becomes

$$\hat{\mathcal{R}}_\varepsilon(\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2) = (\hat{\mathcal{R}}_\varepsilon\Phi''(U_\varepsilon))(\hat{\mathcal{R}}_\varepsilon D_x U_\varepsilon)^2 - \varphi''(u_\varepsilon)C_\varepsilon^2. \quad (2.5.6)$$

Resuming up the equations (2.5.4) and (2.5.6), we obtain the final rearrangement

$$(\partial_t - \partial_{xx})\varphi(u_\varepsilon) = \hat{\mathcal{R}}_\varepsilon(\Phi'(U_\varepsilon)\Xi) - \hat{\mathcal{R}}_\varepsilon(\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2) + \varphi''(u_\varepsilon) (C_\varepsilon^1 - C_\varepsilon^2). \quad (2.5.7)$$

Let us now send  $\varepsilon \rightarrow 0^+$ . the left hand side of (2.5.7) converges in probability to  $(\partial_t - \partial_{xx})\varphi(u)$  thanks to the Proposition 2.4.12 and the fact that the derivative is a continuous operation between Hölder spaces. On the other hand, the local Lipschitz property of the reconstruction operator  $\mathcal{R}$  in (2.4.3) and the convergence (2.4.26) imply

$$\hat{\mathcal{R}}_\varepsilon(\Phi'(U_\varepsilon)\Xi) \xrightarrow{\mathbb{P}} \hat{\mathcal{R}}(\Phi'(U)\Xi), \quad \hat{\mathcal{R}}_\varepsilon(\Phi''(U_\varepsilon)(D_x U_\varepsilon)^2) \xrightarrow{\mathbb{P}} \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2), \quad (2.5.8)$$

with respect to the topology of  $\mathcal{C}^{-3/2-\kappa}((0, T) \times \mathbb{T})$ . Thus the theorem holds as long as the deterministic sequence  $C_\varepsilon^1 - C_\varepsilon^2 \rightarrow 0$ , which is the main consequence of the Lemma 2.6.2.  $\square$

*Remark 2.5.2.* Looking at the identities (2.5.4) and (2.5.6) separately and the convergence result (2.5.8), we obtain the existence of two sequences of random variables  $X_\varepsilon^1, X_\varepsilon^2 \in \mathcal{C}^{-3/2-\kappa}((0, T) \times \mathbb{T})$  converging in probability such that

$$\varphi'(u_\varepsilon)\xi_\varepsilon = X_\varepsilon^1 + \varphi''(u_\varepsilon)C_\varepsilon^1, \quad \varphi''(u_\varepsilon)(\partial_x u_\varepsilon)^2 = X_\varepsilon^2 + \varphi''(u_\varepsilon)C_\varepsilon^2.$$

Since we know from the Lemma 2.6.2 that the deterministic sequences  $C_\varepsilon^1$  and  $C_\varepsilon^2$  are both diverging, we obtain easily

$$\|\varphi'(u_\varepsilon)\xi_\varepsilon\|_{\mathcal{C}^{-3/2-\kappa}((0, T) \times \mathbb{T})} \xrightarrow{\mathbb{P}} +\infty, \quad \|\varphi''(u_\varepsilon)(\partial_x u_\varepsilon)^2\|_{\mathcal{C}^{-3/2-\kappa}((0, T) \times \mathbb{T})} \xrightarrow{\mathbb{P}} +\infty.$$

Thus we can justify rigorously the calculations done in the introduction.



From the formula (2.5.1) we can identify  $\varphi(u)$  with the solution of the following equation

$$\begin{cases} \partial_t v - \partial_{xx} v = \hat{\mathcal{R}}(\Phi'(U)\Xi) - \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2), \\ v(0, x) = \varphi(0) \\ v(t, 0) = v(t, 1) \\ \partial_x v(t, 0) = \partial_x v(t, 1) \end{cases}$$

Using the general results in the section 2.2 we obtain immediately.

**Corollary 2.5.3** (Pathwise integral Itô Formula). *For any  $\varphi \in C_b^7(\mathbb{R})$  and  $(t, x) \in [0, T] \times \mathbb{T}$  we have*

$$\varphi(u(t, x)) = \varphi(0) + (P * \mathbf{1}_{[0,t]} \hat{\mathcal{R}}(\Phi'(U)\Xi))(t, x) - (P * \mathbf{1}_{[0,t]} \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2))(t, x). \quad (2.5.9)$$

## 2.5.2 Identification of the differential formula

Thanks to the explicit Gaussian structure involving the definition of  $u$  in (2.1.2), in order to obtain the Theorem 2.1.1, we can identify the terms  $\hat{\mathcal{R}}(\Phi'(U)\Xi)$  and  $\hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)$  appearing in the formula (2.5.1) with some explicit classical operations of stochastic calculus (the so called identification theorems of the introduction). In what follows we will denote by  $(\mathcal{F}_t)_{t \in \mathbb{R}}$  the natural filtration of  $\xi$ , that is  $\mathcal{F}_t := \sigma(\{\xi(\psi) : \psi|_{(t, +\infty) \times \mathbb{T}} = 0; \psi \in L^2(\mathbb{R} \times \mathbb{T})\})$ . In case of  $\hat{\mathcal{R}}(\Phi'(U)\Xi)$  this identification is done by means of a general result contained in [HP15].

**Proposition 2.5.4.** *Let  $(\hat{\Pi}, \hat{\Gamma})$  be the BPHZ model and  $\varphi \in C_b^7(\mathbb{R})$ . Then for any smooth function  $\psi : \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$  with  $\text{supp}(\psi) \subset (0, +\infty) \times \mathbb{T}$ , one has for any  $t \in (0, T]$*

$$(\mathbf{1}_{[0,t]} \hat{\mathcal{R}}(\Phi'(U)\Xi))(\psi) = \int_0^t \int_{\mathbb{T}} \varphi'(u(s, y)) \psi(s, y) dW_{s,y}. \quad (2.5.10)$$

*Proof.* Thanks to the inclusion of  $\mathcal{V}_\Xi$  and  $\mathcal{U}$  into the regularity structure  $\mathcal{T}_{HP}$  and the identification of the BPHZ model  $(\hat{\Pi}, \hat{\Gamma})$  with the *Itô model*, both defined in [HP15] (see the Remark 2.3.3 and 2.3.20), the identity (2.5.10) is a consequence of [HP15, Theorem 6.2] applied to the modelled distribution  $\Phi'(U) \in \mathcal{D}_U^{\gamma', \eta'}$ , with  $\gamma' = 3/2 + 2\kappa$ ,  $\eta' = 1/2 - \kappa$  given the Proposition 2.4.19. Let us check that  $\Phi'(U)$  satisfy the hypotheses of this theorem. For any  $z \in [0, T] \times \mathbb{R}$ ,  $z = (t, x)$  applying the definition (2.4.17) to the explicit form of  $U$  in (2.4.14) we have

$$\begin{aligned} \Phi'(U)(z) &= \varphi'(\tilde{u}(z)) \mathbf{1} + \varphi''(\tilde{u}(z)) \mathbf{1}_+(z) \dot{\downarrow} + \varphi''(\tilde{u}(z)) v^{(0,1)}(z) \mathbf{X}^{(0,1)} \\ &\quad + \frac{\varphi'''(\tilde{u}(z))}{2} v^{(0,1)}(z) \mathbf{1}_+(z) \dot{\downarrow}_{(0,1)} + \frac{\varphi'''(\tilde{u}(z))}{2} \mathbf{1}_+(z) \dot{\downarrow} \check{\downarrow} \\ &\quad + \frac{\varphi^{(4)}(\tilde{u}(z))}{6} \mathbf{1}_+(z) \dot{\downarrow} \check{\downarrow} \check{\downarrow}. \end{aligned}$$

Since  $\{u(t, x)\}_{(t,x)}$  and  $\{v^{(0,1)}(t, x)\}_{(t,x)}$  are adapted to the filtration  $\mathcal{F}_t$ , so is the process  $\{\Phi'(U)(t, x)\}_{(t,x)}$ . Moreover  $\mathbb{E}|\Phi'(U)|_{\gamma', \eta'}^p < +\infty$  for any  $p > 2$  thanks to the local Lipschitz property of  $\Phi'$ , the lifting of  $\varphi'$ , and the bounds

$$\mathbb{E} \sup_{z \in [0, T] \times \mathbb{T}} |u(z)|^p < +\infty, \quad \mathbb{E} \sup_{z \in [0, T] \times \mathbb{T}} |v^{(0,1)}(z)|^p < +\infty, \quad (2.5.11)$$

that are true because  $u$  and  $v^{(0,1)}$  are centred Gaussian processes whose variance is uniformly bounded on  $[0, T] \times \mathbb{T}$ .  $\square$

We pass to the identification of  $\hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)$ . In this case no general result can be applied and following the same procedure of [Zam06], we can identify this random distribution by means of a different approximation of the process  $u$  using the heat semigroup on  $u$ . For any  $\varepsilon > 0$  we define the process

$$u_t^\varepsilon(x) := \int_{\mathbb{T}} P_\varepsilon(x - y)u(t, y)dy. \quad (2.5.12)$$

**Lemma 2.5.5.** *For any  $\varepsilon > 0$  the process  $u^\varepsilon$  satisfies the following properties:*

- for any  $t > 0$  the process  $\{u_t^\varepsilon(x)\}_{x \in \mathbb{T}}$  has a.s. smooth trajectories, satisfying for any integer  $m \geq 0$  the a.s. identity

$$\partial_x^m u_t^\varepsilon(x) = \int_0^t \int_{\mathbb{T}} \partial_x^m P_{\varepsilon+t-s}(x - y)dW_{s,y}. \quad (2.5.13)$$

- for any  $x \in \mathbb{T}$  the process  $\{u_t^\varepsilon(x)\}_{t \in [0, T]}$  is the strong solution of the equation

$$\begin{cases} du_t^\varepsilon(x) = \partial_{xx}(u_t^\varepsilon)(x)dt + dW_t^\varepsilon(x) \\ u_0^\varepsilon(x) = 0, \end{cases} \quad (2.5.14)$$

where  $W_t^\varepsilon(x)$  is the  $(\mathcal{F}_t)$ -martingale

$$W_t^\varepsilon(x) = \int_0^t \int_{\mathbb{T}} P_\varepsilon(x - y)dW_{s,y}, \quad \langle W_t^\varepsilon(x) \rangle_t = t \int_{\mathbb{T}} P_\varepsilon(x - y)^2 dy = t \|P_\varepsilon(\cdot)\|_{L^2(\mathbb{T})}^2.$$

- By sending  $\varepsilon \rightarrow 0$  one has

$$\sup_{(t,x) \in [0, T] \times \mathbb{T}} |u^\varepsilon(t, x) - u(t, x)| \rightarrow 0 \quad a.s. \quad (2.5.15)$$

*Proof.* We start by considering the trajectories of  $x \rightarrow u_t^\varepsilon(x)$  for any fixed  $t > 0$ . Since  $u$  is a.s. a continuous function, the regularisation property of the heat kernel  $P$  implies the desired property on its trajectories. Moreover for any integer  $m \geq 0$  we can pass the derivative under the Lebesgue integral to obtain

$$\partial_x^m u_t^\varepsilon(x) = \int_{\mathbb{T}} \partial_x^m P_\varepsilon(x - y)u(t, y)dy \quad a.s.$$

Using the straightforward bound

$$\int_{\mathbb{T}} \int_0^t \int_{\mathbb{T}} (\partial_x^m P_\varepsilon(x-y) P_{t-s}(y-v))^2 ds dv dy < +\infty,$$

we can obtain the formula (2.5.13) by writing the stochastic integral in (2.5.13) as a Wiener integral and applying the stochastic Fubini theorem for Wiener integral, as explained in [PT11, Thm. 5.13.1]. For any fixed  $x \in \mathbb{T}$  we study the process  $t \rightarrow u_t^\varepsilon(x)$ . By definition of mild solution of the equation (2.1.1)  $u$  satisfies the equality (2.1.3) for any smooth function  $l: \mathbb{T} \rightarrow \mathbb{R}$ , thus the identity (2.5.14) follows by simply setting  $l(y) = P_\varepsilon(x-y)$  in (2.1.3). Finally for any  $(t, x) \in [0, T] \times \mathbb{T}$  the a.s. Hölder continuity of  $u$  in the space and time implies the convergence (2.5.15), using the classical property of the heat semigroup on continuous functions.  $\square$

**Theorem 2.5.6.** *Let  $(\hat{\Pi}, \hat{\Gamma})$  be the BPHZ model and  $\varphi \in C_b^7(\mathbb{R})$ . Then for any smooth function  $\psi: \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$  with  $\text{supp}(\psi) \subset (0, +\infty) \times \mathbb{T}$ , one has for any  $t \in (0, T]$*

$$\mathbf{1}_{[0,t]} \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)(\psi) = -\frac{1}{2} \int_0^t \int_{\mathbb{T}} \psi(s, y) \varphi''(u(s, y)) C(s) dy ds \quad (2.5.16)$$

$$+ \int_{[0,t]^2 \times \mathbb{T}^2} \left[ \int_{s_2 \vee s_1}^t \int_{\mathbb{T}} \psi(s, y) \varphi''(u(s, y)) \partial_x P_{s-s_1}(y - \mathbf{y}_1) \partial_x P_{s-s_2}(y - \mathbf{y}_2) dy ds \right] dW_{s,y}^2,$$

where  $C: (0, T) \rightarrow \mathbb{R}$  is the deterministic integrable function  $C(s) := \|P_s(\cdot)\|_{L^2(\mathbb{T})}^2$ .

*Proof.* We prove firstly the result when  $\psi = h \otimes l$  where  $h: [0, t] \rightarrow \mathbb{R}$  is a compactly supported smooth function and  $l: \mathbb{T} \rightarrow \mathbb{R}$ .  $\psi$  is compactly supported up to time  $t$ . Therefore we can forget the operator  $\mathbf{1}_{[0,t]}$  on the right hand side of (2.5.16) and we can apply the Theorem 2.5.1 and the Proposition 2.5.4 obtaining

$$\begin{aligned} \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)(\psi) &= \left( -\partial_t(\varphi(u)) + \partial_x^2(\varphi(u)) + \hat{\mathcal{R}}(\Phi'(U)\Xi) \right) (\psi) = \\ &= \int_0^t \int_{\mathbb{T}} \left( \varphi(u(s, y)) h'(s) l(y) + \varphi(u(s, y)) h(s) l''(y) \right) ds dy \\ &+ \int_0^t \int_{\mathbb{T}} \varphi'(u(s, y)) h(s) l(y) dW_{s,y}. \end{aligned} \quad (2.5.17)$$

Let us recover the right hand side of (2.5.17) via a different approximation. Using the process  $u^\varepsilon$  defined in (2.5.12), we can apply the Itô formula to the semimartingale  $h(s)\varphi(u_s^\varepsilon(x))$  and we obtain

$$\begin{aligned} h(t)\varphi(u_t^\varepsilon(y)) - h(0)\varphi(u_0^\varepsilon(y)) &= \int_0^t h'(s)\varphi(u_s^\varepsilon(y)) ds + \int_0^t h(s)\partial_{xx}(u_s^\varepsilon)(y)\varphi'(u_s^\varepsilon(y)) ds \\ &+ \int_0^t h(s)\varphi'(u_s^\varepsilon(y)) dW_s^\varepsilon(y) + \frac{1}{2} \|P_\varepsilon(\cdot)\|_{L^2(\mathbb{T})}^2 \int_0^t h(s)\varphi''(u_s^\varepsilon(y)) ds. \end{aligned} \quad (2.5.18)$$

The left hand side of (2.5.18) is a.s. equal to zero by hypothesis on  $h$  and we can still apply the formula (2.1.9) with  $u^\varepsilon$  instead of  $u_\varepsilon$ . Hence we can rewrite the equation (2.5.18) as

$$\begin{aligned} \int_0^t \left( \varphi(u_s^\varepsilon(y))h'(s) + \partial_{xx}(\varphi(u_s^\varepsilon(y)))h(s) \right) ds + \int_0^t h(s)\varphi'(u_s^\varepsilon(y))dW_s^\varepsilon(y) \\ = \int_0^t \left[ (\partial_x u_s^\varepsilon)^2(y) - \frac{C_\varepsilon(y)}{2} \right] h(s)\varphi''(u_s^\varepsilon(y))ds. \end{aligned} \quad (2.5.19)$$

By multiplying both sides of (2.5.19) with  $l$  and integrating by part over  $\mathbb{T}$  to transfer the second derivative on  $l$ , the equation (2.5.19) becomes

$$\begin{aligned} \int_{\mathbb{T}} \left( \int_0^t \left( \varphi(u_s^\varepsilon(y))h'(s)l(y) + \varphi(u_s^\varepsilon(y))h(s)l''(y) \right) ds \right. \\ \left. + \int_0^t \varphi'(u_s^\varepsilon(y))h(s)l(y)dW_s^\varepsilon(y) \right) dy \\ = \int_0^t \int_{\mathbb{T}} l(y)h(s)\varphi''(u_s^\varepsilon(y)) \left( (\partial_x u_s^\varepsilon)^2(y) - \frac{\|P_\varepsilon(\cdot)\|_{L^2(\mathbb{T})}^2}{2} \right) ds dy. \end{aligned} \quad (2.5.20)$$

Writing the integral with respect to  $dW_s^\varepsilon(x)$  as a Walsh integral, we can apply the boundedness of  $\varphi'$  and  $\varphi$  to apply a stochastic Fubini's theorem on  $dW_{s,y}$  (see [CW75, Thm. 65])

$$\begin{aligned} \int_{\mathbb{T}} \int_0^t \varphi'(u_s^\varepsilon(y))h(s)l(y)dW_s^\varepsilon(y)dy = \int_{\mathbb{T}} \left( \int_0^t \int_{\mathbb{T}} P_\varepsilon(y-z)\varphi'(u_s^\varepsilon(z))h(s)l(y)dW_{s,z} \right) dy \\ = \int_0^t \int_{\mathbb{T}} \left( \int_{\mathbb{T}} P_\varepsilon(z-y)\varphi'(u_s^\varepsilon(y))h(s)l(y)dy \right) dW_{s,z}. \end{aligned} \quad (2.5.21)$$

Let us prove that the left hand side of (2.5.20) converges in  $L^2(\Omega)$  to the right hand side of (2.5.17). From the uniform convergence (2.5.15) of  $u^\varepsilon$ , it is straightforward to show as  $\varepsilon \rightarrow 0$

$$\begin{aligned} \int_0^t \int_{\mathbb{T}} \varphi(u_s^\varepsilon(y))h'(s)l(y)dyds \rightarrow \int_0^t \int_{\mathbb{T}} \varphi(u_s(y))h'(s)l(y)dyds \quad \text{a.s.} \\ \int_0^t \int_{\mathbb{T}} \varphi(u_s^\varepsilon(y))h(s)l''(y)dyds \rightarrow \int_0^t \int_{\mathbb{T}} \varphi(u_s(y))h(s)l''(y)dyds \quad \text{a.s.} \end{aligned}$$

and the convergence holds also in  $L^2(\Omega)$  because these random variables are also uniformly bounded. In case of the stochastic integral in (2.5.21), the same uniform convergence of  $u^\varepsilon$  in (2.5.21) implies that

$$\sup_{(s,z) \in [0,T] \times \mathbb{T}} \left| \int_{\mathbb{T}} P_\varepsilon(z-y)\varphi'(u_s^\varepsilon(y))h(s)l(y)dy - \varphi'(u_s(z))h(s)l(z) \right| \rightarrow 0 \quad \text{a.s.}$$

and bounding these quantity by some constant we obtain by dominated convergence

$$\mathbb{E} \left[ \int_0^t \int_{\mathbb{T}} \left( \int_{\mathbb{T}} P_\varepsilon(z-y) \varphi'(u_s^\varepsilon(y)) h(s) l(y) dy - \varphi'(u_s(z)) h(s) l(z) \right)^2 ds dz \right] \rightarrow 0.$$

Hence the proof is complete as long as the right hand side of (2.5.20) converges in  $L^2(\Omega)$  to the right hand side of (2.5.4). Using the shorthand notations  $p_s^\varepsilon(y) = \partial_x P_{\varepsilon+s}(y)$ ,  $P_s^\varepsilon(y) := P_{\varepsilon+s}(y)$ , the formula (2.5.13) on  $u^\varepsilon$  when  $m = 1$  becomes

$$\partial_x u_s^\varepsilon(y) = \int_0^s \int_{\mathbb{T}} p_{s-r}^\varepsilon(y-z) dW_{r,z}.$$

Thus writing it as a Wiener integral, we can express  $(\partial_x u_s^\varepsilon(y))^2$  using the Wiener chaos decomposition of a product (see [Nua95, Prop. 1.1.2]) obtaining

$$(\partial_x u_s^\varepsilon(y))^2 = I_2 \left( \mathbf{1}_{[0,s]^2 \times \mathbb{T}^2} p_{s-\cdot}^\varepsilon(y-\cdot) p_{s-\cdot}^\varepsilon(y-\cdot) \right) + \int_0^s \int_{\mathbb{T}} [p_{s-r}^\varepsilon(y-z)]^2 dz dr.$$

Hence, recalling the invariance by translations of  $\zeta \rightarrow \int_{\mathbb{T}} P(s, \zeta - y)^2 dy$  we write the right hand side of (2.5.20) as  $A_1^\varepsilon + A_2^\varepsilon$  where

$$\begin{aligned} A_1^\varepsilon &= \int_0^t \int_{\mathbb{T}} l(y) h(s) \varphi''(u_s^\varepsilon(y)) \left( \int_0^s \int_{\mathbb{T}} [p_{s-r}^\varepsilon(y-z)]^2 dr dz - \frac{1}{2} \int_{\mathbb{T}} P_s^\varepsilon(x-y)^2 dy \right) dy ds \\ A_2^\varepsilon &= \int_0^t \int_{\mathbb{T}} l(y) h(s) \varphi''(u_s^\varepsilon(y)) I_2 \left( \mathbf{1}_{[0,s]^2 \times \mathbb{T}^2} p_{s-\cdot}^\varepsilon(y-\cdot) p_{s-\cdot}^\varepsilon(y-\cdot) \right) dy ds. \end{aligned}$$

We treat both terms separately. In case of  $A_1^\varepsilon$  a simple integration by parts on  $\mathbb{T}$  and the smoothness of  $P$  outside the origin imply

$$\begin{aligned} \int_0^s \int_{\mathbb{T}} [p_{s-r}^\varepsilon(y-z)]^2 dr dz &= - \int_\varepsilon^{s+\varepsilon} \int_{\mathbb{T}} P_{s-r}(y-z) \partial_t P_{s-r}(y-z) dr dz = \\ &= - \int_\varepsilon^{s+\varepsilon} \partial_t \left( \int_{\mathbb{T}} \frac{P_r(y-z)^2}{2} dy \right) dr = \int_{\mathbb{T}} \frac{P_\varepsilon(y-z)^2}{2} dy - \int_{\mathbb{T}} \frac{P_{s+\varepsilon}(y-z)^2}{2} dy. \end{aligned}$$

Using again the invariance by translations of  $\zeta \rightarrow \int_{\mathbb{T}} P(s, \zeta - y)^2 dy$ , we can rewrite

$$\begin{aligned} A_1^\varepsilon &= -\frac{1}{2} \int_0^t \int_{\mathbb{T}} l(y) h(s) \varphi''(u_s^\varepsilon(y)) \left( \int_{\mathbb{T}} P_{s+\varepsilon}(y-z)^2 dz \right) ds dx \\ &= -\frac{1}{2} \int_0^t \int_{\mathbb{T}} l(y) h(s) \varphi''(u_s^\varepsilon(y)) \|P_{s+\varepsilon}(\cdot)\|_{L^2(\mathbb{T})}^2 ds dx, \end{aligned}$$

where the function  $C$  is given is defined in the statement. Therefore from the convergence (2.5.15) we obtain

$$A_1^\varepsilon \rightarrow -\frac{1}{2} \int_0^t \int_{\mathbb{T}} \psi(s, y) \varphi''(u(s, y)) C(s) dy ds \quad \text{a.s.}$$

and the convergence holds also in  $L^2(\Omega)$  because the sequence  $A_1^\varepsilon$  is uniformly bounded. We pass to the treatment of  $A_2^\varepsilon$ . In order to identify its limit, we interpret the double Wiener integral  $I_2$  as a multiple Skorohod integral of order 2. Then we want to rewrite the quantity

$$l(y)h(s)\varphi''(u_s^\varepsilon(y))I_2\left(\mathbf{1}_{[0,s]^2}p_{s-\cdot}^\varepsilon(y-\cdot)p_{s-\cdot}^\varepsilon(y-\cdot)\right)$$

using the product formula (2.2.8) for  $\delta^2$  to commute the deterministic integral in  $ds dy$  with the stochastic integration. Defining  $U_s^\varepsilon(y) := l(y)h(s)\varphi''(u_s^\varepsilon(y))$ , one has  $U_s^\varepsilon(y) \in \mathbb{D}^{2,2}$  because  $u^\varepsilon \in \mathbb{D}^{2,2}$  belongs to some fixed Wiener chaos and  $\varphi''$  has both two derivatives bounded. Applying the chain rule formula for the Malliavin derivative (see [Nua95, Proposition 1.2.3]) we have

$$\begin{aligned}\nabla_{\mathbf{s}_1, \mathbf{y}_1} U_s^\varepsilon(y) &= l(y)h(s)\varphi'''(u_s^\varepsilon(y))P_{s-\mathbf{s}_1}^\varepsilon(y-\mathbf{y}_1), \\ \nabla_{\mathbf{s}_1, \mathbf{y}_1, \mathbf{s}_2, \mathbf{y}_2}^2 U_s^\varepsilon &= l(y)h(s)\varphi^{(4)}(u_s^\varepsilon(y))P_{s-\mathbf{s}_1}^\varepsilon(y-\mathbf{y}_1)P_{s-\mathbf{s}_2}^\varepsilon(y-\mathbf{y}_2).\end{aligned}\tag{2.5.22}$$

For any  $\varepsilon > 0$  it is straightforward to check that the hypothesis of the product formula (2.2.8) are satisfied, therefore we can write

$$\begin{aligned}U_s^\varepsilon(y)I_2\left(\mathbf{1}_{[0,s]^2 \times \mathbb{T}^2}p_{s-\cdot}^\varepsilon(y-\cdot)p_{s-\cdot}^\varepsilon(y-\cdot)\right) &= \\ &= \int_{[0,s]^2} \int_{\mathbb{T}^2} U_s^\varepsilon(y)p_{s-\mathbf{s}_1}^\varepsilon(y-\mathbf{y}_1)p_{s-\mathbf{s}_2}^\varepsilon(y-\mathbf{y}_2)dW_{\mathbf{s}_1 \mathbf{y}_1 \mathbf{s}_2 \mathbf{y}_2}^2 \\ &+ 2 \int_{[0,s] \times \mathbb{T}} \left( \int_{[0,s] \times \mathbb{T}} \nabla_{\mathbf{s}_1, \mathbf{y}_1} U_s^\varepsilon(y)p_{s-\mathbf{s}_1}^\varepsilon(y-\mathbf{y}_1)p_{s-\mathbf{s}_2}^\varepsilon(y-\mathbf{y}_2)d\mathbf{s}_1 d\mathbf{y}_1 \right) dW_{\mathbf{s}_2 \mathbf{y}_2} \tag{2.5.23} \\ &+ \int_{[0,s]^2 \times \mathbb{T}^2} \nabla_{\mathbf{s}_1, \mathbf{y}_1, \mathbf{s}_2, \mathbf{y}_2}^2 U_s^\varepsilon(y)p_{s-\mathbf{s}_1}^\varepsilon(y-\mathbf{y}_1)p_{s-\mathbf{s}_2}^\varepsilon(y-\mathbf{y}_2)d\mathbf{s}_1 d\mathbf{y}_1 d\mathbf{s}_2 d\mathbf{y}_2.\end{aligned}$$

Looking at the deterministic deterministic integrals in the right hand side of (2.5.23), they are both zero as a consequence of the trivial identity

$$\int_{\mathbb{T}} P_{s-r}^\varepsilon(y-z)p_{s-r}^\varepsilon(y-z)dz = \int_{\mathbb{T}} \partial_x \frac{(P_{s-r}^\varepsilon(y-z))^2}{2} dz = 0. \tag{2.5.24}$$

Thus we can interchange the product of  $U_s^\varepsilon(y)$  with the multiple Skorohod integral of order 2. For any  $\varepsilon > 0$  the stochastic integrand inside  $dW_{\mathbf{s}_1 \mathbf{y}_1 \mathbf{s}_2 \mathbf{y}_2}^2$  is a smooth function in all its variables  $\mathbf{s}_1, \mathbf{y}_1, \mathbf{s}_2, \mathbf{y}_2, s, y$ , then it is square integrable when we integrate it on its referring domain. Therefore we can apply a Fubini type theorem for Skorohod integrals (see e.g. [NZ98]) to finally obtain

$$A_2^\varepsilon = \int_{[0,t]^2 \times \mathbb{T}^2} \left( \int_{\mathbf{s}_1 \vee \mathbf{s}_2}^t \int_{\mathbb{T}} h(s)l(y)\varphi''(u_s^\varepsilon(y))p_{s-\mathbf{s}_1}^\varepsilon(y-\mathbf{y}_1)p_{s-\mathbf{s}_2}^\varepsilon(y-\mathbf{y}_2)dy ds \right) dW_{\mathbf{s}, \mathbf{y}}^2.$$

Let us explain the convergence of  $A_2^\varepsilon$  to the multiple Skorohod integral of order two in the final formula (2.5.16). On one hand we proved that all the previous terms in the formula converge in  $L^2(\Omega)$ . Then if the sequence of functions

$$F^\varepsilon(\mathbf{s}, \mathbf{y}) := \int_{\mathbf{s}_1 \vee \mathbf{s}_2 + \varepsilon}^{t+\varepsilon} \int_{\mathbb{T}} h(s-\varepsilon)l(y)\varphi''(u_{s-\varepsilon}^\varepsilon(y))p_{s-\mathbf{s}_1}^0(y-\mathbf{y}_1)p_{s-\mathbf{s}_2}^0(y-\mathbf{y}_2)dy ds,$$

where  $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2)$  and  $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$ , converges in  $L^2(\Omega \times [0, t]^2 \times \mathbb{T}^2)$  to the function

$$F(\mathbf{s}, \mathbf{y}) := \int_{\mathbb{T}} \int_{\mathbf{s}_1 \vee \mathbf{s}_2}^t h(s) l(y) \varphi''(u_s(x)) p_{s-\mathbf{s}_1}^0(y - \mathbf{y}_1) p_{s-\mathbf{s}_2}^0(y - \mathbf{y}_2) dy ds,$$

the theorem will follow because the multiple Skorohod integral is a closed operator. From the a.s. convergence of  $u^\varepsilon$  in (2.5.15) it is straightforward to prove that  $F^\varepsilon$  converges to  $F$  a.s. and a.e. Then we conclude by dominated convergence by proving that  $\|F^\varepsilon\|_{L^2}^2$ , the square norm of  $F^\varepsilon$  in  $L^2([0, t]^2 \times \mathbb{T}^2)$  is uniformly bounded in  $\varepsilon$ . Using the symmetry of  $F^\varepsilon$  in the variables  $\mathbf{s}_1$  and  $\mathbf{s}_2$  we introduce the set  $\Delta_{2,t} = \{0 < \mathbf{s}_1 < \mathbf{s}_2 < t\}$  and writing the square of an integral as a double integral one has

$$\begin{aligned} \|F^\varepsilon\|_{L^2}^2 &= 2 \int_{\Delta_{2,t}} \int_{\mathbb{T}^2} F^\varepsilon(\mathbf{s}, \mathbf{y})^2 ds d\mathbf{y} = \\ &= 2 \int_\varepsilon^{t+\varepsilon} ds \int_\varepsilon^{t+\varepsilon} dr \int_{\mathbb{T}} dy \int_{\mathbb{T}} dx h(s-\varepsilon) h(r-\varepsilon) l(y) l(x) \varphi''(u_{s-\varepsilon}^\varepsilon(x)) \varphi''(u_{r-\varepsilon}^\varepsilon(y)) \\ &\times \int_0^{s \wedge r} \int_{\mathbb{T}} \int_0^{\mathbf{s}_2} \int_{\mathbb{T}} p_{s-\mathbf{s}_1}^0(x - \mathbf{y}_1) p_{s-\mathbf{s}_2}^0(x - \mathbf{y}_2) p_{r-\mathbf{s}_1}^0(y - \mathbf{y}_1) p_{r-\mathbf{s}_2}^0(y - \mathbf{y}_2) ds d\mathbf{y}, \end{aligned} \quad (2.5.25)$$

where we adopted the shorthand notation  $ds = ds_1 ds_2$ ,  $d\mathbf{y} = d\mathbf{y}_1 d\mathbf{y}_2$  and we applied the Fubini theorem. Integrating by parts with respect to  $\mathbf{y}_1$  and  $\mathbf{y}_2$  and applying the semigroup property of  $P$  we obtain

$$\begin{aligned} &\int_0^{s \wedge r} \int_{\mathbb{T}} \int_0^{\mathbf{s}_2} \int_{\mathbb{T}} p_{s-\mathbf{s}_1}^0(x - \mathbf{y}_1) p_{s-\mathbf{s}_2}^0(x - \mathbf{y}_2) p_{r-\mathbf{s}_1}^0(y - \mathbf{y}_1) p_{r-\mathbf{s}_2}^0(y - \mathbf{y}_2) ds d\mathbf{y} = \\ &= \int_0^{s \wedge r} \int_{\mathbb{T}} \int_0^{\mathbf{s}_2} \int_{\mathbb{T}} [\partial_t P_{s-\mathbf{s}_1}^0](x - \mathbf{y}_1) P_{r-\mathbf{s}_1}^0(x - \mathbf{y}_1) P_{s-\mathbf{s}_2}^0(y - \mathbf{y}_2) [\partial_t P_{r-\mathbf{s}_2}^0](y - \mathbf{y}_2) ds d\mathbf{y} \\ &= \int_0^{s \wedge r} \int_0^{\mathbf{s}_2} \partial_t P_{s+r-2\mathbf{s}_1}^0(y - x) \partial_t P_{s+r-2\mathbf{s}_2}^0(y - x) ds_1 ds_2 = \frac{1}{8} \left( [(P_\alpha^0(y - x)]_{\alpha=s+r}^{\alpha=|r-s|} \right)^2 \end{aligned} \quad (2.5.26)$$

Bounding  $\varphi l h$  with a deterministic constant and applying the rough estimate

$$(P_{|s-r|}^0(y - x) - P_{s+r}^0(y - x))^2 \leq (P_{|s-r|}^0(y - x))^2 + (P_{s+r}^0(y - x))^2,$$

there exists a constant  $M > 0$  such that for any  $\varepsilon > 0$  one has

$$\|F^\varepsilon\|_{L^2}^2 \leq M \int_0^T \int_{\mathbb{T}^2} (P_s^0(y - x))^2 ds d\mathbf{y} dx = M \int_0^T C(s) ds < +\infty.$$

Thereby obtaining the thesis. To conclude the result when  $\psi$  is a generic smooth function supported on  $(0, t) \times \mathbb{T}$ , we apply the formula (2.5.16) with a sequence of test functions  $h_N \otimes l_N: (0, t) \times \mathbb{T} \rightarrow \mathbb{R}$  converging to  $\psi$  as rapidly decreasing functions. This convergence is very strong and writing  $\hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)(h_N \otimes l_N)$  as the right hand side of (2.5.17) we can use the same argument as before to prove

$$\hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)(h_N \otimes l_N) \xrightarrow{L^2(\Omega)} \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)(\psi).$$

$$-\frac{1}{2} \int_0^t \int_{\mathbb{T}} (h_N(s) l_N(y)) \varphi''(u(s, y)) C(s) dy \xrightarrow{L^2(\Omega)} -\frac{1}{2} \int_0^t \int_{\mathbb{T}} \psi(s, y) \varphi''(u(s, y)) C(s) dy.$$

and then we can repeat the same argument above to prove that the double Skorohod integral converges to the respective quantity. When  $\psi$  is a generic test function defined on  $(0, +\infty) \times \mathbb{T}$  we repeat the same calculations with the sequence of tests function  $\varphi_N \psi$  where  $\varphi_N$  is introduced in (2.2.3) and it converges a.e. to the indicator function  $\mathbf{1}_{(0,t) \times \mathbb{T}}$ .  $\square$

*Remark 2.5.7.* The Proposition 2.5.4 and the Theorem 2.5.6 are formulated when the test function  $\psi: \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$  is supported on positive times in order to be coherent with the statement of the Theorem 2.1.1 and [HP15, Theorem 6.2]. However for any generic test function  $\psi$ , we can apply the identification theorems to the sequence  $\varphi_N \psi$  given in 2.2.3 and the explicit definition of the indicator operator  $\mathbf{1}_{[0,t]}$  to obtain that the same result holds without any restriction on the support of the test function.

*Remark 2.5.8.* The approximating procedure we used to prove this result is very different compared to the proof of [HP15, Theorem 6.2]. In that case the result is obtained by studying of the approximating sequence which is introduced to define the reconstruction map in the proof of the Theorem 2.4.3. That is  $\mathbf{1}_{[0,t]} \hat{\mathcal{R}}(\Phi'(U)\Xi)$  is a.s. the limit in the  $\mathcal{C}^{-3/2-\kappa}$  topology of the smooth random fields

$$\hat{\mathcal{R}}^n(\Phi'(U)\Xi)(z) := \mathbf{1}_{[0,t]}(t) \sum_{\bar{z} \in \Lambda^n([0,t])} \hat{\Pi}_{\bar{z}}(\Phi'(U)\Xi(\bar{z})) (\varphi_{\bar{z}}^n) \varphi_{\bar{z}}^n(z), \quad (2.5.27)$$

where  $\Lambda^n([0, T])$  denotes the dyadic grid on  $[0, T] \times \mathbb{T}$  of order  $n$  and the functions  $\varphi_{\bar{z}}^n(z)$  are obtained by rescaling of a specific compactly supported function  $\varphi: \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$ . It turns out that when we study the sequence (2.5.27) in the  $L^2(\Omega)$  topology the behaviour of this sequence is completely determined by knowing only the terms  $\hat{\Pi}_z(\tau\Xi)(\varphi_{\bar{z}}^n)$  for  $\tau \in U$ , thus we can apply the identity 2.3.44 and conclude. However considering the same approximations of  $\mathbf{1}_{[0,t]} \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)$ , we do not have the same simplification. In particular, the splitting of the heat kernel  $G$  as a sum  $K + R$  as explained in the Lemma 2.3.11 make all the calculations very indirect and it does not allow to use directly some nice properties of  $P$  such as the semigroup property. A general methodology to describe the stochastic properties of the reconstruction operator with respect to the BPHZ model is still missing.

*Remark 2.5.9.* From the formulae (2.5.10) and (2.5.16) we can easily write the periodic lifting of the reconstruction defined above. Indeed for any smooth function  $\psi: \mathbb{R}^2 \rightarrow \mathbb{R}$  with  $\text{supp}(\psi) \subset (0, +\infty) \times \mathbb{R}$  we have the identities

$$\mathbf{1}_{[0,t]} \widetilde{\hat{\mathcal{R}}}(\Phi'(U)\Xi)(\psi) = \int_0^t \int_{\mathbb{R}} \psi(s, y) \varphi'(\tilde{u}_s(y)) d\widetilde{W}_{s,y}, \quad (2.5.28)$$

$$\begin{aligned} \mathbf{1}_{[0,t]} \widetilde{\hat{\mathcal{R}}}(\Phi''(U)(D_x U)^2)(\psi) &= -\frac{1}{2} \int_0^t \int_{\mathbb{R}} \psi(s, y) \varphi''(\tilde{u}_s(y)) C(s) dy ds \\ &+ \int_{[0,t]^2 \times \mathbb{R}^2} \left[ \int_{\mathbf{s}_2 \vee \mathbf{s}_1}^t \int_{\mathbb{R}} \psi(s, y) \varphi''(\tilde{u}_s(y)) \partial_x G_{s-\mathbf{s}_1}(y - \mathbf{y}_1) \partial_x G_{s-\mathbf{s}_2}(y - \mathbf{y}_2) dy ds \right] d\widetilde{W}_{\mathbf{s},\mathbf{y}}^2. \end{aligned} \quad (2.5.29)$$



And the indicator operator on the right hand side tell us that that these identities hold for any smooth function  $\psi$  (see the Remark 2.5.7).

### 2.5.3 Identification of the integral formula

We pass to the identification of the terms involving the convolution with  $P$ . In principle this operation is deterministic and it should be obtained by applying the previous results to the deterministic test function  $\psi: \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$  given by  $\psi(s, y) = P_{t-s}(x - y)$  for some  $(t, x) \in [0, T] \times \mathbb{T}$ . However the function  $\psi$  is not smooth because  $\psi$  has a singularity at  $(t, x)$ . In order to skip this obstacle we recall an additional property of the function  $K: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ , introduced in the Lemma 2.3.11 and the Lemma 2.4.8.

**Lemma 2.5.10.** *There exists a sequence of smooth positive function  $K_n: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $n \geq 0$  satisfying  $\text{supp}(K_n) = \{z = (t, x) \in \mathbb{R}^2: \|z\| \leq 2^{-n}, t > 0\}$  such that for any  $z \in \mathbb{R}^2 \setminus \{0\}$*

$$K(z) = \sum_{n \geq 0} K_n(z). \quad (2.5.30)$$

Moreover for every distribution  $u \in \mathcal{C}^\alpha$  with  $-2 < \alpha < 0$  non integer one has for any  $z \in \mathbb{R}^2$

$$(K * u)(z) = \sum_{n \geq 0} (K_n * u)(z). \quad (2.5.31)$$

*Proof.* The Kernel  $K$  satisfies automatically the property (2.5.30) by construction, as expressed in [Hai14, Ass. 5.1]. Moreover for all test functions  $\psi$  and  $N \geq 0$  we have the identity

$$\left( \sum_{n=0}^N K_n * u \right) (\psi) = \sum_{n=0}^N (K_n * u) (\psi). \quad (2.5.32)$$

Following [Hai14, Lem. 5.19], the right hand side sequence of (2.5.32) is a Cauchy sequence with respect to the topology of  $\mathcal{C}^{\alpha+2}$ . Thus by uniqueness of the limit we obtain the equality

$$(K * u) = \sum_{n \geq 0} (K_n * u), \quad (2.5.33)$$

as elements of  $\mathcal{C}^{\alpha+2}$ . Since  $\alpha + 2 > 0$  (2.5.33) is an equality between functions, thereby obtaining the thesis.  $\square$

**Theorem 2.5.11.** *Let  $\varphi \in C_b^7(\mathbb{R})$ . Then for any  $(t, x) \in [0, T] \times \mathbb{T}$  one has*

$$P * (\mathbf{1}_{[0,t]} \hat{\mathcal{R}}(\Phi'(U)\Xi))(t, x) = \int_0^t \int_{\mathbb{T}} P_{t-s}(x - y) \varphi'(u_s(y)) dW_{s,y}, \quad (2.5.34)$$

$$\begin{aligned} P * (\mathbf{1}_{[0,t]} \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2))(t, x) &= -\frac{1}{2} \int_0^t \int_{\mathbb{R}} \psi(s, y) \varphi''(\tilde{u}_s(y)) C(s) dy ds \\ &+ \int_{[0,t]^2 \times \mathbb{R}^2} \left[ \int_{\mathbf{s}_2 \vee \mathbf{s}_1}^t \int_{\mathbb{R}} \psi(s, y) \varphi''(\tilde{u}_s(y)) \partial_x P_{s-\mathbf{s}_1}(y - \mathbf{y}_1) \partial_x P_{s-\mathbf{s}_2}(y - \mathbf{y}_2) dy ds \right] dW_{\mathbf{s},y}^2. \end{aligned} \quad (2.5.35)$$

*Proof.* We will prove equivalently the identities (2.5.34) and (2.5.35) on the periodic lifting. Using the Lemma 2.3.11 on the distribution  $v = \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)$ ,  $\hat{\mathcal{R}}(\Phi'(U)\Xi)$ , for any  $(t, x) \in [0, T] \times \mathbb{R}$  we have the general identity

$$(P * \widetilde{\mathbf{1}}_{[0,t]}v)(t, x) = (G * \widetilde{\mathbf{1}}_{[0,t]}v)(t, x) = (K * \widetilde{\mathbf{1}}_{[0,t]}v)(t, x) + (R * \widetilde{\mathbf{1}}_{[0,t]}v)(t, x). \quad (2.5.36)$$

Since the function  $R$  is smooth we can apply directly the formulae (2.5.28) (2.5.29). On the other hand the Lemma 2.5.10 implies

$$K * (\widetilde{\mathbf{1}}_{[0,t]}v)(t, x) = \sum_{n \geq 0} K_n * (\widetilde{\mathbf{1}}_{[0,t]}v)(t, x) = \lim_{N \rightarrow +\infty} \widetilde{\mathbf{1}}_{[0,t]}v(\eta_N),$$

where  $\eta_N: \mathbb{R}^2 \rightarrow \mathbb{R}$  is the sequence of compactly smooth functions

$$\eta_N(s, y) := \sum_{n=0}^N K_n(t-s, x-y).$$

Thus we will recover the result by applying the Proposition 2.5.4 and the Theorem 2.5.6 to the function  $\psi = \eta_N$  and studying the convergence of this sequence with respect to the topology of  $L^2(\Omega)$ . In case  $v = \hat{\mathcal{R}}(\Phi'(U)\Xi)$  one has trivially

$$\widetilde{\mathbf{1}}_{[0,t]}v(\eta_N) = \int_0^t \int_{\mathbb{R}} \eta_N(s, y) \varphi'(\tilde{u}(s, y)) d\widetilde{W}_{s,y}.$$

Since  $\varphi'$  is bounded, there exists a constant  $M > 0$  such that for any  $(s, y) \in [0, t] \times \mathbb{R}$  and  $N \geq 0$  one has

$$|\eta_N(s, y) \varphi'(\tilde{u}_s(y))| \leq M G_{t-s}(x-y).$$

The function  $(s, y) \rightarrow G_{t-s}(x-y)$  is  $L^2$  integrable on  $[0, t] \times \mathbb{R}$  and  $\eta_N(s, y)$  converges a.e. to  $K$ . By using the Itô isometry and the dominated convergence we can straightforwardly prove

$$\widetilde{\mathbf{1}}_{[0,t]}v(\eta_N) \xrightarrow{L^2(\Omega)} \int_0^t \int_{\mathbb{R}} K(t-s, x-y) \varphi'(\tilde{u}_s(y)) d\widetilde{W}_{s,y},$$

Thereby obtaining the identity (2.5.34) summing this term with respective containing the function  $R$ . Let us pass to the case of  $v = \hat{\mathcal{R}}(\Phi''(U)(D_x U)^2)$ . To shorten the notation we will adopt the convention  $g_{s-r}(x-y) := \partial_x G_{s-r}(x-y)$  and  $O_t = [0, t] \times \mathbb{R}$ . Looking again at the equation (2.5.29) we have  $\widetilde{\mathbf{1}}_{[0,t]}v(\eta_N) = A_N^1 + A_N^2$  where

$$A_N^1 = -\frac{1}{2} \int_{O_t} \eta_N(s, y) \varphi''(\tilde{u}_s(y)) C(s) dy ds,$$

$$A_N^2 = \int_{O_t \times O_t} \left[ \int_{\mathbf{s}_2 \vee \mathbf{s}_1}^t \int_{\mathbb{R}} \eta_N(s, y) \varphi''(\tilde{u}_s(y)) g_{s-\mathbf{s}_1}(y - \mathbf{y}_1) g_{s-\mathbf{s}_2}(y - \mathbf{y}_2) dy ds \right] d\widetilde{W}_{\mathbf{s},y}^2.$$

From the definition of  $\eta_N$  one has a.e. and a.s.

$$\eta_N(s, y) \varphi''(\tilde{u}(s, y)) C(s) \rightarrow K(t-s, x-y) \varphi''(\tilde{u}(s, y)) C(s).$$

Moreover there exists a constant  $M > 0$  such that for every  $N \geq 0$

$$|A_N^1| \leq M \int_{O_t} G_{t-s}(x-y)C(s)dy ds = M \int_0^t C(s) ds < \infty$$

(the last equality comes by integrating on  $\mathbb{R}$  the density function of a Gaussian random variable). Therefore by dominated convergence we obtain

$$A_N^1 \xrightarrow{L^2(\Omega)} -\frac{1}{2} \int_0^t \int_{\mathbb{R}} K(t-s, x-y)\varphi''(\tilde{u}_s(y))C(s)dy ds.$$

Let us pass to the convergence of the sequence  $A_N^2$ . Introducing the functions  $\{\Phi^N\}_{N \geq 0}$ ,  $\Phi_K$ ,  $\{F^N\}_{N \geq 0}$  and  $F_K$  defined by

$$\Phi^N(s, y, \mathbf{s}, \mathbf{y}) := \eta_N(s, y)\varphi''(\tilde{u}_s(y))g_{s-s_1}(y-\mathbf{y}_1)g_{s-s_2}(y-\mathbf{y}_2),$$

$$\Phi_K(s, y, \mathbf{s}, \mathbf{y}) := K(t-s, x-y)\varphi''(\tilde{u}_s(y))g_{s-s_1}(y-\mathbf{y}_1)g_{s-s_2}(y-\mathbf{y}_2),$$

$$F^N(\mathbf{s}, \mathbf{y}) := \int_{\mathbf{s}_2 \vee \mathbf{s}_1}^t \int_{\mathbb{R}} \Phi^N(s, y, \mathbf{s}, \mathbf{y}) ds dy, \quad F_K(\mathbf{s}, \mathbf{y}) := \int_{\mathbf{s}_2 \vee \mathbf{s}_1}^t \int_{\mathbb{R}} \Phi_K(s, y, \mathbf{s}, \mathbf{y}) ds dy,$$

(as usual  $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2)$  and  $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$  and  $\mathbf{s}_1 \vee \mathbf{s}_2 \leq t$ ), we will prove the last convergence

$$A_N^2 \xrightarrow{L^2(\Omega)} \int_{O_t \times O_t} F_K(\mathbf{s}, \mathbf{y}) d\widetilde{W}_{\mathbf{s}, \mathbf{y}}^2. \quad (2.5.37)$$

The multiple Skorohod integral is a continuous map from  $H = \mathbb{D}^{2,2}(L^2(O_t \times O_t))$  to  $L^2(\Omega)$  (see the definition of  $\mathbb{D}^{2,2}(V)$  and the inequality (2.2.9) in the section 2.2). Then the result (2.5.44) will follow by proving that  $F^N$  and  $F_K$  belong to  $H$  and  $F^N \rightarrow F_K$  in  $H$ . In order to prove these results we calculate the first and second Malliavin derivative of  $F^N$  and  $F_K$  thanks to chain rule formula of the Malliavin derivative (see [Nua95, Proposition 1.2.3]). In particular we have

$$\begin{aligned} \nabla_{\mathbf{t}_1 \mathbf{z}_1} F^N(\mathbf{s}, \mathbf{y}) &= \int_{\mathbf{s}_2 \vee \mathbf{s}_1 \vee \mathbf{t}_1}^t \int_{\mathbb{R}} \eta_N(s, y)\varphi^{(3)}(\tilde{u}_s(y))G_{s-\mathbf{t}_1}(y-\mathbf{z}_1) \\ &\quad \times g_{s-s_1}(y-\mathbf{y}_1)g_{s-s_2}(y-\mathbf{y}_2) dy ds, \end{aligned} \quad (2.5.38)$$

$$\begin{aligned} \nabla_{\mathbf{t}_1 \mathbf{z}_1 \mathbf{t}_2 \mathbf{z}_2}^2 F^N(\mathbf{s}, \mathbf{y}) &= \int_{\mathbf{s}_2 \vee \mathbf{s}_1 \vee \mathbf{t}_1 \vee \mathbf{t}_2}^t \int_{\mathbb{R}} \eta_N(s, y)\varphi^{(4)}(\tilde{u}_s(y))G_{s-\mathbf{t}_2}(y-\mathbf{z}_2)G_{s-\mathbf{t}_1}(y-\mathbf{z}_1) \\ &\quad \times g_{s-s_1}(y-\mathbf{y}_1)g_{s-s_2}(y-\mathbf{y}_2) dy ds. \end{aligned} \quad (2.5.39)$$

and similarly for  $F_K$  by replacing  $\eta_N$  with  $K(t-s, x-y)$ . Bounding uniformly  $\eta_N$  and  $K$  by  $G_{t-s}(x-y)$  we have trivially  $\|F^N\|_K^2 \leq \|F_G\|_K^2$  for every  $N \geq 0$  where

$$\Phi_G(s, y, \mathbf{s}, \mathbf{y}) := G_{t-s}(x-y)\varphi''(\tilde{u}_s(y))g_{s-s_1}(y-\mathbf{y}_1)g_{s-s_2}(y-\mathbf{y}_2),$$

$$F_G(\mathbf{s}, \mathbf{y}) := \int_{\mathbf{s}_2 \vee \mathbf{s}_1}^t \int_{\mathbb{R}} \Phi_G(s, y, \mathbf{s}, \mathbf{y}) ds dy,$$

(The Malliavin derivatives of  $F_G$  are given by (2.5.38) and (2.5.39) where  $\eta_N$  is replaced by  $G$ ) Then in order to show  $F^N$  and  $F_K$  are elements of  $H$ , it is sufficient to prove that the random variables

$$\alpha_1 := \int_{(O_t)^2} (F_G(\mathbf{s}, \mathbf{y}))^2 ds dy, \quad \alpha_2 := \int_{(O_t)^3} (\nabla_{\mathbf{t}_1 \mathbf{z}_1} F_G(\mathbf{s}, \mathbf{y}))^2 ds dy dt_1 dz_1,$$

$$\alpha_3 := \int_{(O_t)^4} (\nabla_{\mathbf{t}_1 \mathbf{z}_1 \mathbf{t}_2 \mathbf{z}_2}^2 F_G(\mathbf{s}, \mathbf{y}))^2 ds dy dt dz, \quad \mathbf{t} = (\mathbf{t}_1, \mathbf{t}_2), \quad \mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2),$$

are uniformly bounded. Let us analyse them separately. Looking at the expression of  $\alpha_1$ , it is possible to express the term  $g_{s-s_1}(y - \mathbf{y}_1)g_{s-s_2}(y - \mathbf{y}_2)$  in the definition of  $F^2$  in the same way as in the equations (2.5.26) (2.5.25) where the kernels  $p$  and  $P$  are replaced by  $g$  and  $G$ , thereby obtaining

$$\alpha_1 = \int_{(O_t)^2} G_{t-r}(x-z)G_{t-s}(x-y)\varphi''(\tilde{u}_s(y))\varphi''(\tilde{u}_r(z))$$

$$\times (G_{|s-r|}(y-z) - G_{s+r}(y-z))^2 dr ds dy dz.$$

By hypothesis on  $\varphi$  and bounding roughly the difference of a square there exists a constant  $M > 0$  such that for every  $N \geq 0$

$$\alpha_1 \leq M \int_{O_t \times O_t} G_{t-s}(x-y)G_{t-r}(x-z)(G_{|s-r|}(z-y))^2 + (G_{s+r}(z-y))^2 dr ds dy dz. \quad (2.5.40)$$

Let us show that the deterministic integral in the right hand side of (2.5.40) is finite. By definition of  $G$  one has

$$G_{|s-r|}(z-y)^2 + G_{s+r}(z-y)^2 = \frac{G_{|s-r|/2}(z-y)}{\sqrt{8\pi|s-r|}} + \frac{G_{(s+r)/2}(z-y)}{\sqrt{8\pi(s+r)}}. \quad (2.5.41)$$

Plugging this identity in the deterministic integral in the right hand side of (2.5.40) we can apply the semigroup property of  $G$  and using a rough estimate on the Heat kernel there exist some constants  $C, C' > 0$  such that

$$\int_{O_t \times O_t} G_{t-s}(x-y)G_{t-r}(x-z) \left( \frac{G_{|s-r|/2}(z-y)}{\sqrt{8\pi|s-r|}} + \frac{G_{(s+r)/2}(z-y)}{\sqrt{8\pi(s+r)}} \right) dr ds dy dz$$

$$= \int_0^t \int_{O_t} G_{t-s}(x-y) \left( \frac{G_{t-r+|s-r|/2}(x-y)}{\sqrt{8\pi|s-r|}} + \frac{G_{t+(s-r)/2}(x-y)}{\sqrt{8\pi(s+r)}} \right) ds dr dy$$

$$\leq C \int_0^t \int_{O_t} \frac{1}{\sqrt{t-s}} \left( \frac{G_{t-r+|s-r|/2}(x-y)}{\sqrt{8\pi|s-r|}} + \frac{G_{t+(s-r)/2}(x-y)}{\sqrt{8\pi(s+r)}} \right) ds dr dy$$

$$\leq C' \int_0^t \int_0^t \left( \frac{1}{\sqrt{t-s}} \frac{1}{\sqrt{|s-r|}} + \frac{1}{\sqrt{t-s}} \frac{1}{\sqrt{s+r}} \right) ds dr < +\infty.$$

(we are again integrating on  $\mathbb{R}$  the density function of a Gaussian random variable). Passing to  $\alpha_2$ , we rewrite  $(\nabla_{\mathbf{t}_1 \mathbf{z}_1} F(\mathbf{s}, \mathbf{y}))^2$  as a double integral and applying again the semigroup property of  $G$  we have

$$\begin{aligned} \alpha_2 &= 2 \int_{(\Delta_{2,t} \times \mathbb{R}^2) \times O_t} (\nabla_{\mathbf{t}_1 \mathbf{z}_1} F^N(\mathbf{s}, \mathbf{y}))^2 ds dy dt_1 dz_1 \\ &= 2 \int_{(O_t)^2} G_{t-s}(x-y) G_{t-r}(x-z) \varphi^{(3)}(\tilde{u}_s(y)) \varphi^{(3)}(\tilde{u}_r(z)) \Gamma_{s,r}^3(z, y) dr ds dy dz, \end{aligned}$$

where the function  $\Gamma_{s,r}^3(z, y)$  is defined through the identities

$$\begin{aligned} \Gamma_{s,r}^3(z, y) &:= \int_0^{s \wedge r} \int_0^{\mathbf{t}_1} \int_0^{\mathbf{s}_2} \Gamma_{s,r,s,\mathbf{t}_1}^3(z, y) ds dt_1 + \int_0^{s \wedge r} \int_0^{\mathbf{s}_2} \int_0^{\mathbf{s}_1} \Gamma_{s,r,s,\mathbf{t}_1}^3(z, y) dt_1 ds \\ &\quad + \int_0^{s \wedge r} \int_0^{\mathbf{s}_2} \int_0^{\mathbf{t}_1} \Gamma_{s,r,s,\mathbf{t}_1}^3(z, y) ds_1 dt_1 ds_2, \end{aligned}$$

$$\Gamma_{s,r,s,\mathbf{t}_1}^3(z, y) := G_{s+r-2\mathbf{t}_1}(y-z) g_{s+r-2\mathbf{s}_1}(y-z) g_{s+r-2\mathbf{s}_2}(y-z).$$

Let us consider the term  $\Gamma_{s,r}^3(z, y)$ . Using the elementary estimates

$$|g_t(y)| \leq \sup_{u \in \mathbb{R}} \left( \frac{u}{\sqrt{4\pi}} \exp^{-u^2} \right) \frac{1}{t}, \quad |G_t(y)| \leq \frac{1}{\sqrt{t}}, \quad (2.5.42)$$

we can bound each term of in the sum defining  $\Gamma_{s,r}^3(z, y)$  by some integrable functions depending only on  $r$  and  $s$ . For example in case of the first term in the sum defining  $\Gamma_{s,r}^3(z, y)$  there exists a constant  $C > 0$  such that

$$\begin{aligned} &\left| \int_0^{s \wedge r} \int_0^{\mathbf{t}_1} \int_0^{\mathbf{s}_2} \Gamma_{s,r,s,\mathbf{t}_1}^3(z, y) ds dt_1 \right| \leq \\ &\leq C \int_0^{s \wedge r} \left[ \int_0^{\mathbf{t}_1} \left[ \int_0^{\mathbf{s}_2} \frac{1}{s+r-2\mathbf{s}_1} ds_1 \right] \frac{1}{s+r-2\mathbf{s}_2} ds_2 \right] \frac{1}{\sqrt{s+r-2\mathbf{t}_1}} dt_1, \end{aligned}$$

and writing explicitly the integral on the right hand side there exists a constant  $C' > 0$  such that this integral is bounded by

$$C' \left[ \ln(s+r)^2 (\sqrt{s+r} - \sqrt{|s-r|}) + \ln(s+r) \int_{|s-r|}^{s+r} \frac{|\ln(y)|}{\sqrt{y}} dy + \int_{|s-r|}^{s+r} \frac{\ln(y)^2}{\sqrt{y}} dy \right].$$

Working exactly in the same way on the other integrals and bounding roughly the functions without singularity by some constant depending on the finite time parameter  $T$ , it is possible to show that there exists a constant  $D_T > 0$  depending on  $T$  such that

$$|\Gamma_{s,r}^3(z, y)| \leq D_T \left( 1 + \int_{|s-r|}^{s+r} \frac{\ln(y)^2}{\sqrt{y} \vee 1} dy + \int_{|s-r|}^{s+r} \frac{|\ln(y)|}{\sqrt{y} \vee 1} dy \right). \quad (2.5.43)$$

let us denote the right hand side of (2.5.43) by  $C_T(s, r)$ . This function is clearly integrable on  $[0, t]^2$ , therefore integrating on the remaining components and bounding the derivatives there exists a constant  $M > 0$  such that

$$\begin{aligned}\alpha_2 &\leq M \int_{(O_t)^2} G_{t-s}(x-y)G_{t-r}(x-z)C_T(s, r)drdsdydz \\ &= M \int_{[0, t]^2} C_T(s, r)drds < +\infty.\end{aligned}$$

Considering  $\alpha_3$ , we write  $(\nabla_{\mathbf{t}, \mathbf{z}}^2 F(\mathbf{s}, \mathbf{y}))^2$  using the same technique to express  $\alpha_2$  and we obtain

$$\begin{aligned}\alpha_3 &= 8 \int_{(\Delta_{2,t} \times \mathbb{R}^2)^2} (\nabla_{\mathbf{t}, \mathbf{z}}^2 F^N(\mathbf{s}, \mathbf{y}))^2 ds dy dt dz \\ &= 8 \int_{(O_t)^2} \eta_N(s, y)\eta_N(r, z)\varphi^{(4)}(\tilde{u}_s(y))\varphi^{(4)}(\tilde{u}_r(z))\Gamma_{s,r}^4(z, y)drdsdydz.\end{aligned}$$

(the factor 8 comes out because the function  $(\nabla_{\mathbf{t}, \mathbf{z}}^2 F^N(\mathbf{s}, \mathbf{y}))^2$  is symmetric under the change of coordinates  $\mathbf{s}_1 \rightarrow \mathbf{s}_2$ ,  $\mathbf{t}_1 \rightarrow \mathbf{t}_2$  and  $\mathbf{s} \rightarrow \mathbf{t}$ ). The function  $\Gamma_{s,r}^4(z, y)$  is defined through the new identities

$$\begin{aligned}\Gamma_{s,r}^4(z, y) &:= \int_0^{s \wedge r} \int_0^{\mathbf{t}_2} \int_0^{\mathbf{t}_1} \int_0^{\mathbf{s}_2} \Gamma_{s,r,\mathbf{s},\mathbf{t}}^4(z, y) ds dt \\ &\quad + \int_0^{s \wedge r} \int_0^{\mathbf{t}_2} \int_0^{\mathbf{s}_2} \int_0^{\mathbf{s}_1} \Gamma_{s,r,\mathbf{s},\mathbf{t}}^4(z, y) dt_1 ds dt_2 \\ &\quad + \int_0^{s \wedge r} \int_0^{\mathbf{t}_2} \int_0^{\mathbf{s}_2} \int_0^{\mathbf{t}_1} \Gamma_{s,r,\mathbf{s},\mathbf{t}}^4(z, y) ds_1 dt_1 ds_2 dt_2,\end{aligned}$$

$$\Gamma_{s,r,\mathbf{s},\mathbf{t}}^4(z, y) := G_{s+r-2\mathbf{t}_1}(y-z)G_{s+r-2\mathbf{t}_2}(y-z)g_{s+r-2\mathbf{s}_1}(y-z)g_{s+r-2\mathbf{s}_2}(y-z).$$

Recalling the elementary estimates in (2.5.42), we can similarly bound every single integral appearing in  $\Gamma_{s,r}^4(z, y)$  in the same way implying there exists an integrable function  $B_T(r, s)$  such that  $|\Gamma_{s,r}^4(z, y)| \leq B_T(s, t)$ . bounding  $\varphi^{(4)}$  we conclude there exists a constant  $M > 0$  such that

$$\begin{aligned}\alpha_3 &\leq M \int_{(O_t)^2} G_{t-s}(x-y)G_{t-r}(x-z)B_T(s, r)drdsdydz \\ &= M \int_{[0, t]^2} B_T(s, r)drds < +\infty.\end{aligned}$$

Thus we conclude that the sequence  $\alpha$  is uniformly bounded and  $F^N, F_K \in H$ . As a matter of fact the previous estimates have a stronger consequence because they imply that the functions  $\Phi^N(s, y, \mathbf{s}, \mathbf{y})$  and  $\Phi_K(s, y, \mathbf{s}, \mathbf{y})$  defined above are a.e. on  $s, y, \mathbf{s}, \mathbf{y}$  and

a.s. dominated by some integrable functions. Therefore looking at the quantity

$$\begin{aligned} \|F^N - F_K\|_H^2 &= \mathbb{E} \int_{(O_t)^2} (F^N(\mathbf{s}, \mathbf{y}) - F_K(\mathbf{s}, \mathbf{y}))^2 ds dy \\ &+ \mathbb{E} \int_{(O_t)^3} (\nabla_{\mathbf{t}_1 \mathbf{z}_1} F^N(\mathbf{s}, \mathbf{y}) - \nabla_{\mathbf{t}_1 \mathbf{z}_1} F_K(\mathbf{s}, \mathbf{y}))^2 ds dy dt_1 dz_1 \\ &+ \mathbb{E} \int_{(O_t)^4} (\nabla_{\mathbf{t}, \mathbf{z}}^2 F^N(\mathbf{s}, \mathbf{y}) - \nabla_{\mathbf{t}, \mathbf{z}}^2 F_K(\mathbf{s}, \mathbf{y}))^2 ds dy dt dz, \end{aligned} \quad (2.5.44)$$

since we have trivially the a.e. a.s. convergence of the functions

$$\begin{aligned} \Phi^N(s, y, \mathbf{s}, \mathbf{y}) &\rightarrow \Phi_K(s, y, \mathbf{s}, \mathbf{y}), \quad \nabla_{\mathbf{t}_1 \mathbf{z}_1} F^N(\mathbf{s}, \mathbf{y}) \rightarrow \nabla_{\mathbf{t}_1 \mathbf{z}_1} F_K(\mathbf{s}, \mathbf{y}), \\ \nabla_{\mathbf{t}, \mathbf{z}}^2 F^N(\mathbf{s}, \mathbf{y}) &\rightarrow \nabla_{\mathbf{t}, \mathbf{z}}^2 F_K(\mathbf{s}, \mathbf{y}), \end{aligned}$$

we obtain  $\|F^N - F_K\|_H^2 \rightarrow 0$  by dominated converge and the theorem is proved.  $\square$

*Proof of the Theorem 2.1.1.* For any  $\varphi \in C_b^7(\mathbb{R})$  the differential and the integral formula are obtained applying straightforwardly the previous results. Looking at their proofs we realise that the Skorohod and the Wiener integrals and their convolution with  $P$ , differently from the reconstructions, are well defined if the derivatives of  $\varphi$  are bounded up to the order 4. Thus for any fixed  $\varphi \in C_b^4(\mathbb{R})$  the formula holds using a classical density argument of the  $C^\infty$  functions with all bounded derivatives and repeating the same calculations.  $\square$

*Remark 2.5.12.* Using the integrability of the random field  $u$  in (2.5.11) and looking carefully at the proof of the identity (2.5.35), we could actually lower down slightly the hypothesis on  $\varphi$  in the Theorem 2.1.1, supposing that  $\varphi$  has only the second, the third and the fourth derivative bounded. Indeed the function  $\varphi'(u)$  will have linear growth and the right hand side of (2.5.34) will be always well defined. In this way the same density argument should provide to extend the Theorem 2.1.1 even in this case. These slight modifications should allow us to obtain a differential and an integral formula even for the random field  $u^2$ , giving an interesting decomposition of this random field.

## 2.6 Renormalisation constants

We calculate the asymptotic behaviour of the renormalisations constants defined in (2.3.30), (2.3.31). A preliminary result to analyse them lies on a remarkable identity on  $G$ , the heat kernel on  $\mathbb{R}$ , interpreted as a function  $G: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ .

**Lemma 2.6.1.** *For any  $z \in \mathbb{R}^2 \setminus \{0\}$  one has*

$$2 \int_{\mathbb{R}^2} G_x(z - \bar{z}) G_x(-\bar{z}) d\bar{z} = G(z) + G(-z) \quad (2.6.1)$$

*Proof.* We verify this identity by calculating the space-time Fourier transform

$$f \rightarrow \widehat{f}(\xi) = \int_{\mathbb{R}^2} e^{-2\pi i(\xi \cdot z)} f(z) dz$$

of both sides. In order to do that, we recall the elementary identity

$$\widehat{G}(\xi) = \frac{1}{2\pi i \xi_1 + 4\pi^2 \xi_2^2}.$$

Using the notation  $\bar{u}(z) = u(-z)$ , for any function  $u: \mathbb{R}^2 \rightarrow \mathbb{R}$ , we rewrite the left hand side of (2.6.1) as  $2G_x * \overline{G_x}(z)$ . Applying the Fourier transform, we then obtain

$$2\widehat{G_x * \overline{G_x}}(\xi) = 2\widehat{G_x}(\xi)\widehat{\overline{G_x}}(\xi) = 2(2\pi i \xi_2 \widehat{G}(\xi))(-2\pi i \xi_2 \widehat{G}(-\xi)) = \frac{8\pi^2 \xi_2^2}{4\pi^2 \xi_1^2 + (4\pi^2 \xi_2^2)^2}.$$

On the other hand, the same operation on the right hand side of (2.6.1) implies

$$\widehat{G}(\xi) + \widehat{G}(-\xi) = \frac{8\pi^2 \xi_2^2}{(2\pi i \xi_1 + 4\pi^2 \xi_2^2)(-2\pi i \xi_1 + 4\pi^2 \xi_2^2)} = \frac{8\pi^2 \xi_2^2}{4\pi^2 \xi_1^2 + (4\pi^2 \xi_2^2)^2}.$$

By uniqueness of the Fourier transform, we conclude.  $\square$

**Lemma 2.6.2.** *Let  $C_\varepsilon^1, C_\varepsilon^2$  be the constants introduced in (2.3.30), (2.3.31). Then the following estimates hold as  $\varepsilon \rightarrow 0^+$*

$$C_\varepsilon^1 = \frac{1}{\varepsilon} \int_{\mathbb{R}^2} G(s, y) \rho^{*2}(s, y) ds dy + o(1); \quad (2.6.2)$$

$$C_\varepsilon^2 = \frac{1}{\varepsilon} \int_{\mathbb{R}^2} (G_x * \rho)^2(s, y) ds dy + o(1); \quad (2.6.3)$$

$$C_\varepsilon^1 = C_\varepsilon^2 + o(1). \quad (2.6.4)$$

*Proof.* All the integrals we consider in the proof will be taken on the whole space  $\mathbb{R}^2$ , therefore we will not write it explicitly. Moreover for any function  $F: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ , any integer  $m$  and  $\varepsilon > 0$ , we introduce the shorthand notation

$$S_\varepsilon^m(F)(t, x) := \varepsilon^m F(\varepsilon^2 t, \varepsilon x).$$

Using the definition of  $C_\varepsilon^1$ , together with the hypothesis  $\rho(-z) = \rho(z)$  one has

$$\begin{aligned} C_\varepsilon^1 &= \int \int K(w) \rho_\varepsilon(z) \rho_\varepsilon(z - w) dw dz \\ &= \int K(w) \int \rho_\varepsilon(z) \rho_\varepsilon(w - z) dz dw = \int K(w) (\rho_\varepsilon)^{*2}(w) dw. \end{aligned}$$

A simple change of variable tells us that  $(\rho_\varepsilon)^{*2}(w) = (\rho^{*2})_\varepsilon(w)$ . Therefore we deduce that

$$C_\varepsilon^1 = \int K(t, x) \varepsilon^{-3} \rho^{*2} \left( \frac{t}{\varepsilon^2}, \frac{x}{\varepsilon} \right) dt dx = \int K(\varepsilon^2 t, \varepsilon x) \rho^{*2}(t, x) dt dx$$



$$= \frac{1}{\varepsilon} \int S_\varepsilon^1(K)(t, x) \rho^{*2}(t, x) dt dx.$$

Since  $S_\varepsilon^1(K)$  is equal to  $S_\varepsilon^1(G)$  as  $\varepsilon \rightarrow 0^+$  and  $G$  satisfies  $S_\varepsilon^1(G) = G$ , one has

$$S_\varepsilon^1(K)(t, x) \rho^{*2}(t, x) \rightarrow G(t, x) \rho^{*2}(t, x) \quad \text{a.e.}$$

Moreover, it is straightforward to show that the function  $G \rho^{*2}$  is integrable and it dominates  $S_\varepsilon^1(K) \rho^{*2}$ , therefore we obtain

$$\int S_\varepsilon^1(K)(t, x) \rho^{*2}(t, x) dt dx \rightarrow \int G(t, x) \rho^{*2}(t, x) dt dx,$$

by dominated convergence. We recover the identity (A.2), by using the decomposition  $G = K + R$ , as explained in the Lemma 2.3.11. Writing again  $S_\varepsilon^1(G) = G$ , we obtain

$$\frac{1}{\varepsilon} \int [G(t, x) - S_\varepsilon^1(K)(t, x)] \rho^{*2}(t, x) dt dx = \int S_\varepsilon^0(R)(t, x) \rho^{*2}(t, x) dt dx.$$

Since the function  $R$  is smooth, bounded and  $R(0, 0) = 0$  and  $\rho^{*2}$  is compactly supported, by sending  $\varepsilon \rightarrow 0^+$  it is straightforward to conclude

$$\int S_\varepsilon^0(R)(t, x) \rho^{*2}(t, x) dt dx \rightarrow R(0, 0) \int \rho^{*2}(t, x) dt dx = 0.$$

Thus the estimate (A.2) holds. Passing to the identity (A.3), we rewrite  $C_\varepsilon^2$  as

$$C_\varepsilon^2 = \int_{\mathbb{R}^2} (K_x * \rho_\varepsilon)^2(z) dz = \frac{1}{\varepsilon} \int \varepsilon (K_x * \rho_\varepsilon)^2(z) dz.$$

Let us express  $\int \varepsilon (K_x * \rho_\varepsilon)^2(z) dz$  in terms of the operator  $S_\varepsilon^2$  and of  $\rho$ . We apply a change of variable to get

$$(K_x * \rho_\varepsilon)(\varepsilon^2 t, \varepsilon x) = \int K_x(\varepsilon^2 t - \varepsilon^2 s, \varepsilon x - \varepsilon y) \rho(s, y) ds dy = (S_\varepsilon^0(K_x) * \rho)(t, x).$$

Therefore for any  $z = (t, x)$  we write

$$(K_x * \rho_\varepsilon)(t, x) = (S_\varepsilon^0(K_x) * \rho) \left( \frac{t}{\varepsilon^2}, \frac{x}{\varepsilon} \right).$$

Applying again the same change of variable we obtain

$$\int \varepsilon (K_x * \rho_\varepsilon)^2(z) dz = \varepsilon^4 \int (S_\varepsilon^0(K_x) * \rho)^2(t, x) dt dx = \int (S_\varepsilon^2(K_x) * \rho)^2(t, x) dt dx.$$

By sending  $\varepsilon \rightarrow 0^+$  the function  $S_\varepsilon^2(K_x)$  becomes equal to  $S_\varepsilon^2(G_x)$  and, using the scaling relation  $S_\varepsilon^2(G_x) = G_x$ , for a.e. couple of points  $(t, x)$ ,  $(s, y)$  one has

$$S_\varepsilon^2(K_x)(t - s, x - y) \rho(s, y) \rightarrow G_x(t - s, x - y) \rho(s, y).$$

The function  $G_x(t-s, x-y)\rho(s, y)$  is clearly integrable in both variables  $(t, x)$   $(s, y)$ . Therefore as a consequence of Fubini's theorem we get

$$(S_\varepsilon^2(K_x) * \rho)(t, x) \rightarrow (G_x * \rho)(t, x) \quad \text{a.e.}$$

which implies trivially

$$(S_\varepsilon^2(K_x) * \rho)^2(t, x) \rightarrow (G_x * \rho)^2(t, x) \quad \text{a.e.}$$

The function  $(G_x * \rho)^2$  is also integrable and it dominates  $(S_\varepsilon^2(K_x) * \rho)^2$ , then

$$\int (S_\varepsilon^2(K_x) * \rho)^2(t, x) dt dx \rightarrow \int (G_x * \rho)^2(z) dz.$$

Writing again the decomposition  $G = K + R$  as explained in the Lemma 2.3.11 and then using Cauchy-Schwarz inequality we deduce

$$\begin{aligned} & \frac{1}{\varepsilon} \int (G_x * \rho)^2(z) dz - \frac{1}{\varepsilon} \int \varepsilon (K_x * \rho_\varepsilon)^2(z) dz \\ &= \frac{1}{\varepsilon} \int (S_\varepsilon^2(G_x) * \rho)^2(z) - (S_\varepsilon^2(K_x) * \rho)^2(z) dz \\ &= \frac{1}{\varepsilon} \int 2(S_\varepsilon^2(K_x) * \rho)(z)(S_\varepsilon^2(R_x) * \rho)(z) dz + \frac{1}{\varepsilon} \int (S_\varepsilon^2(R_x) * \rho)^2(z) dz \\ &= \int 2(S_\varepsilon^2(K_x) * \rho)(z)(S_\varepsilon^1(R_x) * \rho)(z) dz + \varepsilon \int (S_\varepsilon^1(R_x) * \rho)^2(z) dz \\ &\leq 2 \left( \int (S_\varepsilon^2(K_x) * \rho)^2(z) dz \right)^{1/2} \left( \int (S_\varepsilon^1(R_x) * \rho)^2(z) dz \right)^{1/2} + \varepsilon \int (S_\varepsilon^1(R_x) * \rho)^2(z) dz \end{aligned}$$

Now, the function  $R$  is smooth, bounded and it satisfies  $R(0, 0) = 0$ . Using the identity  $S_\varepsilon^1(R_x) * \rho = S_\varepsilon^0(R) * \rho_x$ , it is straightforward to show that

$$(S_\varepsilon^1(R_x) * \rho)(z) \rightarrow 0 \quad \text{a.e.}$$

Moreover we can bound  $(S_\varepsilon^0(R) * \rho_x)^2$  by an integrable function. Thus we obtain

$$\int (S_\varepsilon^1(R_x) * \rho)^2(z) dz \rightarrow 0,$$

and the identity (A.3) follows. To finally prove the identity (A.4), it is sufficient to show

$$\int G(s, y) \rho^{*2}(s, y) ds dy = \int (G_x * \rho)^2(s, y) ds dy.$$

Starting from the identity (2.6.1), we convolve both sides of the equation with the function  $\rho^{*2} = \rho * \rho$ . Therefore for any  $u \in \mathbb{R}^2$  the left hand side of (2.6.1) becomes

$$2(G_x * \overline{G_x}) * \rho^{*2}(u) = \int \int \int 2G_x(u-v-w)G_x(-w)\rho(v-x)\rho(x) dx dv dw.$$

Choosing the following change of variable

$$\begin{cases} x' = x \\ v' = v - x \\ w' = w + x \end{cases} \quad \begin{cases} x = x' \\ v = v' + x' \\ w = w' - x' \end{cases} \quad dx'dv'dw' = dx'dv'dw',$$

the integral becomes

$$\int \int \int 2G_x(u - v' - w')G_x(-w' + x')\rho(v')\rho(x')dx'dv'dw'.$$

Using the identity  $\rho(x') = \rho(-x')$  the above integral equals

$$\begin{aligned} 2 \int \int \int G_x(u - v' - w')G_x(-w' - x')\rho(v')\rho(x')dx'dv'dw' = \\ 2 \int (G_x * \rho)(u - w')(G_x * \rho)(-w')dw'. \end{aligned}$$

On the other hand, the right hand side of (2.6.1) convolved with  $\rho^{*2}$  gives

$$\begin{aligned} \int G(u - w)\rho^{*2}(w)dw + \int G(w - u)\rho^{*2}(w)dw = \\ \int G(u - w)\rho^{*2}(w)dw + \int G(-u - w)\rho^{*2}(w)dw = (G * \rho^{*2})(u) + (G * \rho^{*2})(-u). \end{aligned}$$

Evaluating both sides in  $u = 0$ , we finally conclude. □

# Chapter 3

## Rough change of variable formulae and quasi-geometric rough paths

### 3.1 Introduction

Let  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  be a smooth function. The classical change of variable formula states that for any  $C^1$  path  $Y: [0, T] \rightarrow \mathbb{R}^d$ ,  $Y = (Y_1, \dots, Y_d)$  one has the identity

$$f(Y(t)) = f(Y(s)) + \sum_{i=1}^d \int_s^t \partial_i f(Y(r)) Y'_i(r) dr, \quad (3.1.1)$$

where  $Y' = (Y'_1, \dots, Y'_d)$  is the derivative of  $Y$  and  $\partial_i f = \partial f / \partial x_i$ . The formula (3.1.1) is a milestone of standard calculus and it holds if and only if  $Y$  is an absolutely continuous function. In particular when the path  $Y$  is not absolutely continuous, the Lebesgue integral in (3.1.1) is not well defined because  $Y$  might not be differentiable.

A remarkable example where this phenomenon happens comes from the stochastic analysis when  $Y$  is the realisation of a continuous semi-martingale and the trajectories of  $Y$  are not even a bounded variation function. One of the big achievement of the stochastic calculus theory was to show that even in this case there exists a new change of variable formula on  $Y$ , the well celebrated Itô formula (see [RY04, Thm 3.3])

$$f(Y_t) = f(Y_s) + \sum_{i=1}^d \int_s^t \partial_i f(Y_r) dY_r^i + \frac{1}{2} \sum_{i,j=1}^d \int_s^t \partial_{ij} f(Y_r) d\langle Y^i, Y^j \rangle_r, \quad (3.1.2)$$

where differently from (3.1.1) for any any given sequence of partitions  $\mathcal{P}_n$  of  $[s, t]$  such that its mesh size  $|\mathcal{P}_n|$  converges to zero, the integrals of (3.1.2) are continuous processes which can be defined through the identities

$$\begin{aligned} \int_s^t \partial_i f(Y_r) dY_r^i &:= \lim_{n \rightarrow \infty} \sum_{t_k \in \mathcal{P}_n} \partial_i f(Y_{t_k}) (Y_{t_{k+1}}^i - Y_{t_k}^i), \\ \int_s^t \partial_{ij} f(Y_r) d\langle Y^i, Y^j \rangle_r &:= \lim_{n \rightarrow \infty} \sum_{t_k \in \mathcal{P}_n} \partial_{ij} f(Y_{t_k}) (Y_{t_{k+1}}^i - Y_{t_k}^i) (Y_{t_{k+1}}^j - Y_{t_k}^j). \end{aligned}$$

The convergence of these discrete sums requires deeply the semimartingale hypothesis and it does not a.s. but it holds only as a convergence in probability uniformly on  $s, t \in [0, T]$  (see [RY04]). The formula (3.1.2) is not the unique way to express a change of variable formula. Using the so called Stratonovich integration (see [Str66]) for any semimartingale  $Y$  we can write also a formula which looks formally equivalent to (3.1.1)

$$f(Y_t) = f(Y_s) + \sum_{i=1}^d \int_s^t \partial_i f(Y_r) d^\circ Y_r^i, \quad (3.1.3)$$

where the integrals of (3.1.3) can be defined similarly by means of the limit

$$\int_s^t \partial_i f(Y_r) d^\circ Y_r^i := \lim_{n \rightarrow \infty} \sum_{t_k \in \mathcal{P}_n} \frac{\partial_i f(Y_{t_k}) + \partial_i f(Y_{t_{k+1}})}{2} (Y_{t_{k+1}}^i - Y_{t_k}^i).$$

The formulae (3.1.2) and (3.1.3) were two important breakthroughs in the modern probability theory and they have been extended in many directions (see e.g. [ERV02] for càdlàg processes or [GRV03] for the fractional Brownian motion).

In this chapter we present an alternative approach to the stochastic change of variable formulae expressed in (3.1.2) and we will explain how to extend the formula (3.1.1) when  $Y$  is a  $\gamma$ -Hölder function for some  $\gamma \in (0, 1)$  satisfying some additional hypotheses. The fundamental tool to express these new identities will be the rough path theory. Introduced in [Lyo98] and reformulated in [Gub04], the main focus of this theory was to give a rigorous meaning to this type of dynamical systems, called *rough differential equations*

$$\begin{cases} \dot{Y}(t) = \sum_{i=1}^e f_i(Y(t)) \dot{X}^i(t) \\ Y(0) = y_0 \end{cases} \quad (3.1.4)$$

where  $f_i: \mathbb{R}^d \rightarrow \mathbb{R}^d$  are smooth vector fields and  $X: [0, T] \rightarrow \mathbb{R}^e$  is generic  $\gamma$ -Hölder function. In particular it is possible to define a well posed integration theory of a path  $Z$  with respect to the increments of  $X$ , the *rough integral*, as long as we start from two enlarged objects: one containing the path  $X$  and some other objects representing its higher order integrals, a *rough path*; the other one, a *controlled rough path*, obtained from  $Z$  and other additional paths associated to it (see the Section 3.3). This rough integration will be the key notion to formulate a so called rough change of variable formula (see also [FH14, Chap. 5] for an introduction on this topic).

A first general theory to express these new identities was given in the last chapter of David Kelly's Phd Thesis [Kel12, Chap. 5]. Following its results it turns out that even in this context it is possible to formulate two general types of formulae very similar to the semimartingale case: a first one formally equal to (3.1.1) and (3.1.3) (see Proposition 3.4.1) and a second one similar to (3.1.2), where it appears a correction containing some higher order derivatives of the function  $f$  (see Proposition 3.4.10). These two different results are indeed related to two different families of rough path where the original problem of the change of variable formula was posed: the geometric and the branched rough paths (see [HK15], [Gub10]). In order to obtain a formula in

the branched case the author introduced the notion of *bracket extension* (see Definition 3.4.8) which extend in an algebraic and analytical context the classical notion of quadratic variation of a semimartingale. The notion of bracket extension presented in these works, however turns out to be quite technical to check on some explicit examples and for the time being it has not been linked with any other properties of branched rough paths.

In what follows we will introduce rigorously a change of variable formula starting from an another class of rough paths: the so called *quasi-geometric rough paths* (see Definition 3.5.7). The origin of this family can be traced back in [BCEF18] and its definition was already sketched by David Kelly in an oral talk of 2013 at the ICMAT of Madrid. Essentially the main idea behind their definition comes by simply modifying the definition of a geometric rough path in order to preserve an alternative product defined on the tensor algebra: the *quasi-shuffle product* (see [Hof00] [HI17]). Using some elementary algebraic properties of this operation, we can state two remarkable properties of this new class of rough paths: first we can associate to each quasi-geometric rough path an explicit bracket extension. This property will imply the existence of a non-geometric change of variable formula over this new class of rough paths. Then we will show the existence of a one to one bijection (the so called *Hoffman isomorphism*) between this new family of rough paths and a class of the geometric rough path over an extended path (see Proposition 3.5.8). The resulting isomorphism will allow us to manipulate quasi-geometric rough path as if they were geometric showing the existence of an intrinsic change of variable formula without using Kelly's results.

The present chapter will be organised as follows. We first recall the main properties of the rough paths theory in Section 2 and 3. We then briefly detail the main ideas behind Kelly's theory in Section 3.4, and we conclude with the last two sections where we study the general properties of quasi-geometric rough paths.

As mentioned at the beginning, all the results in this chapter are deterministic but in principle they can be applied to the realisation of any a.s. Hölder continuous stochastic process. We remark that a recent paper [CP19] studies the same problem of rough change of variable formulae, establishing some interesting connections with the theory of Föllmer calculus (see [Föl81]).

## 3.2 Rough paths of any Hölder regularity

We introduce the basic constructions and the notations of the rough path theory. Unlike the classical case, we want to describe the family of higher order integrals starting from a path  $X = (X^a)_{a \in A}$  whose components are indexed by a an *alphabet*  $A$ , which is not necessarily  $\{1, \dots, d\}$  but a generic finite set. By analogy the elements of  $A$  are called *letters*.

### 3.2.1 Algebraic preliminaries

The two main algebraic structures where the rough paths theory are formulated depend on two specific *Hopf Algebras*: the tensor algebra and the Butcher-Connes-

Kreimer Hopf algebra (see e.g. [Man04] for a general introduction to the theory of Hopf algebras). We recall briefly their definition.

### Tensor algebra

For any real vector space  $V$  we will adopt the notation  $T(V)$  to denote the tensor algebra of  $V$  and  $\otimes$  for the usual tensor product of  $T(V)$ . In order to distinguish between  $\otimes$  and the tensor product of two generic vector spaces  $W, U$  we will adopt the notation  $W \tilde{\otimes} U$  for the latter. In this way every bilinear product on  $T(V)$  can be described as a function  $m: T(V) \tilde{\otimes} T(V) \rightarrow T(V)$ . The duality between one vector space and its dual is denoted by  $\langle \cdot, \cdot \rangle$ . In case of an alphabet  $A$  we denote by  $\mathbb{R}^A$  the free vector space on  $A$ , namely the set of formal linear combination of elements of  $A$ . We define the tensor algebra  $T(A)$  as the tensor algebra of  $\mathbb{R}^A$ , that is

$$T(A) := \bigoplus_{k=0}^{+\infty} T_k(A),$$

where, by convention,  $T_k(A) = ((\mathbb{R}^A)^{\otimes k})$  and we set  $(T_0(A))^{\otimes 0} = \mathbb{R} \simeq \text{span}\{\mathbf{1}\}$ . There are two possible ways to denote the elements of  $T(A)$ : on one hand, using the canonical basis of  $\mathbb{R}^A$   $\{e_a\}_{a \in A}$ , we can express an element  $v \in T(A)$  as a finite linear combination of tensors  $e_{a_1} \otimes \cdots \otimes e_{a_k}$  for some letters  $a_1, \dots, a_k \in A$  and  $k \geq 1$ . On the other hand we can use the identification

$$e_{a_1} \otimes \cdots \otimes e_{a_n} \sim a_1 \cdots a_n, \quad \mathbf{1} \sim \text{empty word}$$

and the elements of  $T(A)$  can be represented as a finite linear combination of *words* built from the alphabet  $A$  (we will use both notations). Thus  $T(A)$  can be also seen as the free vector space on  $W(A)$ , the set of words given by  $A$ . Three operations can be naturally defined on  $T(A)$ :

- The concatenation product  $\otimes: T(A) \tilde{\otimes} T(A) \rightarrow T(A)$ , defined for any couple of words  $v = a_1 \cdots a_n, w = b_1 \cdots b_m \in W(A)$  as the usual tensor product in term of words, that is

$$v \otimes w := a_1 \cdots a_n b_1 \cdots b_m$$

and extended linearly. In order to simplify the notation we will denote this operation as a simple juxtaposition on  $T(A)$ .

- The shuffle product  $\sqcup: T(A) \tilde{\otimes} T(A) \rightarrow T(A)$ , defined by the recursive relations

$$\begin{aligned} \mathbf{1} \sqcup v &:= v \sqcup \mathbf{1} := v, \\ av \sqcup bw &:= (v \sqcup bw)a + (av \sqcup w)b, \end{aligned} \tag{3.2.1}$$

for any couple of words  $v, w \in W(A)$  and letters  $a, b \in A$  and extended linearly.

- The deconcatenation coproduct  $\nabla: T(A) \rightarrow T(A) \tilde{\otimes} T(A)$ , defined by the relations

$$\begin{aligned} \nabla \mathbf{1} &= \mathbf{1} \tilde{\otimes} \mathbf{1} \\ \nabla w &:= \sum_{i=0}^m a_1 \cdots a_i \tilde{\otimes} a_{i+1} \cdots a_m, \end{aligned} \tag{3.2.2}$$

for any word  $w = a_1 \cdots a_m \in W(A)$  and extended linearly

Using the last two operations it is a standard result in the algebraic literature (see e.g. [Reu93]) that the triple  $(T(A), \sqcup, \nabla)$  is a commutative bialgebra where the empty word  $\mathbf{1}$  and its dual element  $\mathbf{1}^*: T(A) \rightarrow \mathbb{R}$  are the unit and the counit. We recall that the shuffle product  $\sqcup$  admits also an explicit expression in terms of permutations. Indeed for all couples of words  $w, v \in T(A)$  such that the length of  $w$  is  $n$  and the length of  $v$  is  $k$  we have the identity

$$w \sqcup v := \sum_{\sigma \in \text{Shuf}(w,v)} (wv)_\sigma,$$

where  $\sigma \in \text{Shuf}(w,v)$  is a permutation of  $n + k$  element which preserves the original ordering of the sequences in  $w$  and  $v$  and for any word  $u$   $(u)_\sigma$  is the word where the indices are permutated following  $\sigma$ . The bialgebra  $T(A)$  has a natural graduation by definition and we remark that  $(\mathbb{R}^A)^{\otimes k}$  is generated by all words  $w$  whose length  $|w|$  is equal to  $k$ . Moreover it is also straightforward to show that the operations  $\sqcup$  and  $\nabla$  are compatible with the bialgebra structures, meaning that the operations  $\sqcup$  and  $\nabla$  satisfy

$$\sqcup: T_k(A) \tilde{\otimes} T_h(A) \rightarrow T_{k+h}(A), \quad \nabla: T_k(A) \rightarrow \bigoplus_{p+q=k} T_p(A) \tilde{\otimes} T_q(A),$$

for any integers  $k, h \geq 0$ . Thus the triple  $(T(A), \sqcup, \nabla)$  becomes a connected and graduated bialgebra. From a classical result on bialgebra (see e.g. [Abe04]) these properties imply automatically that  $T(A)$  is also a Hopf algebra.

We denote the algebraic dual of  $T(A)$  by  $T((A))$  and the algebraic dual of  $T_k(A)$  by  $T_k((A))$ . By definition any element of  $T((A))$  can be expressed as a formal series of terms in  $T_k((A))$ . The dual space  $T((A))$  carries a general product operation  $*$ :  $T((A)) \tilde{\otimes} T((A)) \rightarrow T(A)$  which is defined by duality from the relation

$$\langle u * v, x \rangle := \langle v \tilde{\otimes} w, \nabla x \rangle,$$

for any  $x \in T(A)$  and  $u, v \in T((A))$ . We will call it the convolution product. In case of the tensor algebra  $T(A)$  this convolution product can be explicitly described as the concatenation product  $\otimes$  applied to the dual words. Therefore we can adopt this second notation without ambiguity. A fundamental object for the theory of the rough paths is given by the set of real characters  $\mathbb{G}(A)$ , defined as

$$\mathbb{G}(A) := \{h \in T((A)): \text{for every } a, b \in T(A) \quad \langle h, a \sqcup b \rangle = \langle h, a \rangle \langle h, b \rangle\}, \quad (3.2.3)$$

From a classical Hopf Algebra result (and in this case we can explicitly check it with a straightforward calculation) the couple  $(\mathbb{G}(A), \otimes)$  is a group where the counit  $\mathbf{1}^*$  is the identity element. For any  $N \geq 0$  we can consider the truncated dual tensor algebra  $T^N(A) := \bigoplus_{n=0}^N T_n((A))$  and its dual space  $T^N((A))$ . We define the *A-decorated step-N free nilpotent group*  $\mathbb{G}_N(A)$  by the quotient

$$\mathbb{G}_N(A) := \left( \mathbb{G}(A) / \bigoplus_{n=N+1}^{+\infty} T_n((A)) \right) \subset T^N((A)), \quad (3.2.4)$$



where all the dual words whose homogeneity is bigger than  $N + 1$  are identified to 0. We stress the fact that the quotient in (3.2.4) is not a group quotient but, including  $\mathbb{G}(A)$  in  $T((A))$ , the set  $\mathbb{G}_N(A)$  is the image of the vector space quotient map

$$\pi_N: T((A)) \rightarrow \left( T((A)) / \bigoplus_{n=N+1}^{+\infty} T_n((A)) \right) = T^N((A)).$$

Remarkably the group operation  $\otimes$  is still a group operation on  $\mathbb{G}_N(A)$ . We will adopt the shorthand notation  $(\mathbb{G}_N(A), \otimes)$  to recall this group.

**Butcher-Connes-Kreimer Hopf Algebra**

For any given alphabet  $A$ , an  $A$ -labelled rooted trees is given by a triple  $(\tau, r, \mathcal{L})$  where  $\tau$  is a combinatorial tree (a finite connected simple graph) with a non-empty set of nodes  $N_\tau$  and a set of edges  $E_\tau$ ,  $r \in N_\tau$  is a fixed vertex called the *root* and  $\mathcal{L}: N_\tau \rightarrow A$  is a fixed function called *labelling*. The set of  $A$ -labelled trees is denoted by  $\mathcal{T}(A)$ . A finite disjoint union of  $A$ -labelled rooted trees is called an  $A$ -labelled *forest* and the set of all  $A$ -labelled forest is denoted by  $\mathcal{F}(A)$ . By convention we introduce also the empty forest  $\mathbf{1}$  (just the same notation of the empty word) and we consider it as an element of  $\mathcal{F}(A)$ . For any finite family of trees  $\tau_1, \dots, \tau_n$  the forest obtained by the disjoint union of the previous trees is denoted by  $f = \tau_1 \cdots \tau_n$  independently of the order of the trees (we represent the disjoint union as a commutative operation on forests). We can graphically represent trees or forests by simply putting the root at the bottom and decorating each node with the corresponding label. For instance if  $A = \{1, 2\}$  one has

$$\bullet_1, \mathbf{!}_1^2, \mathbf{!}_2^1 \in \mathcal{T}(A); \quad \mathbf{!}_1^2 \in \mathcal{F}(A).$$

Starting from the empty forest  $\mathbf{1}$  we can recursively describe  $\mathcal{T}(A)$  by means of the *grafting* maps  $\{B_+^a: \mathcal{F}(A) \rightarrow \mathcal{T}(A)\}_{a \in A}$ . That is for any  $a \in A$  we define  $B_+^a(\mathbf{1}) := \bullet_a$  and for any forest  $\tau_1 \cdots \tau_n$  the tree  $B_+^a(\tau_1 \cdots \tau_n)$  is graphically given by

$$B_+^a(\tau_1 \cdots \tau_n) = \begin{array}{c} \tau_1 \quad \dots \quad \tau_n \\ \diagup \quad \vdots \quad \diagdown \\ \bullet_a \end{array}$$

(since we consider combinatorial graphs, all the graphical representations above are identified to a single tree). The free vector space on  $\mathcal{F}(A)$  is denoted by  $\mathcal{H}(A)$  and the grafting maps can be extended linearly on  $\mathcal{H}(A)$ . Three main operations are defined on  $\mathcal{H}(A)$ .

- The free commutative product  $m: \mathcal{H}(A) \tilde{\otimes} \mathcal{H}(A) \rightarrow \mathcal{H}(A)$ , which coincides with the linear extension of the disjoint union of graphs.
- The Butcher-Connes-Kreimer coproduct  $\Delta: \mathcal{H}(A) \rightarrow \mathcal{H}(A) \tilde{\otimes} \mathcal{H}(A)$ , which is given by the base condition  $\Delta \mathbf{1} = \mathbf{1} \tilde{\otimes} \mathbf{1}$  and satisfies recursively the identities

$$\Delta f := \Delta \tau_1 \cdots \Delta \tau_n, \quad \Delta(B_+^a(\sigma)) := B_+^a(\sigma) \tilde{\otimes} \mathbf{1} + (\text{id} \tilde{\otimes} B_+^a) \Delta \sigma,$$

for any forest  $f = \tau_1 \cdots \tau_n$ .

- The reduced coproduct  $\Delta': \mathcal{H}(A) \rightarrow \mathcal{H}(A) \tilde{\otimes} \mathcal{H}(A)$ , defined for any forest  $\sigma$  as

$$\Delta' \sigma := \Delta \sigma - \sigma \tilde{\otimes} \mathbf{1} - \mathbf{1} \tilde{\otimes} \sigma, \quad (3.2.5)$$

and extended linearly.

For any tree  $\tau$  we can equivalently rewrite  $\Delta \tau$  working directly on the graphical structure of tree  $\tau$  (see e.g. [Man04]). A cut  $C$  of a tree is a subset of  $E_\tau$ . A cut  $C$  is said admissible if  $C$  is non empty and any path starting from the root of  $\tau$  to any other node contains at most one element of  $C$ . We denote by  $Ad(\tau)$  the set of all admissible cuts of the tree  $\tau$ . Given a tree  $\tau$  and an admissible cut  $C$ , we can erase the edge of  $C$  from  $\tau$  to obtain a forest. We will denote by  $R^C(\tau)$  the tree containing the root after the cancellation and by  $P^C(\tau)$  the remaining forest. Thus we obtain the alternative formula

$$\Delta \tau = \mathbf{1} \tilde{\otimes} \tau + \tau \otimes \mathbf{1} + \sum_{C \in Ad(\tau)} P^C(\tau) \tilde{\otimes} R^C(\tau).$$

We will abbreviate the definition of  $\Delta$  using the Sweedler notation  $\Delta \tau = \tau^{(1)} \tilde{\otimes} \tau^{(2)}$ . The triple  $(\mathcal{H}(A), \cdot, \Delta)$  defined above, together with the natural modification of  $\mathbf{1}$  and  $\mathbf{1}^*$  in this context, is again a bialgebra. The set  $\mathcal{H}(A)$  is naturally graded with respect to  $|\sigma|$ , the number of nodes of the forest  $\sigma$ . That is

$$\mathcal{H}(A) = \bigoplus_{k=0}^{+\infty} \mathcal{H}_k(A), \quad \mathcal{H}_k(A) := \text{span}\{v \in \mathcal{F}(A) : |v| = k\}.$$

Also in this case it is straightforward to show that this decomposition is compatible with the bialgebra structure and  $\mathcal{H}_0(A) = \mathbb{R}\mathbf{1}$ . Thus for the same reasons as before  $(\mathcal{H}(A), \cdot, \Delta)$  is also a Hopf Algebra. We will call it the *A-decorated Butcher-Connes-Kreimer Hopf Algebra* (see [CK98], [But72] for the first definition of this algebra).

In analogy with the notation for the tensor algebra, we denote the algebraic dual of  $\mathcal{H}(A)$  and  $\mathcal{H}_k(A)$  by  $\mathcal{H}((A))$ ,  $\mathcal{H}_k((A))$ . The presence of a coproduct allows to define equivalently a convolution product  $*$ :  $\mathcal{H}((A)) \tilde{\otimes} \mathcal{H}((A)) \rightarrow \mathcal{H}((A))$  by duality in the same manner as before:

$$\langle \tau_1 * \tau_2, \sigma \rangle := (\tau_1 \tilde{\otimes} \tau_2) \Delta \sigma.$$

In the case of the Butcher-Connes-Kreimer Hopf algebra this operation is called in the literature the *Grossman-Larson* product and we will henceforth denote it by the symbol  $*$ . The set of real characters on  $\mathcal{H}(A)$  is denoted by  $G(A)$  and it is defined as in (3.2.3) with  $\mathcal{H}((A))$  and  $m$  instead of  $T((A))$  and  $\sqcup$ . Again the couple  $(G(A), *)$  is a group (which we call the *A-decorated Butcher group*), due to the Hopf algebra structure of  $\mathcal{H}(A)$ . For any  $N \geq 0$  we introduce the vector space  $\mathcal{H}^N(A) := \bigoplus_{n=0}^N \mathcal{H}_n(A)$  and  $\mathcal{H}^N((A))$  its algebraic dual. We denote by  $G_N(A)$  the quotient

$$G_N(A) := \left( G(A) / \bigoplus_{n=N+1}^{+\infty} \mathcal{H}_n((A)) \right) \subset \mathcal{H}^N((A)). \quad (3.2.6)$$

We call this set the *A decorated step-N Butcher group* and the quotient is interpreted in the same manner as  $\mathbb{G}(A)$ . By analogy we adopt the same notation for  $\pi_N$ , the quotient map of vector spaces and for the couple  $(G_N(A), *)$  is still a group.

We decided to specify the algebraic constructions of  $T(A)$  and  $\mathcal{H}(A)$  starting from a generic locally finite set  $A$  in order to be coherent with the framework of the quasi-shuffle algebras as explained in [HI17] (see Section 3.5.1). Nevertheless in what follows we will formulate our results on *finite* alphabets  $A$ . This hypothesis does not change anything in the *algebraic* construction above but it implies simply that each vector subspace  $T_k((A))$ ,  $\mathcal{H}_k((A))$  will be finite dimensional. If  $A = \{1, \dots, d\}$  the resulting algebraic structure are denoted by the classical notation  $\mathcal{H}(\mathbb{R}^d)$ ,  $\mathcal{H}((\mathbb{R}^d))$  and  $T(\mathbb{R}^d)$ ,  $T((\mathbb{R}^d))$ .

If  $A \subset A'$  we can always identify  $T(A)$  with a subspace of  $T(A')$ . This inclusion is also an injective Hopf algebra morphism and we will write it as  $T(A) \subset T(A')$ . Since the same reasoning applies also for  $\mathcal{H}(A)$  and  $\mathcal{H}(A')$ , we will adopt the same notation  $\mathcal{H}(A) \subset \mathcal{H}(A')$ . By duality we can define the inclusions  $j: T((A')) \rightarrow T((A))$ ,  $j': \mathcal{H}((A')) \rightarrow \mathcal{H}((A))$  defined for any  $h \in T((A'))$  or  $k \in \mathcal{H}((A'))$  as

$$\langle j(h), w \rangle = \langle h, w \rangle, \quad \langle j'(k), \sigma \rangle = \langle k, \sigma \rangle,$$

for any  $w \in W(A)$  and  $\sigma \in \mathcal{F}(A)$ . These inclusions are also Hopf algebra morphisms and in case of the groups  $\mathbb{G}(A')$  and  $\mathbb{G}(A)$  the maps  $j$  and  $j'$  are an injective homomorphism. We denote these inclusions as  $\mathbb{G}(A') \subset \mathbb{G}(A)$  and  $G(A') \subset G(A)$ .

### Anisotropic grading

In order to take into account the case where the letters of  $A$  are associated to some continuous path  $X^a$  with some Hölder regularity  $\gamma_a \in (0, 1)$  we introduce a different type of groups related to a different quotient of  $\mathbb{G}(A)$  and  $G(A)$ . We suppose that  $A$  is endowed with a sequence  $\Gamma = \{\gamma_a\}_{a \in A}$  of real numbers such that  $\gamma_a \in (0, 1)$  for all  $a \in A$  and  $\hat{\gamma} = \inf_{a \in A} \gamma_a > 0$ . We call the elements of  $\Gamma$  *weights* and we define every couple  $(A, \Gamma)$  satisfying these properties a *weighted alphabet*. Under these hypotheses for any word  $v = a_1 \cdots a_k \in W(A)$  and any forest  $\sigma \in \mathcal{F}(A)$  we can define the real numbers

$$\begin{aligned} |v|_\Gamma &:= \frac{\gamma_{a_1} + \dots + \gamma_{a_k}}{\hat{\gamma}} = \frac{1}{\hat{\gamma}} \sum_{a \in A} n_a(v) \gamma_a, \\ |\sigma|_\Gamma &:= \sum_{u \in N_\sigma} \frac{\gamma(\mathcal{L}(u))}{\hat{\gamma}} = \frac{1}{\hat{\gamma}} \sum_{a \in A} n_a(\sigma) \gamma_a, \end{aligned} \tag{3.2.7}$$

where  $n_a(v) = |\{j : v_j = a\}|$  and  $n_a(\sigma) = |\{v \in N_\sigma : \mathcal{L}(v) = a\}|$  and  $|\mathbf{1}|_\Gamma = 0$ . In both cases the function  $|\cdot|_\Gamma$  defined above has an additive property on words and forests because  $|uv|_\Gamma = |u|_\Gamma + |v|_\Gamma$  for each pair of words  $u, v$  and  $|\sigma\sigma'|_\Gamma = |\sigma|_\Gamma + |\sigma'|_\Gamma$  for any couple of forests  $\sigma, \sigma'$ . Thus for any word or forest  $\sigma$  the coproducts  $\nabla\sigma$  or  $\Delta\sigma$  is a sum of terms  $\sigma^{(1)} \otimes \sigma^{(2)}$  such that any term satisfies  $|\sigma^{(1)}|_\Gamma + |\sigma^{(2)}|_\Gamma = |\sigma|_\Gamma$ . This function

allows to define the sets

$$\begin{aligned} W_\Gamma(A) &:= \{v \in W(A) : |v|_\Gamma > \hat{\gamma}^{-1}\}, & W^\Gamma(A) &:= \{v \in W(A) : |v|_\Gamma \leq \hat{\gamma}^{-1}\}, \\ \mathcal{F}_\Gamma(A) &:= \{\sigma \in \mathcal{F}(A) : |\sigma|_\Gamma > \hat{\gamma}^{-1}\}, & \mathcal{F}^\Gamma(A) &:= \{\sigma \in \mathcal{F}(A) : |\sigma|_\Gamma \leq \hat{\gamma}^{-1}\}. \end{aligned} \quad (3.2.8)$$

We will denote by  $T_\Gamma(A)$ ,  $\mathcal{H}_\Gamma(A)$ ,  $T^\Gamma(A)$ ,  $\mathcal{H}^\Gamma(A)$  the subspaces generated respectively by  $W_\Gamma(A)$ ,  $\mathcal{F}_\Gamma(A)$ ,  $W^\Gamma(A)$  and  $\mathcal{F}^\Gamma(A)$  in the spaces  $T(A)$  and  $\mathcal{H}(A)$ . Moreover we can easily define the subspaces  $T_\Gamma((A))$ ,  $\mathcal{H}_\Gamma((A))$ ,  $T^\Gamma((A))$ ,  $\mathcal{H}^\Gamma((A))$  by duality. In analogy with the definition of the  $N$ -step free nilpotent group and the  $N$ -step Butcher group in (3.2.4) and (3.2.6) we introduce the corresponding quotient sets

$$\mathbb{G}_\Gamma(A) := \mathbb{G}(A)/T_\Gamma((A)) \subset T^\Gamma((A)), \quad G_\Gamma(A) := G(A)/\mathcal{H}_\Gamma((A)) \subset \mathcal{H}^\Gamma((A)). \quad (3.2.9)$$

and we call them the *anisotropic  $\Gamma$ -step free nilpotent group* (resp. the *anisotropic  $\Gamma$ -step Butcher group*). The additivity property of  $|\cdot|_\Gamma$  implies immediately that the operations  $\otimes$  and  $*$  are well defined on  $\mathbb{G}_\Gamma(A)$  and  $G_\Gamma(A)$  and the couples  $(\mathbb{G}_\Gamma(A), \otimes)$  and  $(G_\Gamma(A), *)$  are groups. We will adopt the notation  $\pi_\Gamma$  to denote the quotient map of vector spaces in both cases. By construction one has  $|v|_\Gamma \geq |v|$  and  $|\sigma|_\Gamma \geq |\sigma|$  for any word  $v$  and any forest  $\sigma$  (In case  $\Gamma$  is a constant sequence we have an equality). Thus it follows that  $T^\Gamma((A)) \subset T^N((A))$  and  $\mathcal{H}^\Gamma((A)) \subset \mathcal{H}^N((A))$  where  $N := \lfloor \hat{\gamma}^{-1} \rfloor$ . Using the classical graduation of  $T(A)$  and  $\mathcal{H}(A)$  we can write also

$$T^\Gamma(A) = \bigoplus_{n=0}^N T_n^\Gamma(A) \quad \mathcal{H}^\Gamma(A) = \bigoplus_{n=0}^N \mathcal{H}_n^\Gamma(A),$$

where  $T_n^\Gamma(A) := T_n(A) \cap T^\Gamma(A)$  and  $\mathcal{H}_n^\Gamma(A) := \mathcal{H}_n(A) \cap \mathcal{H}^\Gamma(A)$ .

The definition of  $\mathbb{G}_\Gamma(A)$  was already given in the recent work [TZ18, pag. 22] in an equivalent form and we simply extended this construction in the forest case. If  $\Gamma$  is a constant sequence equal to  $\gamma$  (which we call also *isotropic case*) then  $G_\Gamma(A) = G_M(A)$ , where  $M := \lfloor \gamma^{-1} \rfloor$ . In case we have two different weighted alphabets  $(A, \Gamma)$ ,  $(A', \Gamma')$  where  $A \subset A'$ ,  $\Gamma = \{\gamma_a\}_{a \in A}$  and  $\Gamma' = \{\gamma'_a\}_{a \in A}$  we will always impose  $\gamma'_a = \gamma_a$  for any  $a \in A$ , to be coherent with the anisotropic grading.

### 3.2.2 Geometric and branched rough paths

Let us introduce the formal definition of a rough path in the case of a weighted alphabet  $(A, \Gamma)$  where  $A$  is a finite set and  $\Gamma = \{\gamma_a\}_{a \in A}$  is a finite number of weights. Following the previous section the parameter  $\hat{\gamma} = \min_{a \in A} \gamma_a$  and the function  $|\cdot|_\Gamma$  given in (3.2.7) are uniquely identified. Starting from  $(A, \Gamma)$  we can easily consider  $\mathcal{C}^\Gamma(A)$ , the set of continuous functions  $X : [0, T] \rightarrow \mathbb{R}^A$ ,  $X_t = (X_t^a)_{a \in A}$  such that for any  $a \in A$  one has

$$\sup_{s \neq t} \frac{|X_t^a - X_s^a|}{|t - s|^{\gamma_a}} < +\infty.$$

We call the elements of  $\mathcal{C}^\Gamma(A)$   $\Gamma$ -Hölder path. The notion of a rough path wants to encode into a formal object all the *increments* and the *iterated integrals* of  $\Gamma$ -Hölder

paths and, as explained before, the geometric or the branched nature of a rough path are related to the integration by parts formulae that the coordinates of this path should satisfy.

**Definition 3.2.1.** A  $\Gamma$ -regular geometric rough path  $\mathbb{X}$  is a function  $\mathbb{X}: [0, T]^2 \rightarrow T((A))$  satisfying the following conditions:

- 1g) for every  $w_1, w_2 \in T(A)$   $\langle \mathbb{X}_{st}, w_1 \sqcup w_2 \rangle = \langle \mathbb{X}_{st}, w_1 \rangle \langle \mathbb{X}_{st}, w_2 \rangle$ ,
- 2g) for every  $s, t, u \in [0, T]^3$ ,  $\mathbb{X}_{st} = \mathbb{X}_{su} \otimes \mathbb{X}_{ut}$ ,
- 3g) for every word  $w \in W(A)$  one has

$$\sup_{s \neq t} \frac{|\langle \mathbb{X}_{st}, w \rangle|}{|t - s|^{\hat{\gamma}|w|_\Gamma}} < +\infty. \quad (3.2.10)$$

**Definition 3.2.2.** A  $\Gamma$ -regular branched rough path  $\mathbf{X}$  is a function  $\mathbf{X}: [0, T]^2 \rightarrow \mathcal{H}((A))$  satisfying the following conditions:

- 1b) for every  $\tau_1, \tau_2 \in \mathcal{H}(A)$   $\langle \mathbf{X}_{st}, \tau_1 \tau_2 \rangle = \langle \mathbf{X}_{st}, \tau_1 \rangle \langle \mathbf{X}_{st}, \tau_2 \rangle$ ,
- 2b) for every  $s, t, u \in [0, T]^3$   $\mathbf{X}_{st} = \mathbf{X}_{su} * \mathbf{X}_{ut}$ ,
- 3b) for every forest  $\sigma \in \mathcal{F}(A)$  one has

$$\sup_{s \neq t} \frac{|\langle \mathbf{X}_{st}, w \rangle|}{|t - s|^{\hat{\gamma}|w|_\Gamma}} < +\infty. \quad (3.2.11)$$

We denote the set of  $\Gamma$ -regular geometric and branched rough path by the symbols  $\mathcal{R}_g^\Gamma(A)$ ,  $\mathcal{R}_b^\Gamma(A)$ . We call the elements of both  $\mathcal{R}_g^\Gamma(A)$  and  $\mathcal{R}_b^\Gamma(A)$  rough paths and we adopt the shorthand notation  $\mathcal{R}^\Gamma(A)$  for these objects. An alternative to state the properties 2g) or 2b) in the literature is also given by saying that  $\mathbf{X}$  satisfies the Chen's relations.

*Remark 3.2.3.* Following [HK15] and [Gub10], the original definition of geometric and branched rough path is done when the underlying alphabet is  $A = \{1, \dots, d\}$  and  $\Gamma$  is a constant sequence with value  $\gamma \in (0, 1)$ . In this case we adopt the notation  $\mathcal{R}^\gamma(\mathbb{R}^d)$ ,  $\mathcal{R}_g^\gamma(\mathbb{R}^d)$ ,  $\mathcal{R}_b^\gamma(\mathbb{R}^d)$ . The name  $\Gamma$ -regular is given as a generalisation of [CEFMMK18, Def 4.1.]. We remark that in other works such as [TZ18] the elements of  $\mathcal{R}^\Gamma(A)$  are called *anisotropic rough paths*. An equivalent version of the geometric  $\Gamma$  rough path in the context of the  $p$  variation topology is given by the notion of  $\Pi$ -rough paths (see [Gyu16]) and the original formulation of the geometric rough paths in [Lyo98].

*Remark 3.2.4.* Recalling the inclusions  $\mathbb{G}(A') \subset \mathbb{G}(A)$  and  $G(A') \subset G(A)$  for any couple of weighted alphabets  $(A, \Gamma)$ ,  $(A', \Gamma')$  such that  $A \subset A'$ , we can identify any element of  $\mathcal{R}^{\Gamma'}(A')$  as an element of  $\mathcal{R}^\Gamma(A)$ . Under this identification if there exists  $\mathbf{X}' \in \mathcal{R}^{\Gamma'}(A')$  such that its inclusion in  $\mathcal{R}^\Gamma(A)$  coincides with  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  we say that  $\mathbf{X}'$  is over  $\mathbf{X}$ . The link between a rough path  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  and the increments of a function  $X \in \mathcal{C}^\Gamma(A)$  is trivially encoded in Chen's relations. Indeed for all  $a \in A$  and  $(s, t, u) \in [0, T]^3$  it is straightforward to show the identities

$$\langle \mathbf{X}_{st}, a \rangle = \langle \mathbf{X}_{su}, a \rangle + \langle \mathbf{X}_{ut}, a \rangle \quad \text{or} \quad \langle \mathbf{X}_{st}, \bullet_a \rangle = \langle \mathbf{X}_{su}, \bullet_a \rangle + \langle \mathbf{X}_{ut}, \bullet_a \rangle.$$

Thus there exists at least a function  $X' \in \mathcal{C}^\Gamma(A)$  such that  $\langle \mathbf{X}_{st}, a \rangle = (X')_t^a - (X')_s^a$  or  $\langle \mathbf{X}_{st}, \bullet_a \rangle = (X)_t^a - (X)_s^a$ . The resulting path  $X'$  is not unique but it is defined up to a translation constant. By choosing a specific element  $X \in \mathcal{C}^\Gamma(A)$  in this class, we resume this condition by saying that  $\mathbf{X}$  is a rough path over  $X$ .

Using the definition of the groups  $G(A)$  and  $\mathbb{G}(A)$ , an equivalent and more synthetic reformulation of the properties (1g) (2g) (1b) (2b) comes by simply stating that a rough path is a group-valued path  $\mathcal{X}: [0, T] \rightarrow (G, \circ)$  (where  $(G, \circ) = (\mathbb{G}(A), \otimes)$  or  $G(A) = (G(A), *)$ ) such that  $\mathcal{X}_0 = \mathbf{1}^*$ . Indeed for any given path  $\mathcal{X}$ , the definition

$$\mathbf{X}_{st} := \mathcal{X}_t \circ (\mathcal{X}_s)^{-1}$$

satisfies (1g) (2g) (1b) (2b) and for any fixed rough path  $\mathbf{X}$  the path  $t \rightarrow \mathbf{X}_{0t}$  is a group valued path. By definition the knowledge of one element  $g \in \mathbb{G}(A)$  or  $G(A)$  requires potentially the knowledge of a infinite number of real values (e.g. in case of  $G(A)$  a character is determined as long we know  $\langle g, \tau \rangle$  for any tree  $\tau \in \mathcal{T}(A)$ ). However the analytical conditions (3g) (3b) and the finiteness of  $A$  simplify radically the structure of a rough path, letting all the values of  $\mathbf{X}_{st}$  depend uniquely only on the trees or words  $h$  such that  $\hat{\gamma}|h|_\Gamma < 1$ . An algebraic way to consider only these terms and to formulate this dependency consists in replacing the groups  $\mathbb{G}(A)$   $G(A)$  with the quotients  $\mathbb{G}_\Gamma(A)$  and  $G_\Gamma(A)$ . Thus we can write the following definitions:

**Definition 3.2.5.** A geometric  $\Gamma$ -rough path  $\mathbb{X}$  is a function  $\mathbb{X}: [0, T]^2 \rightarrow \mathbb{G}_\Gamma(A)$  satisfying the following three conditions:

- 1)  $\mathbb{X}_{tt} = \mathbf{1}^*$  for every  $t \in [0, T]$ ;
- 2) for every  $(s, t, u) \in [0, T]^3$ ,  $\mathbb{X}_{st} = \mathbb{X}_{su} \otimes \mathbb{X}_{ut}$ ;
- 3) for every word  $w \in W^\Gamma(A)$ ,  $\sup_{s \neq t} |\langle \mathbb{X}_{st}, w \rangle| / |t - s|^{\hat{\gamma}|w|_\Gamma} < +\infty$ .

**Definition 3.2.6.** A branched  $\Gamma$ -rough path  $\mathbf{X}$  is a function  $\mathbf{X}: [0, T]^2 \rightarrow G_\Gamma(A)$  satisfying the following three conditions:

- 1)  $\mathbf{X}_{tt} = \mathbf{1}^*$  for every  $t \in [0, T]$ ;
- 2) for every  $(s, t, u) \in [0, T]^3$ ,  $\mathbf{X}_{st} = \mathbf{X}_{su} * \mathbf{X}_{ut}$
- 3) for every forest  $\sigma \in \mathcal{F}^\Gamma(A)$ ,  $\sup_{s \neq t} |\langle \mathbf{X}_{st}, \sigma \rangle| / |t - s|^{\hat{\gamma}|\sigma|_\Gamma} < +\infty$ .

The Definitions 3.2.5 and 3.2.6 are very similar to the definition of  $\mathcal{R}^\Gamma(A)$  but they are not exactly the same because they are defined in principle on two different groups. In general the projection  $\pi_\Gamma$  of an element  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  provides immediately a  $\Gamma$ -rough path. A non trivial theorem shows that the viceversa is also true.

**Theorem 3.2.7** (Extension Theorem). *Let  $\tilde{\mathbf{X}}$  be a geometric or a branched  $\Gamma$ -rough path. Then there exists a unique  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  such that  $\pi_\Gamma \tilde{\mathbf{X}} = \mathbf{X}$ .*

*Proof.* Let us adopt the same notation  $H$  for the spaces  $T((A))$  or  $\mathcal{H}((A))$ ,  $G$  for the groups  $\mathbb{G}(A)$  or  $\mathcal{H}(A)$  and  $\mathcal{B}$  for  $W(A)$  or  $\mathcal{F}(A)$ . In case when  $\Gamma$  is a constant sequence this result was already partially proven in [Lyo98] for the geometric case and in [Gub10], [Boe18] for the branched case by showing the existence of a function

$\mathbf{X}: [0, T]^2 \rightarrow H$  satisfying the Chen's relations (2g) (2b) and the conditions (3g) (3b). The last properties (1g) (1b) are then obtained following [CEFMMK18, Thm. 4.1]. Passing to a weighted alphabet  $(A, \Gamma)$ , the properties of the function  $|\cdot|_\Gamma$  allow to apply the same proof of [CEFMMK18, Thm. 4.1] to obtain for any  $\sigma \in \mathcal{B}$  such that  $|\sigma|_\Gamma > \hat{\gamma}^{-1}$  a unique function  $f_\sigma: [0, T]^2 \rightarrow \mathbb{R}$  by means of the sewing lemma (see the Theorem 3.3.4) such that the function  $\mathbf{X}: [0, T]^2 \rightarrow H$  given by

$$\langle \mathbf{X}_{st}, \sigma \rangle := \begin{cases} \langle \tilde{\mathbf{X}}_{st}, \sigma \rangle & \text{if } \sigma \in W^\Gamma(A) \text{ or } \mathcal{F}^\Gamma(A) \\ f_\sigma(s, t) & \text{if } \sigma \in W_\Gamma(A) \text{ or } \mathcal{F}_\Gamma(A) \end{cases}$$

belong to  $G$ . Another consequence of the sewing lemma implies that there exists a constant  $c_\Gamma > 0$  depending on  $\Gamma$  such that one has for all  $\sigma \in \mathcal{B}$

$$|\langle \mathbb{X}_{st}, \sigma \rangle| \leq c_\Gamma^{|\sigma|} q_{\hat{\gamma}}(\sigma) |t - s|^{\hat{\gamma}|\sigma|_\Gamma} \tag{3.2.12}$$

where the function  $q_{\hat{\gamma}}(\sigma)$  is defined as follows: for any letter  $a$  or any tree  $\bullet_a$  we set  $q_{\hat{\gamma}}(a) = q_{\hat{\gamma}}(\bullet_a) = 1$  and for any other  $\sigma \in \mathcal{B}$   $q_{\hat{\gamma}}(\sigma)$  is obtained recursively by the identity

$$q_{\hat{\gamma}}(\sigma) := \frac{1}{2^{\hat{\gamma}|\sigma|} - 2} q_{\hat{\gamma}}(\sigma^{(1)}) q_{\hat{\gamma}}(\sigma^{(2)}). \tag{3.2.13}$$

Thereby obtaining the thesis. □

*Remark 3.2.8.* Thanks to this theorem we can easily identify the notion of  $\Gamma$ -rough path with the the notion of rough path of regularity  $\Gamma$  adopting the same notation  $\mathcal{R}^\Gamma(A)$  for both objects. We recall that there exists in the literature another equivalent definition of a rough path. Indeed writing an element  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  as a path with values in  $\mathbb{G}_\Gamma(A)$  or  $G_\Gamma(A)$ , it is possible to rewrite also the analytic conditions in the definitions 3.2.5 and 3.2.6 in an intrinsic way. It turns out that the groups  $\mathbb{G}_\Gamma(A)$  and  $G_\Gamma(A)$  are actually *Lie Groups* (more precisely  $\mathbb{G}_\Gamma(A)$  is a *Carnot group* and when  $\Gamma$  is constant  $G_\Gamma(A)$  is a *Homogeneous group*). These geometric properties allow the existence of an intrinsic metric  $d$  and the analytic conditions in the Definitions 3.2.5 and 3.2.6 are equivalent to state that the path  $\mathcal{X}$  is really a  $\hat{\gamma}$ -Hölder path with respect the metric  $d$ . For more details see [TZ18].

In order to prove that the notion of a rough path is not useless we recall the Lyons-Victoir extension Theorem (see [LV07]). This result gives a sufficient condition to prove in an extremely general setting the existence of a geometric rough path  $\mathbb{X} \in$  over a given path  $X \in \mathcal{C}^\Gamma(A)$ . We state this result as in [TZ18, Thm. 5.7], where it is formulated and proved for a generic weighted alphabet  $(A, \Gamma)$ .

**Theorem 3.2.9** (Lyons-Victoir extension Theorem). *Supposing  $1 \notin \sum_{a \in A} \gamma_a \mathbb{N}$ , for any  $X \in \mathcal{C}^\Gamma(A)$  there exists  $\mathbb{X} \in \mathcal{R}_g^\Gamma(A)$  over  $X$ .*

*Remark 3.2.10.* Differently from the Theorem 3.2.7 the rough path  $\mathbb{X}$  given by the Theorem 3.2.9 is clearly not unique but it is obtained trough an explicit construction.

### 3.3 Controlled rough paths

As the notion of rough path generalises the iterated integrals of a driving continuous path  $X$ , in order to state a change of variable formula we recall a class of paths  $Y$  where we can give a rigorous meaning to the notion of *integration* with respect to  $X$  and *the composition*  $\varphi(Y)$  with respect to a smooth function. Known in the literature as *controlled rough paths*, (see [Gub04] for its first formulation) we introduce them in the case of a weighted alphabet  $(A, \Gamma)$ .

#### 3.3.1 General definitions

The main idea behind their definition is that a controlled rough path should be a path  $Y: [0, T] \rightarrow \mathbb{R}$  where we impose some additional Taylor's formula relations between  $Y$  and the iterated integrals of  $X^a$ , making them dependent on an underlying rough path  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$ . To recall its definition we introduce the following sets of words and forests for any integer  $m \geq 1$

$$W_m^\Gamma(A) := \{v \in W^\Gamma(A) : [|v|_\Gamma] < m\}, \quad \mathcal{F}_m^\Gamma(A) := \{\sigma \in \mathcal{F}(A) : [|\sigma|_\Gamma] < m\}, \quad (3.3.1)$$

and the corresponding subspaces  $T_m^\Gamma(A)$  and  $\mathcal{H}_m^\Gamma(A)$  generated by these sets in  $T(A)$  and  $\mathcal{H}(A)$ . We recall that by construction  $T_m^\Gamma(A) \subset T^\Gamma(A)$  and  $\mathcal{H}_m^\Gamma(A) \subset \mathcal{H}^\Gamma(A)$  and denoting  $N = \lfloor \hat{\gamma}^{-1} \rfloor$  we have the equality  $T_{N+1}^\Gamma(A) = T^\Gamma(A)$ ,  $\mathcal{H}_{N+1}^\Gamma(A) = \mathcal{H}^\Gamma(A)$ .

**Definition 3.3.1.** Let  $\mathbb{X} \in \mathcal{R}_g^\Gamma(A)$ . An  $\mathbb{X}$ -controlled rough path of regularity  $m \geq 1$  is a path  $\mathbb{Y}: [0, T] \rightarrow T_m^\Gamma(A)$  such that for every dual word  $w \in W_m^\Gamma(A)$  one has

$$\sup_{s \neq t} \frac{|\langle w, \mathbb{Y}_t \rangle - \langle \mathbb{X}_{st} \otimes w, \mathbb{Y}_s \rangle|}{|t - s|^{(m - |w|_\Gamma)\hat{\gamma}}} < +\infty. \quad (3.3.2)$$

If  $\langle \mathbf{1}^*, \mathbb{Y}_t \rangle = Y_t$  we say that  $\mathbb{Y}$  is a controlled rough path over  $Y$ .

**Definition 3.3.2.** Let  $\mathbf{X} \in \mathcal{R}_b^\Gamma(A)$ . An  $\mathbf{X}$ -controlled rough path of regularity  $m \geq 1$  is a path  $\mathbf{Y}: [0, T] \rightarrow \mathcal{H}_m^\Gamma(A)$  such that for every dual forest  $\sigma \in \mathcal{F}_m^\Gamma(A)$  one has

$$\sup_{s \neq t} \frac{|\langle \sigma, \mathbf{Y}_t \rangle - \langle \mathbf{X}_{st} * \sigma, \mathbf{Y}_s \rangle|}{|t - s|^{(m - |\sigma|_\Gamma)\hat{\gamma}}} < +\infty. \quad (3.3.3)$$

If  $\langle \mathbf{1}^*, \mathbf{Y}_t \rangle = Y_t$  we say that  $\mathbf{Y}$  is a controlled rough path over  $Y$ .

For any  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  we denote by  $\mathcal{D}^{m\Gamma}(\mathbf{X})$  the set of all  $\mathbf{X}$ -controlled rough path of regularity  $m$ . Since  $\mathcal{D}^{m\Gamma}(\mathbf{X})$  is a vector space and it generalises the notion of a real valued path, we can easily define vectors of controlled rough path using product spaces.

*Remark 3.3.3.* The conditions (3.3.3) and (3.3.2) evaluated on  $\mathbf{1}^*$  mean that the function  $Y$  satisfies the generalised Taylor expansion

$$Y_t = Y_s + \sum_{1 \leq [|w|_\Gamma] < m} \mathbf{Y}_s^w \langle \mathbf{X}_{st}, w \rangle + O(|t - s|^{m\hat{\gamma}}),$$



for some functions  $\mathbf{Y}_t^w$ . These functions are known in the literature as the Gubinelli derivatives of  $Y$  (see [FH14, Chap. 4]) and an informal way to define a controlled rough path is given by a path  $Y$  and all its possible Gubinelli derivatives up to a certain order. The other estimates in the Definition 3.3.1 and the Definition 3.3.2 encode a similar development for the functions  $\mathbf{Y}_t^w$  but with a remainder of order  $O(|t-s|^{(m-|\sigma|_\Gamma)\hat{\gamma}})$ . An equivalent way to restate the bounds (3.3.2) and (3.3.3) can be given using the notion of *modelled distribution* developed in [Hai14]. In order to obtain such description in the branched and in the geometric setting we adopt the notation  $\mathcal{V}_m$  and  $V_m$  for the codomain of a controlled rough path and its basis. Then we can easily define the two parameter family of linear maps  $M_{st}^{\mathbf{X}}: \mathcal{V}_m \rightarrow \mathcal{V}_m$  defined for any  $v \in V_m$  by means of the Sweedler notation as

$$M_{st}^{\mathbf{X}}w := \langle \mathbf{X}_{st}, w^{(1)} \rangle w^{(2)}.$$

Since the set  $V_m$  is finite, we can easily show that the conditions (3.3.2) and (3.3.3) are equivalent to

$$\sup_{\sigma \in V_m} \sup_{s \neq t} \frac{|\langle \sigma, \mathbf{Y}_t - M_{st}^{\mathbf{X}} \mathbf{Y}_s \rangle|}{|t-s|^{(m-|\sigma|_\Gamma)\hat{\gamma}}} < +\infty. \quad (3.3.4)$$

For further details on the link between controlled rough path and regularity structures in a simplified setting see [FH14].

The study of all possible controlled rough paths over a fixed rough path  $\mathbf{X}$  or the existence and uniqueness of a controlled rough path above a fixed path  $Y: [0, T] \rightarrow \mathbb{R}$  might be a very technical topic (see e.g. [FH14, Chap. 6]). However for any  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  over a path  $X \in \mathcal{C}^\Gamma(A)$  there exists always a trivial  $\mathbf{X}$ -controlled rough path, the *coordinate path of  $\mathbf{X}$* . In the case when  $\mathbf{X}$  is a geometric rough path we define it as the function  $\mathcal{X}: [0, T] \rightarrow (T_m^\Gamma(A))^A$  given by

$$\langle k, (\mathcal{X}_t)^a \rangle = \begin{cases} X_t^a & \text{if } k = \mathbf{1}^*, \\ 1 & \text{if } k = a \\ 0 & \text{for any other } k \in W_m^\Gamma(A), \end{cases} \quad (3.3.5)$$

for every component  $a \in A$ . In the branched case we define it similarly as a function  $\mathcal{X}': [0, T] \rightarrow (\mathcal{H}_m^\Gamma(A))^A$  such that

$$\langle \sigma, (\mathcal{X}'_t)^a \rangle = \begin{cases} X_t^a & \text{if } \sigma = \mathbf{1}^*, \\ 1 & \text{if } \sigma = \bullet_a, \\ 0 & \text{for any other } \sigma \in \mathcal{F}_m^\Gamma(A), \end{cases} \quad (3.3.6)$$

Applying the basic definitions of the products  $*$  and  $\otimes$  we obtain

$$\langle w, (\mathcal{X}_t)^a \rangle - \langle \mathbf{X}_{st} \otimes w, (\mathcal{X}_s)^a \rangle = 0, \quad \langle \sigma, (\mathcal{X}'_t)^a \rangle - \langle \mathbf{X}_{st} * \sigma, (\mathcal{X}'_s)^a \rangle = 0,$$

for any  $a \in A$  and for all words  $w \in \mathcal{F}_m^\Gamma(A)$  and forests  $\sigma \in \mathcal{F}_m^\Gamma(A)$ . Therefore  $\mathcal{X}$  and  $\mathcal{X}'$  are two controlled rough paths for any  $m \geq 1$ . In order to simplify the notation, we will denote them by the shorthand notation  $X \in (\mathcal{D}^{m\Gamma}(\mathbf{X}))^A$ .

### 3.3.2 Integration and composition

Let us explain how the notion of controlled rough path allows to give a deterministic theory of integration of one path  $Y: [0, T] \rightarrow \mathbb{R}$  with respect to the increments of  $X^a: [0, T] \rightarrow \mathbb{R}$  for some  $X \in C^\Gamma(A)$  and  $a \in A$ .

In general for any given sequence of partitions  $\mathcal{P}_n$  of  $[0, T]$ , represented as  $\mathcal{P}_n = \{0 = t_0^n < \dots < t_{p(n)}^n = T\}$  such that its mesh size  $|\mathcal{P}_n| := \sup_i(t_{i+1}^n - t_i^n)$  converges to zero, we would like to define the integral from a generalisation of the Riemann summation:

$$\int_0^T Y_r dX_r^a := \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} Y_u (X_v^a - X_u^a), \quad (3.3.7)$$

provided that the limit does not depend on the approximating sequence. This strategy is at the basis of *Young's integration* [You36]. A fundamental result of this operation is that the limit (3.3.7) is well defined if and only if  $Y$  is  $\beta$ -Hölder and  $\beta + \gamma_a > 1$ . However when  $2\gamma_a < 1$  this characterisation prevents us from considering integrand functions of the form  $\varphi(X_t^a)$  for some smooth function  $\varphi$  and then any possible formulation of change of variable formula is allowed. In order to explain the exact way to formulate an integration operator we recall the *Sewing Lemma* (see [Gub04, Prop. 1]). This fundamental theorem provides a general criterion to study the behaviour as  $|\mathcal{P}_n| \rightarrow 0$  of the following sums

$$S_{\mathcal{P}_n} := \sum_{[u,v] \in \mathcal{P}_n} \mathcal{Z}_{uv}$$

for a fixed continuous function  $\mathcal{Z}: [0, T]^2 \rightarrow \mathbb{R}$ . We formulate the theorem in the same way as [CEFMMK18, Prop. 2.1], referring the reader to the same citation for the proof.

**Theorem 3.3.4** (Sewing Lemma). *Let  $\mathcal{Z}: [0, T]^2 \rightarrow \mathbb{R}$  be a continuous function and  $\varepsilon > 0$ . Suppose that there exists a collection of real numbers  $a_i, b_i \geq 0$  for  $i = 1, \dots, n$  and  $\gamma \in (0, 1)$  such that  $a_i + b_i = 1 + \varepsilon$  and*

$$|\mathcal{Z}_{st} - \mathcal{Z}_{su} - \mathcal{Z}_{ut}| \leq \sum_{i=1}^n C_i |t - u|^{a_i} |u - s|^{b_i} \quad |\mathcal{Z}_{st}| \leq C_0 |t - s|^\gamma \quad (3.3.8)$$

for some positive constants  $\{C_i\}_{i=1, \dots, n} \geq 0$  and  $C_0 \geq 0$  and uniformly on  $s, t, u \in [0, T]$  with  $s \leq u \leq t$  or  $t \leq u \leq s$ . Then there exists a unique  $\gamma$ -Hölder function  $Z: [0, T] \rightarrow \mathbb{R}$  such that  $Z_0 = 0$  and for any  $s, t \in [0, T]$

$$|Z_t - Z_s - \mathcal{Z}_{st}| \leq C |t - s|^{1+\varepsilon},$$

where the constant  $C = (2^{1+\varepsilon} - 2)^{-1} \sum_{i=1}^n C_i$ . Moreover for any sequence of partitions  $\mathcal{P}_n$  of  $[s, t]$  such that  $|\mathcal{P}_n| \rightarrow 0$  it follows that

$$Z_t - Z_s = \lim_{n \rightarrow +\infty} \sum_{[u,v] \in \mathcal{P}_n} \mathcal{Z}_{uv}, \quad (3.3.9)$$

independently on the approximating sequence.

*Remark 3.3.5.* The convergence of the series in (3.3.9) holds uniformly on  $t \in [0, T]$ . The path  $Z$  in (3.3.9) depends clearly on the function  $\mathcal{Z}_{st}$  but this dependency holds only up to a correction of order  $o(|t - s|)$ . Indeed for any other continuous function  $\mathcal{Z}'_{st}$  such that

$$|\mathcal{Z}_{st} - \mathcal{Z}'_{st}| \leq C|t - s|^\delta, \quad (3.3.10)$$

for some  $\delta > 1$  and  $C > 0$ , there exists  $C' > 0$  such that the function  $\mathcal{Z}'$  satisfies

$$|\mathcal{Z}'_{st} - \mathcal{Z}'_{su} - \mathcal{Z}'_{ut}| \leq |\mathcal{Z}_{st} - \mathcal{Z}_{su} - \mathcal{Z}_{ut}| + C'|t - u|^\delta + C'|u - s|^\delta.$$

Therefore the function  $\mathcal{Z}'$  satisfies again the hypothesis of the Sewing Lemma and there exists a continuous function  $Z': [0, T] \rightarrow \mathbb{R}$ ,  $Z'_0 = 0$  and  $C'' > 0$  such that

$$|Z'_t - Z'_s - \mathcal{Z}'_{st}| \leq C''|t - s|^{\varepsilon+1}. \quad (3.3.11)$$

Combining the conditions (3.3.11) and (3.3.10) we deduce that

$$|(Z_t - Z_s) - (Z'_t - Z'_s)| = o(|t - s|)$$

and the paths  $Z$  and  $Z'$  must coincide because  $Z'_0 = Z_0 = 0$ .

Combining the sewing lemma with the definition of a controlled rough path we obtain the so called *rough integral*.

**Proposition 3.3.6.** *Let  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  be a fixed rough path,  $a \in A$  and  $N = \lfloor \hat{\gamma}^{-1} \rfloor$ . For any  $\mathbf{Y} \in \mathcal{D}^{(N - \lfloor |a|_\Gamma| + 1)\Gamma}(\mathbf{X})$  and any partition  $\mathcal{P}_n$  of  $[0, t]$  such that  $|\mathcal{P}_n| \rightarrow 0$  the following sums*

$$\sum_{[u,v] \in \mathcal{P}_n} \sum_{w \in W_a^\Gamma(A)} \langle w, \mathbf{Y}_u \rangle \langle \mathbf{X}_{uv}, w a \rangle \quad \text{or} \quad \sum_{[u,v] \in \mathcal{P}_n} \sum_{\sigma \in \mathcal{F}_a^\Gamma(A)} \langle \sigma, \mathbf{Y}_u \rangle \langle \mathbf{X}_{uv}, B_a^+(\sigma) \rangle,$$

where we introduce the finite sets  $W_a^\Gamma(A) := \{w \in W^\Gamma(A) : \lfloor |w|_\Gamma \rfloor \leq N - \lfloor |a|_\Gamma \rfloor\}$ ,  $\mathcal{F}_a^\Gamma(A) := \{\sigma \in \mathcal{F}(A) : \lfloor |\sigma|_\Gamma \rfloor \leq N - \lfloor |\bullet_a|_\Gamma \rfloor\}$  converge to a limit independently on the sequence of partitions  $\mathcal{P}_n$ . We call this limit rough integral and we denote it by the symbol

$$\int_0^t \mathbf{Y}_r d\mathbf{X}_r^a.$$

The rough integral is a  $\gamma_a$ -Hölder function whose value in 0 is 0 and there exists  $\varepsilon > 0$  such that for any  $t, s \in [0, T]$ , the increments

$$\int_s^t \mathbf{Y}_r d\mathbf{X}_r^a := \int_0^t \mathbf{Y}_r d\mathbf{X}_r^a - \int_0^s \mathbf{Y}_r d\mathbf{X}_r^a$$

satisfy the relations

$$\int_s^t \mathbf{Y}_r d\mathbf{X}_r^a = \begin{aligned} & \sum_{w \in W_a^\Gamma(A)} \langle w, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, w a \rangle + O(|t - s|^{1+\varepsilon}) \\ & \text{or} \\ & \sum_{\sigma \in \mathcal{F}_a^\Gamma(A)} \langle \sigma, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, B_a^+(\sigma) \rangle + O(|t - s|^{1+\varepsilon}) \end{aligned} \quad (3.3.12)$$

*Proof.* Depending on the geometric or branched setting, the result is obtained by checking that the functions

$$\mathcal{Z}_{st}^g := \sum_{w \in W_a^\Gamma(A)} \langle w, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, w a \rangle, \quad \mathcal{Z}_{st}^b := \sum_{\sigma \in \mathcal{F}_a^\Gamma(A)} \langle \sigma, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, B_a^+(\sigma) \rangle$$

satisfy respectively the hypotheses of the Sewing lemma. We will prove firstly the geometric case. For any fixed  $s \leq u \leq t$  one has by definition.

$$\mathcal{Z}_{st}^g - \mathcal{Z}_{su}^g - \mathcal{Z}_{ut}^g = \sum_{w \in W_a^\Gamma(A)} \langle w, \mathbf{Y}_s \rangle (\langle \mathbf{X}_{st}, w a \rangle - \langle \mathbf{X}_{su}, w a \rangle) - \langle w, \mathbf{Y}_u \rangle \langle \mathbf{X}_{ut}, w a \rangle.$$

Let us fix a word  $w \in W_a^\Gamma(A)$ . By definition of  $\nabla$  and using the geometric Chen's property of  $\mathbf{X}$  we have

$$\begin{aligned} \langle \mathbf{X}_{st}, w a \rangle - \langle \mathbf{X}_{su}, w a \rangle &= \langle \mathbf{X}_{su} \tilde{\otimes} \mathbf{X}_{ut}, \nabla(w a) \rangle - \langle \mathbf{X}_{su}, w a \rangle \\ &= \langle \mathbf{X}_{su} \tilde{\otimes} \mathbf{X}_{ut}, w a \tilde{\otimes} \mathbf{1} + w^{(1)} \tilde{\otimes} w^{(2)} a \rangle - \langle \mathbf{X}_{su}, w a \rangle \\ &= \langle \mathbf{X}_{su}, w^{(1)} \rangle \langle \mathbf{X}_{ut}, w^{(2)} a \rangle. \end{aligned}$$

Therefore we can write

$$\mathcal{Z}_{st}^g - \mathcal{Z}_{su}^g - \mathcal{Z}_{ut}^g = \sum_{w \in W_a^\Gamma(A)} \langle w, \mathbf{Y}_s \rangle \langle \mathbf{X}_{su}, w^{(1)} \rangle \langle \mathbf{X}_{ut}, w^{(2)} a \rangle - \langle w, \mathbf{Y}_u \rangle \langle \mathbf{X}_{ut}, w a \rangle. \quad (3.3.13)$$

In order to understand the right hand side of (3.3.13) we split the sum in two terms. Writing on the same dual basis the first term becomes

$$\begin{aligned} \sum_{w \in W_a^\Gamma(A)} \langle w, \mathbf{Y}_s \rangle \langle \mathbf{X}_{su}, w^{(1)} \rangle \langle \mathbf{X}_{ut}, w^{(2)} a \rangle &= \sum_{k, w \in W_a^\Gamma(A)} \langle w, \mathbf{Y}_s \rangle \langle \mathbf{X}_{st}, w^{(1)} \rangle \langle k, w^{(2)} \rangle \langle \mathbf{X}_{ut}, k a \rangle \\ &= \sum_{w \in W_a^\Gamma(A)} \langle w, \mathbf{Y}_s \rangle \langle \mathbf{X}_{su} \otimes k, w \rangle \langle \mathbf{X}_{ut}, k a \rangle \\ &= \sum_{k \in W_a^\Gamma(A)} \langle \mathbf{X}_{st} \otimes k, \mathbf{Y}_s \rangle \langle \mathbf{X}_{ut}, k a \rangle. \end{aligned}$$

Since the index in the summation is mute, we can finally apply the bound (3.3.2) and the Hölder property of a rough path in (3.2.10) to obtain the existence of a sequence of positive constants  $\{C_w\}_{w \in W_a^\Gamma(A)}$  such that

$$\begin{aligned} |\mathcal{Z}_{st}^g - \mathcal{Z}_{su}^g - \mathcal{Z}_{ut}^g| &\leq \sum_{w \in W_a^\Gamma(A)} |(\langle \mathbf{X}_{st} \otimes w, \mathbf{Y}_s \rangle - \langle w, \mathbf{Y}_u \rangle) \langle \mathbf{X}_{ut}, w a \rangle| \\ &\leq \sum_{w \in W_a^\Gamma(A)} C_w |u - s|^{(N+1 - \lfloor |a|_\Gamma \rfloor - |w|_\Gamma) \hat{\gamma}} |t - u|^{(|w|_\Gamma + |a|_\Gamma) \hat{\gamma}}. \end{aligned}$$

Using the trivial estimation  $(N+1 + |a|_\Gamma - \lfloor |a|_\Gamma \rfloor) \hat{\gamma} \geq (N+1) \hat{\gamma} > 1$ , we can apply the sewing lemma. In case of the function  $\mathcal{Z}^b$  we can repeat exactly the same proof because the coproduct  $\Delta$  satisfies again the relation  $\Delta B_a(\sigma) = B_a(\sigma) \tilde{\otimes} \mathbf{1} + \sigma^{(1)} \tilde{\otimes} B_a(\sigma^{(2)})$  by construction.  $\square$

*Remark 3.3.7.* For any  $\mathbf{Y} \in \mathcal{D}^{m\Gamma}(\mathbf{X})$  we introduce the norms (depending on the nature of the underlying rough path  $\mathbf{X}$ )

$$\|\mathbf{Y}\|_{m\gamma} := \begin{aligned} & |\langle \mathbf{1}^*, \mathbf{Y}_0 \rangle| + \sum_{w \in W_m^\Gamma(A)} \sup_{s \neq t} |\langle w, \mathbf{Y}_t \rangle - \langle \mathbf{X}_{st} \otimes w, \mathbf{Y}_s \rangle| / |t - s|^{(m-|w|_\Gamma)\hat{\gamma}} \\ & \text{or} \\ & |\langle \mathbf{1}^*, \mathbf{Y}_0 \rangle| + \sum_{\sigma \in \mathcal{F}_m^\Gamma(A)} \sup_{s \neq t} |\langle \sigma, \mathbf{Y}_t \rangle - \langle \mathbf{X}_{st} \otimes \sigma, \mathbf{Y}_s \rangle| / |t - s|^{(m-|\sigma|_\Gamma)\hat{\gamma}}. \end{aligned}$$

The couple  $(\mathcal{D}^{m\Gamma}(\mathbf{X}), \|\cdot\|_{m\gamma})$  becomes then a Banach space and for any fixed  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$ ,  $a \in A$  the integration map  $\mathbf{Y} \rightarrow \int \mathbf{Y} d\mathbf{X}^a$  becomes a linear continuous map from to the space  $\mathcal{D}^{(N-|a|_\Gamma+1)\Gamma}(\mathbf{X})$  where  $N = \lfloor \hat{\gamma}^{-1} \rfloor$  to  $\mathcal{C}\gamma_a(\mathbb{R})$ , as a consequence of a quantitative estimation of the sewing lemma. More generally if  $\mathbf{Y} \in \mathcal{D}^{(N-|a|_\Gamma+1)\Gamma}(\mathbf{X})$  we can introduce the path  $\mathcal{I}_a^{\mathbf{X}}(\mathbf{Y})_t$  defined by duality as

$$\langle k, \mathcal{I}_a^{\mathbf{X}}(\mathbf{Y})_t \rangle = \begin{cases} \int_0^t \mathbf{Y}_r d\mathbf{X}_r^a & \text{if } k = \mathbf{1}^*, \\ \langle w, \mathbf{Y}_t \rangle & \text{for any } k \in W^\Gamma(A), \mathcal{F}^\Gamma(A) \text{ such that } k = wa \text{ or } B_a^+(k), \\ 0 & \text{for any other } w \in W^\Gamma(A) \text{ or } \mathcal{F}^\Gamma(A). \end{cases}$$

Using the estimate (3.3.12) and the properties of  $\mathbf{Y}$ , we can easily check that the map  $\mathbf{Y} \rightarrow \mathcal{I}_a^{\mathbf{X}}(\mathbf{Y})$  is a bounded linear map  $\mathcal{I}_a^{\mathbf{X}}: \mathcal{D}^{(N-|a|_\Gamma+1)\Gamma}(\mathbf{X}) \rightarrow \mathcal{D}^{(N+1)\Gamma}(\mathbf{X})$ . In this way we can encode the relation (3.3.12) in a higher order algebraic structure.

The second fundamental operation we can define on controlled rough paths is the composition with a smooth function. That is for any fixed  $\mathbf{Y} \in (\mathcal{D}^{m\Gamma}(\mathbf{X}))^e$  over a path  $Y: [0, T] \rightarrow \mathbb{R}^e$  and a smooth function  $\varphi: \mathbb{R}^e \rightarrow \mathbb{R}$  for some integer  $e \geq 1$ , we want to define  $\mathbf{Y}^\varphi \in \mathcal{D}^{m\Gamma}(\mathbf{X})$  such that  $\langle \mathbf{1}^*, \mathbf{Y}_t^\varphi \rangle = \varphi(Y_t)$ . Using the same notation of the remark 3.3.3 for any  $a \in (\mathcal{V}_m)^e$  there exists a unique decomposition  $a = a_0 \mathbf{1} + a_+$  where  $a_0 \in \mathbb{R}^e$  and  $a_+ = (a_+^1, \dots, a_+^e)$  is an  $e$ -dimensional vector whose components are generated by all words or forests  $\sigma \in \mathcal{V}_m$  such that  $|\sigma|_\Gamma \geq 1$ . Given such a decomposition, we define the *lift* of  $\varphi$  as the function  $\Phi: (\mathcal{V}_m)^e \rightarrow \mathcal{V}_m$

$$\Phi(a) := \sum_{n=0}^m \frac{1}{n!} \sum_{i_1, \dots, i_n=1}^e \partial_{i_1} \cdots \partial_{i_n} \varphi(a_0) (a_+^{i_1} \star \cdots \star a_+^{i_n})_{< m}, \quad (3.3.14a)$$

where  $\star$  is the intrinsic product of  $\mathcal{V}_m$  (the shuffle product or the disjoint union of forests),  $(\cdot)_{< m}$  is the projection on  $\mathcal{V}_m$  and

$$\partial_{i_1} \cdots \partial_{i_n} \varphi(x) = \frac{\partial^{i_1}}{\partial x^{i_1}} \cdots \frac{\partial^{i_n}}{\partial x^{i_n}} \varphi(x)$$

Since the product  $\star$  is commutative and the expression inside the second sum is symmetrical in the choice of  $i_1, \dots, i_n$ , this definition is equivalent to the following

$$\Phi(a) = \sum_k \frac{\partial^k \varphi(a_0)}{k_1! \cdots k_e!} ((a_+^1)^{\star k_1} \star \cdots \star (a_+^e)^{\star k_e})_{< m} \quad (3.3.14b)$$

where the sum is done over all multi-indices  $k \in \mathbb{N}^e$ ,  $k = (k_1, \dots, k_e)$  and

$$\partial^k \varphi(x) := \partial_1^{k_1} \cdots \partial_e^{k_e} \varphi(x) = \frac{\partial^{k_1}}{\partial x^1} \cdots \frac{\partial^{k_e}}{\partial x^e} \varphi(x).$$

(See Section 3.4 for a careful explanation of the equivalence between equalities (3.3.14a) and (3.3.14b)). This operation has then the right behaviour in the controlled rough path setting. In what follows we will denote by  $C_b^k(\mathbb{R}^d, \mathbb{R})$  the set of  $C^k$  functions  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  with all bounded derivatives up to the  $k$ -th order.

**Theorem 3.3.8.** *For any  $\mathbf{Y} \in (\mathcal{D}^{m\Gamma}(\mathbf{X}))^e$  over a path  $Y: [0, T] \rightarrow \mathbb{R}^e$  and any function  $\varphi \in C_b^m(\mathbb{R}^e, \mathbb{R})$  the function  $t \rightarrow \Phi(\mathbf{Y}_t)$  is an  $\mathbf{X}$ -controlled rough path over  $\varphi(Y)$ .*

*Proof.* By construction all the derivatives appearing in  $\Phi(\mathbf{Y}_t)$  are at most of order  $m - 1$ , thus this function is well defined. In order to give a synthetic description of the branched and the geometric case, we keep the notation of the remark 3.3.3. Using the equivalent condition (3.3.4) and the symbol  $\lesssim$  to denote the inequality up to a constant, we want to show for any  $\sigma \in V_m$

$$|\langle \sigma, \Phi(\mathbf{Y}_t) - M_{st}^{\mathbf{X}} \Phi(\mathbf{Y}_s) \rangle| \lesssim |t - s|^{(m-|\sigma|_{\Gamma})\hat{\gamma}}.$$

Writing  $\mathbf{Y}_s = \tilde{\mathbf{Y}}_s + Y_s \mathbf{1}$  for some fixed function  $\tilde{\mathbf{Y}}: [0, T] \rightarrow (\mathcal{V}_m)^e$ ,  $\tilde{\mathbf{Y}} = (\tilde{\mathbf{Y}}^1, \dots, \tilde{\mathbf{Y}}^e)$  and using the compact multi-index notation we have the identity

$$\Phi(\mathbf{Y}_s) = \sum_{k \in \mathbb{N}^e} \frac{\partial^k \varphi(Y_s)}{k!} ((\tilde{\mathbf{Y}}_s)_{< m})^{\star k}.$$

Applying the linear maps  $M_{st}^{\mathbf{X}}$  to  $\Phi(\mathbf{Y}_s)$ , we remark that for any  $\sigma, \sigma' \in V_m$   $M_{st}^{\mathbf{X}}(\sigma \star \sigma') = M_{st}^{\mathbf{X}}(\sigma) \star M_{st}^{\mathbf{X}}(\sigma')$  as a consequence of the character property of  $\mathbf{X}$  and the multiplicative property of the coproduct. Thus there exists a value  $C > 0$  such that

$$M_{st}^{\mathbf{X}} \Phi(\mathbf{Y}_s) = \sum_{|k| \leq C} \frac{\partial^k \varphi(Y_s)}{k!} (M_{st}^{\mathbf{X}}(\tilde{\mathbf{Y}}_s))^{\star k}, \quad (3.3.15)$$

where  $|k| = \sum_{i=1}^e k_i$ . Since  $M_{st}^{\mathbf{X}} \mathbf{1} = \mathbf{1}$  and  $\mathbf{Y} \in (\mathcal{D}^{m\Gamma}(\mathbf{X}))^e$  there exists a function  $R: [0, T]^2 \rightarrow (\mathcal{V}_m)^e$ ,  $R = (R^1, \dots, R^e)$  such that

$$M_{st}^{\mathbf{X}}(\tilde{\mathbf{Y}}_s) = M_{st}^{\mathbf{X}}(\mathbf{Y}_s) - Y_s \mathbf{1} = \tilde{\mathbf{Y}}_t + (Y_t - Y_s) \mathbf{1} + R_{st}, \quad \sup_{s \neq t} \frac{|\langle \sigma, R_{st}^i \rangle|}{|t - s|^{(m-|\sigma|_{\Gamma})\hat{\gamma}}} < +\infty,$$

uniformly on  $i = 1, \dots, e$  and  $\sigma \in V_m$ . Plugging this expression in (3.3.15) there exists a function  $\tilde{R}: [0, T]^2 \rightarrow \mathcal{V}_m$  satisfying  $|\langle \sigma, \tilde{R}_{st} \rangle| \lesssim |t - s|^{(m-|\sigma|_{\Gamma})\hat{\gamma}}$  for any  $\sigma \in V_m$  such that

$$M_{st}^{\mathbf{X}} \Phi(\mathbf{Y}_s) = \sum_{|k| \leq C} \frac{\partial^k \varphi(Y_s)}{k!} (\tilde{\mathbf{Y}}_t + (Y_t - Y_s) \mathbf{1})^{\star k} + \tilde{R}_{st}. \quad (3.3.16)$$

Performing a Taylor expansion between  $Y_s = (Y_s^1, \dots, Y_s^e)$  and  $Y_t = (Y_t^1, \dots, Y_t^e)$ , we have

$$\partial^k \varphi(Y_s) = \sum_{|k+l| \leq C} \frac{\partial^{k+l} \varphi(Y_t)}{l!} (Y_s - Y_t)^{\star l} + O(|t - s|^{(m-|k|)\hat{\gamma}}), \quad (3.3.17)$$

as a consequence of the Hölder regularity  $|Y_s^i - Y_t^i| \lesssim |t - s|^{\hat{\gamma}}$  for any  $i = 1, \dots, e$ . Using the same bound we obtain also for any  $\sigma \in V_m$  such that  $|\sigma| \leq |k|$

$$\langle \sigma, (\tilde{\mathbf{Y}}_t^i + (Y_t^i - Y_s^i) \mathbf{1})^{\star k} \rangle \lesssim |t - s|^{(|k| - |\sigma|)\hat{\gamma}} \lesssim |t - s|^{(|k| - |\sigma|)\hat{\gamma}}. \quad (3.3.18)$$

Combining the bounds (3.3.17) and (3.3.18) in (3.3.16) one has

$$\langle \sigma, M_{st}^{\mathbf{X}} \Phi(\mathbf{Y}_s) - \sum_{|k+l| \leq C} \frac{\partial^{k+l} \varphi(Y_t)}{k! l!} (Y_s - Y_t)^{\star l} (\tilde{\mathbf{Y}}_t + (Y_t - Y_s) \mathbf{1})^{\star k} \rangle \lesssim |t - s|^{(m-|\sigma|)\hat{\gamma}}.$$

From this bound we apply the multinomial identity

$$\sum_{k+l=m} \frac{1}{k! l!} (Y_s - Y_t)^{\star l} (\tilde{\mathbf{Y}}_t + (Y_t - Y_s) \mathbf{1})^{\star k} = \frac{\partial^m \varphi(Y_t)}{m!} (\tilde{\mathbf{Y}}_t)^{\star m}$$

and we obtain the thesis.  $\square$

*Remark 3.3.9.* The proof of this result was already given in [Gub10, Lem. 8.4] for the branched rough path case but we decided to give a full account because our definition of  $\Phi(\mathbf{Y}_t)$  does not coincide with the equivalent formulation given in [Gub10] and [HK15] and it requires only a commutative product  $\star$ . This proof is a simple rewriting of the proof of [Hai14, Thm. 4.16] without referring explicitly to the theory of regularity structures. Looking at this operation as a linear map  $\mathbf{Y} \rightarrow \Phi(\mathbf{Y})$  for any fixed  $\mathbf{X}$ , we obtain that this operation is also a bounded operator with respect to the norm  $\|\cdot\|_{m\gamma}$ .

As a direct application of the Theorem 3.3.8, for any fixed  $\mathbf{X} \in \mathcal{R}^\Gamma(A)$  we calculate explicitly  $\Phi(X) \in \mathcal{D}^{m\Gamma}(\mathbf{X})$ , the composition of the coordinate path of  $\mathbf{X}$  in the geometric and in the branched case for any  $m \geq 1$ .

**Proposition 3.3.10.** *Let  $\mathbf{X} \in \mathcal{R}_g^\Gamma(A)$  and  $m \geq 1$ , Then for any  $\varphi \in C_b^m(\mathbb{R}^A, \mathbb{R})$  one has*

$$\Phi(X_t) = \sum_{w \in W_m^\Gamma(A)} \partial_w \varphi(X_t) w, \quad (3.3.19)$$

where for any word  $w \in W_m^\Gamma(A)$  we define

$$\partial_w \varphi(X_t) := \begin{cases} \varphi(X_t) & \text{if } w = \mathbf{1}, \\ \partial_{a_{i_1}} \cdots \partial_{a_{i_n}} \varphi(X_t) & \text{if } w = a_{i_1} \cdots a_{i_n}. \end{cases}$$

*Proof.* In what follows we write  $A = \{a_1, \dots, a_{|A|}\}$ . Following the definition of  $\Phi$  in (3.3.14b) in the case of the coordinate path  $X \in (\mathcal{D}^{m\Gamma}(\mathbf{X}))^A$  defined in (3.3.5) we obtain

$$\Phi(X_t) = \sum_{k \in N_m^\Gamma(A)} \frac{\partial^k \varphi(X_t)}{k_1! \cdots k_{|A|}!} a_1^{\sqcup k_1} \sqcup \cdots \sqcup a_{|A|}^{\sqcup k_{|A|}},$$

where  $N_m^\Gamma(A) = \{k \in \mathbb{N}^A: 0 \leq [k_1|a_1|_\Gamma + \dots + k_{|A|}|a_{|A|}|_\Gamma] < m\}$ . Using the identity  $a_i^{\sqcup k} = k!a_i^{\otimes k}$ , which holds for every  $i = 1 \dots |A|$  and  $k \geq 0$ , the theorem is proved as long as we can verify the identity

$$\sum_{k \in N_m^\Gamma(A)} \partial^k \varphi(X_t) a_1^{\otimes k_1} \sqcup \dots \sqcup a_{|A|}^{\otimes k_{|A|}} = \sum_{w \in W_m^\Gamma(A)} \partial_w \varphi(X_t) w, \quad (3.3.20)$$

We check this identity by induction on  $|A|$ . If  $|A| = 1$  the result is trivial. Let us prove the inductive hypothesis. Supposing that the result is true for any alphabet  $A'$  such that  $|A'| < |A|$  we use the linear property of  $\sqcup$  to write the left hand side of (3.3.20) as

$$\sum_{l=0}^L \left( \sum_{k \in N_{m-l}^\Gamma(|A| \setminus \{a_{|A|}\})} \partial^k \partial_{a_{|A|}}^l \varphi(X_t) a_1^{\otimes k_1} \sqcup \dots \sqcup a_{|A|-1}^{\otimes k_{|A|-1}} \right) \sqcup a_{|A|}^{\otimes l}$$

for some fixed value  $L \geq 0$  depending on  $m$ . Applying the induction hypothesis this sum becomes

$$\sum_{l=0}^L \left( \sum_{w' \in W_{m-l}^\Gamma(|A| \setminus \{a_{|A|}\})} \partial_{w'} \partial_{a_{|A|}}^l \varphi(X_t) w' \right) \sqcup a_{|A|}^{\otimes l}. \quad (3.3.21)$$

Using the direct definition of shuffle product, the sum (3.3.21) becomes

$$\sum_{l=0}^L \sum_{w'} \partial_{w'} \partial_{a_{|A|}}^l \varphi(X_t) \sum_{\sigma \in \text{Shuf}(w', l)} (w' a_{|A|}^{\otimes l})_\sigma =: \sum_{l, w' \sigma} \partial_{w'} \partial_{a_{|A|}}^l \varphi(X_t) (w' a_{|A|}^{\otimes l})_\sigma,$$

where  $w'$  belongs in the same set as in (3.3.21). Since the triple  $(l, w', \sigma)$  is in a one to one correspondence with the words  $w \in W_m^\Gamma(A)$ , we conclude.  $\square$

*Remark 3.3.11.* In case  $\mathbf{X} \in \mathcal{R}_b^\Gamma(A)$ , using the same notation of  $A$ , we do not have any further simplification and we can only apply the composition formula (3.3.14a) obtaining

$$\begin{aligned} \Phi(X_t) &= \sum_{n=0}^m \frac{1}{n!} \sum_{i_1, \dots, i_n=1}^{|A|} \partial_{a_{i_1}} \dots \partial_{a_{i_n}} \varphi(X_t) (\bullet_{a_{i_1}} \dots \bullet_{a_{i_n}})_{< m} \\ &= \sum_{n=0}^m \sum_{|\bullet_{a_{i_1}} \dots \bullet_{a_{i_n}}|_\Gamma| < m} \frac{1}{n!} \sigma(a_{i_1} \dots a_{i_n}) \partial_{a_{i_1}} \dots \partial_{a_{i_n}} \varphi(X_t) \bullet_{a_{i_1}} \dots \bullet_{a_{i_n}}, \end{aligned} \quad (3.3.22)$$

where  $\sigma(a_{i_1} \dots a_{i_n})$  is the number of distinct anagrams of the word  $a_{i_1} \dots a_{i_n}$ .

## 3.4 Change of variable formulae for rough paths

### 3.4.1 Kelly's change of variable formula

We apply the integration and the composition operations to recall the general change of variable formulae in the rough path setting starting from  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$  for



some  $\gamma \in (0, 1)$ . In what follows we will suppose that there exists some rough path  $\mathbf{X} \in \mathcal{R}^\gamma(\mathbb{R}^d)$  over  $X$ . The geometric setting will be obtained as a simple exercise, whereas to express the formula in the branched setting, we will need the crucial notion of *simple bracket extension* as developed in David Kelly's Phd thesis [Kel12], a branched rough path defined on a weighted alphabet  $(A, \Gamma)$ .

### The geometric case

Since geometric rough paths aim to generalise the notion of iterated integrals for smooth paths, it is reasonable to think that we can prove a classical change of variable formula as if  $X$  were a smooth path.

**Proposition 3.4.1.** *Let  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$  and  $N = \lfloor \gamma^{-1} \rfloor$ . For any geometric rough path  $\mathbf{X} \in \mathcal{R}_g^\gamma(\mathbb{R}^d)$  over  $X$  and any  $\varphi \in C_b^{N+1}(\mathbb{R}^d, \mathbb{R})$  one has*

$$\varphi(X_t) = \varphi(X_s) + \sum_{i=1}^d \int_s^t \partial_i \Phi(X_r) d\mathbf{X}_r^i, \quad (3.4.1)$$

where  $\{\partial_i \Phi\}_{i=1, \dots, d}$  is the lift of  $\partial_i \varphi: \mathbb{R}^d \rightarrow \mathbb{R}$ .

*Proof.* Since  $\partial_i \varphi \in C_b^N(\mathbb{R}^d, \mathbb{R})$ , the controlled rough path  $\partial_i \Phi(X_r) \in \mathcal{D}^{N\gamma}(\mathbf{X})$  as a consequence of the Theorem 3.3.8 and the rough integrals in (3.4.1) are well defined. Using the explicit definition of  $\partial_i \Phi(X_r)$  in (3.3.19) and the development (3.3.12) for the rough integral one has

$$\int_s^t \partial_i \Phi(X_r) d\mathbf{X}_r^i = \sum_{n=0}^{N-1} \sum_{|w|=n} \partial_w \partial_i \varphi(X_s) \langle \mathbf{X}_{st}, wi \rangle + o(|t-s|),$$

where we parametrised the set of words  $W_i^\gamma(\{1, \dots, d\}) = \{w \in W(\{1, \dots, d\}) : |w| \leq N-1\}$  using the word's length. Summing on  $i$  we obtain

$$\begin{aligned} \sum_{i=1}^d \int_s^t \partial_i \Phi(X_r) d\mathbf{X}_r^i &= \sum_{i=1}^d \sum_{n=0}^{N-1} \sum_{|w|=n} \partial_w \partial_i \varphi(X_s) \langle \mathbf{X}_{st}, wi \rangle + o(|t-s|) \\ &= \sum_{n=1}^N \sum_{|w|=n} \partial_w \varphi(X_s) \langle \mathbf{X}_{st}, w \rangle + o(|t-s|). \end{aligned} \quad (3.4.2)$$

On the other hand, interpreting  $X_t$  as a controlled rough path of regularity  $N+1$  the hypothesis of  $\varphi$  implies that  $\Phi(X_t) \in \mathcal{D}^{(N+1)\gamma}(\mathbf{X})$  by means of the Theorem 3.3.8. Thus we can apply the bound (3.3.2) on  $\langle \mathbf{1}^*, \Phi(X_t) \rangle$  to obtain

$$\varphi(X_t) - \varphi(X_s) = \sum_{n=1}^N \sum_{|k|=n} \partial_k \varphi(X_s) \langle \mathbf{X}_{st}, k \rangle + o(|t-s|), \quad (3.4.3)$$

Then subtracting (3.4.2) to both sides of (3.4.3), we obtain that the difference between the left hand side of (3.4.1) and its right side is equal to  $o(|t-s|)$  for any  $t, s$ . Since this difference is the increment of a path which is zero in  $t = s$ , we obtain the identity (3.4.1).  $\square$

*Remark 3.4.2.* Even if this simple proof is a direct consequence of Proposition 3.3.10, we want to stress that these identities hold because we have encoded in  $\mathbf{X}$  all classical integration by parts formulae by means of the shuffle product. Therefore *products of coordinates* might be rewritten as a sum of *iterated integrals*.

### The branched case

In case  $\mathbf{X} \in \mathcal{R}_b^\gamma(\mathbb{R}^d)$ , the different expansion of  $\Phi(X)$  in (3.3.22) does not imply the same result of the Proposition 3.4.1 but it really depends on the parameter  $\gamma$ . Indeed writing the controlled rough path expansions of  $\langle \mathbf{1}, \Phi(X_t) \rangle$  and the expansion (3.3.12) for  $\int_s^t \partial_i \Phi(X_r) d\mathbf{X}_r^i$  one has the following identities modulo  $o(|t-s|)$

$$\begin{aligned} \varphi(X_t) - \varphi(X_s) &= \sum_{n=1}^N \frac{1}{n!} \sum_{i_1, \dots, i_n=1}^d \partial_{i_1} \cdots \partial_{i_n} \varphi(X_t) \langle \mathbf{X}_{st}, \bullet_{i_1} \cdots \bullet_{i_n} \rangle, \quad (3.4.4) \\ \sum_{i=1}^d \int_s^t \partial_i \Phi(X_r) d\mathbf{X}_r^i &= \sum_{n=1}^N \frac{1}{(n-1)!} \sum_{i_1, \dots, i_n=1}^d \partial_{i_1} \cdots \partial_{i_n} \varphi(X_t) \langle \mathbf{X}_{st}, B_+^{i_1}(\bullet_{i_2} \cdots \bullet_{i_n}) \rangle \\ &= \sum_{n=1}^N \frac{1}{n!} \sum_{i_1, \dots, i_n=1}^d \partial_{i_1} \cdots \partial_{i_n} \varphi(X_t) \langle \mathbf{X}_{st}, \sum_{j=1}^n B_+^{i_{(\sigma(j)(1))}}(\bullet_{i_{\sigma(j)(2)}} \cdots \bullet_{i_{\sigma(j)(n)}}) \rangle, \end{aligned}$$

where  $\sigma = (1 \cdots n)$  is the cyclic permutation of  $n$  elements and  $N = \lfloor \gamma^{-1} \rfloor$ . Unfortunately the right hand sides of the above sums only coincide when  $N = 1$ . Therefore if  $N > 1$  one has:

$$\begin{aligned} \varphi(X_t) - \varphi(X_s) - \sum_{i=1}^d \int_s^t \partial^i \Phi(X_r) dX_r^i &= \\ \sum_{n=2}^N \frac{1}{n!} \sum_{i_1, \dots, i_n=1}^d \partial_{i_1 \dots i_n} \varphi(X_t) \langle \mathbf{X}_{st}, \bullet_{i_1} \cdots \bullet_{i_n} - \sum_{j=1}^n B_+^{i_{(\sigma(j)(1))}}(\bullet_{i_{\sigma(j)(2)}} \cdots \bullet_{i_{\sigma(j)(n)}}) \rangle \end{aligned} \quad (3.4.5)$$

obtaining a non trivial remainder in the formula. The key idea of [Kel12] to get rid of this remainder is then to subtract recursively to both sides of (3.4.5) the expansions associated to other rough integrals in order to obtain finally a remainder of order  $o(|t-s|)$  as before. This is possible by *extending* the actual alphabet  $\{1, \dots, d\}$  to a bigger weighted alphabet  $(\mathcal{A}, \Gamma)$  and choosing  $\widehat{\mathbf{X}} \in \mathcal{R}_b^\Gamma(\mathcal{A})$ , a possible extension of  $\mathbf{X}$ , which satisfies an explicit algebraic condition between the new letters and the old ones. We will simply recall the basic definition to state the theorem and we refer to [Kel12] for the heuristics behind these definitions. In order to state them we introduce a weighted alphabet depending on some parameter  $\gamma$

**Definition 3.4.3.** For any  $\gamma \in (0, 1)$  we define the finite alphabet  $\mathcal{A}_N$  as

$$\mathcal{A}_N = \bigsqcup_{n=1}^N \{1, \dots, d\}^n = \{1, \dots, d\} \sqcup \cdots \sqcup \{(i_1 \cdots i_N) : i_1, \dots, i_N \in \{1, \dots, d\}\},$$

where  $N = \lfloor \gamma^{-1} \rfloor$  and the symbols  $(i)$  and  $i$  are identified. The weight  $\Gamma = \{\gamma_a\}_{a \in \mathcal{A}_N}$  is defined for any  $(i_1 \cdots i_m) \in \mathcal{A}_N$  as  $\gamma_{(i_1 \cdots i_m)} := m\gamma$ .

The definition of  $\mathcal{A}_N$  depends in principle on  $\gamma$  but we suppress  $\gamma$  from the notation to make the notation easier. We simply remark that  $\{1, \dots, d\} \subset \mathcal{A}_N$  for any  $\gamma \in (0, 1)$  obtaining an equality if  $1/2 < \gamma < 1$  and  $\mathcal{A}_N = \mathcal{A}_1$ . The weight  $\Gamma$  is directly compatible with the isotropic setting for  $\{1, \dots, d\}$  and trivially  $|(i_1 \cdots i_m)|_\Gamma = m$  for any  $(i_1 \cdots i_m) \in \mathcal{A}_N$ . This alphabet comes also with a natural notion of symmetry.

**Definition 3.4.4.** Two elements of  $a, b \in \mathcal{A}_N$ ,  $a = (i_1 \cdots i_n)$ ,  $b = (j_1 \cdots j_m)$  are said to be *symmetrical* if  $m = n$  and there exists a permutation  $\sigma \in S_n$  such that  $a_\sigma = (i_{\sigma(1)} \cdots i_{\sigma(n)}) = b$ . We denote it by the symbol  $a \sim b$ .

It is straightforward to check that  $\sim$  is an equivalence relation over the set  $\mathcal{A}_N$  and on any subset  $\{1, \dots, d\}^n \subset \mathcal{A}_N$ . Moreover we can easily extend it to an equivalence relation  $\sim$  on  $\mathcal{T}(\mathcal{A}_N)$  and  $\mathcal{F}(\mathcal{A}_N)$  considering the trees and the forests with two symmetrical labelling. Using the letters of the alphabet  $\mathcal{A}_N$ , we can now define a specific linear combination of forests: the *bracket polynomials*.

**Definition 3.4.5.** For any  $(i_1 \cdots i_n) \in \mathcal{A}_N$  we define the forest *the bracket polynomial*  $\langle\langle i_1 i_2 \cdots i_n \rangle\rangle$

$$\langle\langle i_1 i_2 \cdots i_n \rangle\rangle = \bullet_{i_1} \cdots \bullet_{i_n} - \sum_{\{a,b\}=\{i_1, \dots, i_n\}} B_+^{(b_1 \cdots b_k)}(\bullet_{a_1} \cdots \bullet_{a_{n-k}}), \tag{3.4.6}$$

where the sum is done over all ways of splitting the set  $\{i_1, \dots, i_n\} = \{a, b\}$  into two non-empty sets  $a = \{a_1, \dots, a_{n-k}\}$  and  $b = \{b_1, \dots, b_k\}$  for every  $k \geq 1$ .

By definition  $\langle\langle i_1 \rangle\rangle = \bullet_{i_1}$  for any  $i_1 \in \{1, \dots, d\}$ . Thus we can expect that the bracket polynomial  $\langle\langle i_1 i_2 \cdots i_n \rangle\rangle$  is related with the tree  $\bullet_{(i_1 i_2 \cdots i_n)}$ . Let us recall the main properties of the bracket polynomial.

**Proposition 3.4.6.** *The bracket polynomials  $\{\langle\langle i_1 i_2 \cdots i_n \rangle\rangle\}_{(i_1 \cdots i_n) \in \mathcal{A}_N}$  satisfy the following properties:*

- a) For all  $a, b \in \mathcal{A}_N$  such that  $a \sim b$  one has  $\langle\langle a \rangle\rangle = \langle\langle b \rangle\rangle$ .
- b) Let  $\mathbf{B}_+ : \mathcal{H}(\mathbb{R}^d) \widetilde{\otimes} \mathcal{H}(\mathbb{R}^d) \rightarrow \mathcal{H}(\mathcal{A}_N)$  be the unique linear map defined for any forest  $\sigma \in \mathcal{H}(\mathbb{R}^d)$  as

$$\mathbf{B}_+(\sigma \otimes \tau) := \begin{cases} B_+^{(i_1 \cdots i_n)}(\sigma) & \text{if } \tau = \bullet_{i_1} \cdots \bullet_{i_n} \text{ for any } 1 \leq n \leq N \\ 0 & \text{otherwise .} \end{cases}$$

Then we have the identity

$$\langle\langle i_1 \cdots i_n \rangle\rangle := \bullet_{i_1} \cdots \bullet_{i_n} - \mathbf{B}_+ \Delta'(\bullet_{i_1} \cdots \bullet_{i_n}), \tag{3.4.7}$$

where  $\Delta'$  is the reduced coproduct defined in (3.2.5).

c) Applying the coproduct  $\Delta: \mathcal{H}(\mathcal{A}_N) \widetilde{\otimes} \mathcal{H}(\mathcal{A}_N) \rightarrow \mathcal{H}(\mathcal{A}_N)$  to  $\langle\langle i_1 i_2 \cdots i_n \rangle\rangle$  we have

$$\begin{aligned} \Delta \langle\langle i_1 i_2 \cdots i_n \rangle\rangle &= \langle\langle i_1 i_2 \cdots i_n \rangle\rangle \widetilde{\otimes} \mathbf{1} + \mathbf{1} \widetilde{\otimes} \langle\langle i_1 i_2 \cdots i_n \rangle\rangle \\ &+ \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \bullet_{a_1} \cdots \bullet_{a_{n-k}} \widetilde{\otimes} (\langle\langle b_1 \cdots b_k \rangle\rangle - \bullet_{(b_1 \cdots b_k)}), \end{aligned} \quad (3.4.8)$$

where the sum is done over the same set of symbols as in (3.4.6).

*Proof.* We will prove the three properties a), b) and c) separately. Writing  $a = (i_1 \cdots i_n)$  and  $b = (j_1 \cdots j_m)$  if  $a \sim b$  one has  $m = n$  and

$$\bullet_{i_1} \cdots \bullet_{i_n} = \bullet_{j_1} \cdots \bullet_{j_m},$$

since the forest product is commutative. Moreover we have the equality between the sets  $\{i_1, \dots, i_n\}$  and  $\{j_1, \dots, j_m\}$  implying the identity

$$\sum_{\{a,b\}=\{i_1, \dots, i_n\}} B_+^{(b_1 \cdots b_k)}(\bullet_{a_1} \cdots \bullet_{a_{n-k}}) = \sum_{\{a',b'\}=\{j_1, \dots, j_m\}} B_+^{(b'_1 \cdots b'_k)}(\bullet_{a'_1} \cdots \bullet_{a'_{n-k}}),$$

Therefore  $\langle\langle a \rangle\rangle = \langle\langle b \rangle\rangle$ .

In order to prove b) we can iterate the multiplicative property of  $\Delta$  to obtain

$$\Delta(\bullet_{i_1} \cdots \bullet_{i_n}) = \bullet_{i_1} \cdots \bullet_{i_n} \widetilde{\otimes} \mathbf{1} + \mathbf{1} \widetilde{\otimes} \bullet_{i_1} \cdots \bullet_{i_n} + \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \bullet_{a_1} \cdots \bullet_{a_{n-k}} \widetilde{\otimes} \bullet_{b_1} \cdots \bullet_{b_k}$$

Thus by definition of the reduced coproduct we obtain

$$\Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) = \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \bullet_{a_1} \cdots \bullet_{a_{n-k}} \widetilde{\otimes} \bullet_{b_1} \cdots \bullet_{b_k} \quad (3.4.9)$$

and by definition of the operator  $\mathbf{B}_+$  we obtain the identity (3.4.7).

In case of c), writing  $\Delta\sigma = \sigma \widetilde{\otimes} \mathbf{1} + \mathbf{1} \widetilde{\otimes} \sigma + \Delta'\sigma$  and applying (3.4.7) we obtain

$$\Delta \langle\langle i_1 i_2 \cdots i_n \rangle\rangle = \langle\langle i_1 i_2 \cdots i_n \rangle\rangle \widetilde{\otimes} \mathbf{1} + \mathbf{1} \widetilde{\otimes} \langle\langle i_1 i_2 \cdots i_n \rangle\rangle + \Delta'(\bullet_{i_1} \cdots \bullet_{i_n} - \mathbf{B}_+ \Delta'(\bullet_{i_1} \cdots \bullet_{i_n})).$$

Using the simple identity  $\Delta'(B_+^b(\sigma)) = (\text{id} \widetilde{\otimes} B_+^b) \Delta'\sigma + \sigma \widetilde{\otimes} \bullet_b$  and the coassociativity of the reduced coproduct  $(\Delta' \widetilde{\otimes} \text{id}) \Delta' = (\text{id} \widetilde{\otimes} \Delta') \Delta'$  one has

$$\begin{aligned} \Delta' \mathbf{B}_+ \Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) &= \\ &= (\text{id} \widetilde{\otimes} \mathbf{B}_+)(\Delta' \widetilde{\otimes} \text{id}) \Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) + \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \bullet_{a_1} \cdots \bullet_{a_{n-k}} \widetilde{\otimes} \bullet_{(b_1 \cdots b_k)} \\ &= (\text{id} \widetilde{\otimes} \mathbf{B}_+)(\text{id} \widetilde{\otimes} \Delta') \Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) + \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \bullet_{a_1} \cdots \bullet_{a_{n-k}} \widetilde{\otimes} \bullet_{(b_1 \cdots b_k)}. \end{aligned}$$

Thus using the combinatorial identity (3.4.9) and the alternative definition (3.4.7) we obtain

$$\begin{aligned} \Delta'(\bullet_{i_1} \cdots \bullet_{i_n} - \mathbf{B}_+ \Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) &= \\ &= \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \bullet_{a_1} \cdots \bullet_{a_{n-k}} \widetilde{\otimes} (\bullet_{b_1} \cdots \bullet_{b_k} - \mathbf{B}_+(\Delta'(\bullet_{b_1} \cdots \bullet_{b_k})) - \bullet_{(b_1 \cdots b_k)}) \\ &= \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \bullet_{a_1} \cdots \bullet_{a_{n-k}} \widetilde{\otimes} (\langle\langle b_1 \cdots b_k \rangle\rangle - \bullet_{(b_1 \cdots b_k)}). \end{aligned}$$

Thereby obtaining the thesis.  $\square$

*Remark 3.4.7.* All these properties as well as their proofs were already given in several parts of [Kel12, Chap. 5] and we decided to encode them in a unique proposition to keep the discussion self-contained.

The identity (3.4.8) has a very deep consequence. Indeed for any fixed  $\widehat{\mathbf{X}} \in \mathcal{R}_b^\Gamma(\mathcal{A}_N)$  and any triplet  $s, u, t \in [0, T]$ , the formula (3.4.8) together with the algebraic property of  $\widehat{\mathbf{X}}$  implies

$$\begin{aligned} \langle \widehat{\mathbf{X}}_{st}, \langle\langle i_1 i_2 \cdots i_n \rangle\rangle \rangle &= \langle \widehat{\mathbf{X}}_{su}, \langle\langle i_1 i_2 \cdots i_n \rangle\rangle \rangle + \langle \widehat{\mathbf{X}}_{ut}, \langle\langle i_1 i_2 \cdots i_n \rangle\rangle \rangle \\ &+ \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \langle \widehat{\mathbf{X}}_{su}, \bullet_{a_1} \cdots \bullet_{a_{n-k}} \rangle \langle \widehat{\mathbf{X}}_{ut}, \langle\langle b_1 \cdots b_k \rangle\rangle - \bullet_{(b_1 \cdots b_k)} \rangle. \end{aligned}$$

Therefore the function  $(s, t) \rightarrow \langle \widehat{\mathbf{X}}_{st}, \langle\langle i_1 i_2 \cdots i_n \rangle\rangle \rangle$  may be reinterpreted as the increment of a path as long as  $\langle \widehat{\mathbf{X}}_{st}, \langle\langle b_1 \cdots b_k \rangle\rangle \rangle = \langle \widehat{\mathbf{X}}_{st}, \bullet_{(b_1 \cdots b_k)} \rangle$  for any letter  $(b_1 \cdots b_k)$  with  $k < n$ . This observation justifies the following definition of *bracket extension*.

**Definition 3.4.8.** For any  $\gamma \in (0, 1)$  an element  $\widehat{\mathbf{X}} \in \mathcal{R}_b^\Gamma(\mathcal{A}_N)$  is said to be a bracket extension if it satisfies the following properties:

- 1) for all  $i_1, \dots, i_n \in \{1, \dots, d\}$ ,  $1 \leq n \leq N$

$$\langle \widehat{\mathbf{X}}_{st}, \bullet_{(i_1 \cdots i_n)} \rangle = \langle \widehat{\mathbf{X}}_{st}, \langle\langle i_1 \cdots i_n \rangle\rangle \rangle. \quad (3.4.10)$$

- 2) The rough path  $\widehat{\mathbf{X}}$  is *symmetrical*. That is for any couple of forests  $\sigma, \tau \in \mathcal{H}(\mathcal{A}_N)$  such that  $\tau \sim \sigma$  then  $\langle \widehat{\mathbf{X}}_{st}, \tau \rangle = \langle \widehat{\mathbf{X}}_{st}, \sigma \rangle$ .

*Remark 3.4.9.* For any  $\mathbf{X} \in \mathcal{R}_g^\gamma(\mathbb{R}^d)$  over a path  $X: [0, T] \rightarrow \mathbb{R}^d$  an informal way to describe a bracket extension over  $\mathbf{X}$  is then to add a family of extra paths  $\widehat{X} = \{\widehat{X}^{(i_1 \cdots i_n)}\}_{(i_1 \cdots i_n) \in \mathcal{A}_N}$  with higher Hölder regularity satisfying some symmetry hypothesis and a family of higher order iterated integrals from  $\mathbf{X}$  in order to satisfy the identities

$$\langle \mathbf{X}_{st}, \bullet_{i_1} \cdots \bullet_{i_n} \rangle = \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \langle \widehat{\mathbf{X}}_{st}, B_+^{(b_1 \cdots b_k)}(\bullet_{a_1} \cdots \bullet_{a_{n-k}}) \rangle + \widehat{X}_t^{(i_1 \cdots i_n)} - \widehat{X}_s^{(i_1 \cdots i_n)}, \quad (3.4.11)$$

for any  $i_1, \dots, i_n \in \{1, \dots, d\}$ . In this way we restore partially *the integration by parts* formulae for the coordinates paths of  $\mathbf{X}$ . By analogy with the stochastic calculus setting, we call the path  $\widehat{X}^{(i_1 \cdots i_n)}$  the  $(i_1 \cdots i_n)$ -*variation* of  $X$ .

This notion is then sufficient to get a general change of variable formula for any sufficiently smooth function  $\varphi$ .

**Theorem 3.4.10** (Kelly's change of variable formula). *Let  $\gamma \in (0, 1)$ ,  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$  and  $\mathbf{X} \in \mathcal{R}_b^\gamma(\mathbb{R}^d)$  over  $X$ . Then for any  $\varphi \in C_b^{N+1}(\mathbb{R}^d, \mathbb{R})$  where  $N = \lfloor \gamma^{-1} \rfloor$  and any bracket extension  $\widehat{\mathbf{X}}$  over  $\mathbf{X}$  one has*

$$\varphi(X_t) = \varphi(X_s) + \sum_{i=1}^d \int_s^t \partial_i \Phi(X_r) d\mathbf{X}_r^i + \sum_{n=2}^N \frac{1}{n!} \sum_{i_1, \dots, i_n=1}^d \int_s^t \partial_{i_1} \cdots \partial_{i_n} \Phi(X_r) d\widehat{\mathbf{X}}_r^{(i_1 \cdots i_n)}, \quad (3.4.12)$$

where  $\partial_{i_1} \cdots \partial_{i_n} \Phi$  is the lift of the function  $\partial_{i_1} \cdots \partial_{i_n} \varphi: \mathbb{R}^d \rightarrow \mathbb{R}$ .

*Proof.* Our proof retraces Kelly's original one via a slightly different notation, we repeat here for sake of completeness. As before we will write down all the equalities modulo  $o(|t-s|)$ . For any fixed element  $(i_1 \cdots i_n) \in \mathcal{A}_N$  the function  $\partial_{i_1} \cdots \partial_{i_n} \varphi \in C_b^{N+1-n}(\mathbb{R}^d, \mathbb{R})$  and the rough integral  $\int_s^t \partial_{i_1} \cdots \partial_{i_n} \Phi(X_r) d\widehat{X}_r^{(i_1 \cdots i_n)}$  is well defined. Applying the identity (3.3.12) to it we obtain

$$\int_s^t \partial_{i_1} \cdots \partial_{i_n} \Phi(X_r) d\widehat{X}_r = \sum_{m=0}^{N-n} \sum_{j_1, \dots, j_m=1}^d \frac{\partial_{j_1, \dots, j_m} \partial_{i_1 \cdots i_n} \varphi(X_s)}{m!} \langle \widehat{\mathbf{X}}_{st}, B_+^{(i_1 \cdots i_n)}(\bullet_{j_1} \cdots \bullet_{j_m}) \rangle.$$

This formula holds also when  $n = 1$  and  $(i_1 \cdots i_n) = i_1$  and the left hand side is the rough integral with respect to  $X^{i_1}$ . By summing over all indices  $i_1, \dots, i_n$  and  $n$ , we use the substitution  $k = m + n$  obtaining

$$\begin{aligned} & \sum_{n=1}^N \frac{1}{n!} \sum_{i_1, \dots, i_n=1}^d \int_s^t \partial_{i_1} \cdots \partial_{i_n} \Phi(X_r) d\widehat{X}_r^{(i_1 \cdots i_n)} = \\ & = \sum_{k=1}^N \frac{1}{k!} \sum_{i_1, \dots, i_k=1}^d \partial_{i_1} \cdots \partial_{i_k} \varphi(X_s) \langle \widehat{\mathbf{X}}_{st}, \sum_{n=1}^k \binom{k}{n} B_+^{(i_1 \cdots i_n)}(\bullet_{i_{n+1}} \cdots \bullet_{i_k}) \rangle. \end{aligned} \quad (3.4.13)$$

For any fixed  $k$  and  $i_1, \dots, i_k$ , we consider the sum of trees on the right hand side of (3.4.13). Since this expression does not change in its equivalence class by applying any permutation of  $k$  elements we can write

$$\begin{aligned} & \sum_{n=1}^k \binom{k}{n} B_+^{(i_1 \cdots i_n)}(\bullet_{i_{n+1}} \cdots \bullet_{i_k}) \sim \frac{1}{k!} \sum_{\sigma \in S_k} \sum_{n=1}^k \binom{k}{n} B_+^{(i_{\sigma(1)} \cdots i_{\sigma(n)})}(\bullet_{i_{\sigma(n+1)}} \cdots \bullet_{i_{\sigma(k)}}) \\ & = \sum_{\sigma \in S_k} \sum_{n=1}^{k-1} \frac{1}{(k-n)!n!} B_+^{(i_{\sigma(1)} \cdots i_{\sigma(n)})}(\bullet_{i_{\sigma(n+1)}} \cdots \bullet_{i_{\sigma(k)}}) + \sum_{\sigma \in S_k} \frac{1}{k!} \bullet_{(i_{\sigma(1)} \cdots i_{\sigma(k)})}. \end{aligned} \quad (3.4.14)$$

Using again the symmetry relation  $\bullet_{(i_{\sigma(1)} \cdots i_{\sigma(k)})} \sim \bullet_{(i_1 \cdots i_k)}$  and

$$B_+^{(i_{\sigma(1)} \cdots i_{\sigma(n)})}(\bullet_{i_{\sigma(n+1)}} \cdots \bullet_{i_{\sigma(k)}}) \sim B_+^{(b_1 \cdots b_n)}(\bullet_{a_1} \cdots \bullet_{a_{n-k}}),$$

where  $a = \{a_1, \dots, a_{n-k}\}$ ,  $b = \{b_1, \dots, b_n\}$  is a non empty partition of  $\{i_1, \dots, i_k\}$  into two set, it is straightforward to show that the sum (3.4.14) is then symmetrical with

$$\sum_{\{i_1, \dots, i_k\}=\{a,b\}} B_+^{(b_1 \cdots b_n)}(\bullet_{a_1} \cdots \bullet_{a_{n-k}}) + \bullet_{(i_1 \cdots i_k)},$$

by counting the respective multiplicity of the indexes in the two sums. Therefore using the symmetrical invariance of  $\widehat{\mathbf{X}}$  and the identity (3.4.11) one has

$$\langle \widehat{\mathbf{X}}_{st}, \sum_{n=1}^k \binom{k}{n} B_+^{(i_1 \cdots i_n)}(\bullet_{i_{n+1}} \cdots \bullet_{i_k}) \rangle = \langle \mathbf{X}_{st}, \bullet_{i_1} \cdots \bullet_{i_k} \rangle.$$

Thus the right hand side of (3.4.13) becomes equal to the right hand side of (3.4.4) up to a remainder of order  $o(|t-s|)$ . We conclude finally using the same final argument given in the proof of the Proposition (3.4.1).  $\square$

*Remark 3.4.11.* Looking at the Definition 3.4.8 and the Theorem 3.4.10, it is not obvious to prove apriori the existence of one bracket extension  $\widehat{\mathbf{X}}$  over some  $\mathbf{X} \in \mathcal{R}_b^\gamma(\mathbb{R}^d)$ . However in the simple case  $\gamma \in (1/3, 1/2]$  where  $N = 2$ , the condition (3.4.10) becomes simply

$$\langle \widehat{\mathbf{X}}_{st}, \bullet_{(ij)} \rangle := \langle \widehat{\mathbf{X}}_{st} \bullet_i \bullet_j - \mathfrak{!}_i^j - \mathfrak{!}_j^i \rangle = \langle \mathbf{X}_{st} \bullet_i \bullet_j - \mathfrak{!}_i^j - \mathfrak{!}_j^i \rangle, \quad (3.4.15)$$

for every  $i, j \in \{1, \dots, d\}$ . In this way recalling the fact that the set  $\mathcal{F}^\Gamma(\mathcal{A}_2)$  is given by

$$\mathcal{F}^\Gamma(\mathcal{A}_2) = \{ \bullet_i, \bullet_{(ij)}, \bullet_i \bullet_j, \mathfrak{!}_i^j : i, j \in \{1, \dots, d\} \},$$

we can easily define a function  $\widehat{\mathbf{X}}: [0, T]^2 \rightarrow \mathcal{H}^\Gamma((\mathcal{A}_2))$  satisfying the properties of the Definition 3.2.6 only in terms of the underlying rough path  $\mathbf{X}$  and consequently there exists a unique bracket extension over  $\mathbf{X}$  by means of the extension theorem. The related change of variable formula (3.4.12) in this case will depend only on  $\mathbf{X}$  and it looks formally identical to the classical Itô formula for a semimartingale  $X$ . The case  $N = 2$  was already known in the literature and an equivalent formulation of (3.4.12) in this case was obtained in [FH14, Prop. 5.6] by means of the notion of *reduced rough paths*. When  $N > 2$  the result given in [Kel12, Prop. 5.2.14] shows the existence of a bracket extension but this object is not unique anymore.

### 3.4.2 Symmetrical paths

In order to simplify the Definition 3.4.8, we want to encode immediately the notion of symmetry hypothesis for a bracket extension  $\widehat{\mathbf{X}}$  by changing the underlying alphabet. Indeed we can describe directly the equivalence classes of  $\mathcal{A}_N$  with respect to  $\sim$  in terms of a new alphabet.

**Definition 3.4.12.** We denote by  $\mathbb{A} = \mathbb{N}^d \setminus \{0\}$  the set of  $d$  dimensional non zero multi-index  $\alpha = (\alpha_1, \dots, \alpha_d)$ . For any and  $\gamma \in (0, 1)$  we define the finite alphabet  $\mathbb{A}_N$  as

$$\mathbb{A}_N = \{ \alpha \in \mathbb{A} : |\alpha| \leq N \}$$

where  $N = \lfloor \gamma^{-1} \rfloor$  and  $|\alpha| := \sum_{i=1}^d \alpha_i$ . The weight  $\Gamma = \{ \gamma_\alpha \}_{\alpha \in \mathbb{A}_N}$  is defined for any  $\alpha \in \mathbb{A}_N$  as  $\gamma_\alpha := |\alpha| \gamma$ .

Similarly to the definition of  $\mathcal{A}_N$ , we suppress the parameter  $\gamma$  from the notation of  $\mathbb{A}_N$ . Moreover we keep the same notation  $\Gamma$  for the weight of  $\mathbb{A}_N$ . We note also that  $\{1, \dots, d\} \subset \mathbb{A}_N$  for any  $\gamma \in (0, 1)$  by means of the canonical identification  $i \rightarrow e_i$  where  $\{e_i\}_{i=1, \dots, d}$  is the canonical basis, obtaining an equality if  $1/2 < \gamma < 1$  and  $\mathbb{A}_N = \mathbb{A}_1$ . The weight  $\Gamma$  is directly compatible with the isotropic setting for  $\{1, \dots, d\}$  and trivially  $|\alpha|_\Gamma = |\alpha|$  for any  $\alpha \in \mathbb{A}_N$ . The link between  $\mathcal{A}_N$  and  $\mathbb{A}_N$  is done by means of an explicit function  $\mathcal{S}: \mathcal{A}_N \rightarrow \mathbb{A}_N$ .

**Definition 3.4.13.** Let  $\gamma \in (0, 1)$ . For any  $a \in \mathcal{A}_N$ ,  $a = (i_1 \cdots i_n)$ , we set for any  $j \in \{1, \dots, d\}$  the integers

$$m_j := \#\{l \in \{1, \dots, n\} : i_l = j\}, \quad (3.4.16)$$

and we define  $\mathcal{S}(a) := (m_1, \dots, m_d)$ . We call the map  $\mathcal{S}$  the *symmetrization map*.

For instance if  $d = 3$  and  $\gamma \in (1/4, 1/3]$  one has

$$\mathcal{S}((12)) = (1, 1, 0), \quad \mathcal{S}((21)) = (1, 1, 0), \quad \mathcal{S}((333)) = (0, 0, 3).$$

By definition the symmetrization map is a well defined surjective application and for any  $a \in \mathcal{A}_N$   $|a|_\Gamma = |\mathcal{S}(a)|_\Gamma$  where the function  $|\cdot|_\Gamma$  is associated to the different weights. The function  $\mathcal{S}$  is also a bijection when we pass to the quotient.

**Lemma 3.4.14.** *For any  $\gamma \in (0, 1)$ ,  $\mathcal{S}$  is a bijective map from  $\mathcal{A}_N / \sim$  to  $\mathbb{A}_N$ .*

*Proof.* In order to prove the bijection it is sufficient to show that for every  $1 \leq n \leq N$ ,  $\mathcal{S}$  is a bijection between  $A_n$  and  $A'_n$  where

$$A_n = \{1, \dots, d\}^n / \sim, \quad A'_n = \{v \in \mathbb{A}_N : |v| = n\}.$$

That is for any  $a, b \in \{1, \dots, d\}^n$ , one has  $a \sim b$  if and only if  $\mathcal{S}(a) = \mathcal{S}(b)$ . Thanks to the definition (3.4.16), if  $a \sim b$  then both  $\mathcal{S}(a)$  and  $\mathcal{S}(b)$  belong to  $A'_n$  and their components do not change under permutation of the writing of  $a$ , therefore  $\mathcal{S}(a) = \mathcal{S}(b)$ . On the other hand, if  $\mathcal{S}(a) = \mathcal{S}(b) = (m_1, \dots, m_d)$ , then for any  $j = \{1, \dots, d\}$  both  $a$  and  $b$  contain the same amount of coordinates with the value  $j$ . Rearranging the components of  $a$  and  $b$ , there exist two permutations  $\sigma_a$  and  $\sigma_b$  such that

$$a_{\sigma_a} = b_{\sigma_b} = \left( \underbrace{1 \cdots 1}_{m_1 \text{ times}} \underbrace{2 \cdots 2}_{m_2 \text{ times}} \cdots \underbrace{d \cdots d}_{m_d \text{ times}} \right). \quad (3.4.17)$$

As a conclusion, if we define  $\sigma = \sigma_b^{-1} \circ \sigma_a$ , then  $a_\sigma = b$  by construction.  $\square$

Looking at the proof of the above lemma, we associated to every  $\beta \in \mathbb{A}_N$  a specific element of  $\mathcal{S}^{-1}(\beta)$ , that is the right hand side of the identity (3.4.17). We will refer to this element of  $\mathcal{A}_N$  with the symbol  $\beta'$ . Another trivial consequence of the Lemma 3.4.14 is then a useful identity we already used in the definition of the lift  $\Phi$  of a smooth function  $\varphi$ .

**Proposition 3.4.15.** *If  $c_{i_1 \dots i_n}$  is a sequence of values indexed by  $\{1, \dots, d\}^n$  invariant by permutation then*

$$\sum_{i_1, \dots, i_n=1}^d c_{i_1 \dots i_n} = \sum_{\beta \in \mathbb{A} : |\beta|=n} \frac{n!}{\beta_1! \cdots \beta_d!} c_{\beta'} \quad (3.4.18)$$

*Proof.* Writing the left hand side of (3.4.18) as a sum on  $\{1, \dots, d\}^n \subset \mathcal{A}_N$  for some  $N \geq n$ , and denoting by  $[a]$  the equivalence class of an element  $a \in \{1, \dots, d\}^n$  we use the proof of the Lemma 3.4.14 and invariance by permutation to obtain

$$\sum_{i_1, \dots, i_n=1}^d c_{i_1 \dots i_n} = \sum_{[a] \in \{1, \dots, d\}^n / \sim} \sum_{[a]=a} c_a = \sum_{\beta \in \mathbb{A} : |\beta|=n} \sum_{a \in \mathcal{S}^{-1}(\beta)} c_a = \sum_{\beta \in \mathbb{A} : |\beta|=n} \#\{\mathcal{S}^{-1}(\beta)\} c_{\beta'}.$$

Now using the elementary combinatorial identity

$$\#\{\mathcal{S}^{-1}(\beta)\} = \#\{\text{anagrams of } \underbrace{1 \cdots 1}_{\beta_1 \text{ times}} \underbrace{2 \cdots 2}_{\beta_2 \text{ times}} \cdots \underbrace{d \cdots d}_{\beta_d \text{ times}}\} = \frac{|\beta|!}{\beta_1! \cdots \beta_d!},$$

which holds for any  $\beta \in \mathbb{A}_N$  we conclude.  $\square$



This simple identity allows to justify the equivalence between definitions (3.3.14a) and (3.3.14b). Even if the map  $\mathcal{S}$  is originally defined at the level of the alphabets, by simply applying  $\mathcal{S}$  at each decoration of any forest appearing in  $\mathcal{H}(\mathcal{A}_N)$  and extending it linearly, we can easily define a linear surjective map  $\mathcal{S}: \mathcal{H}(\mathcal{A}_N) \rightarrow \mathcal{H}(\mathbb{A}_N)$  (we will denote it by the same notation). For instance

$$\mathcal{S}(\mathbf{V}_{(i_3 i_4 i_5)}^{i_2}) = \mathbf{V}_{[i_3 i_4 i_5]}^{i_2}, \quad \mathcal{S}(\mathbf{V}_{(i_1 i_2)}^{i_3} \bullet_{(i_4 i_5)} + 2 \bullet_{(i_6 i_7)}) = \mathbf{V}_{[i_1 i_2]}^{i_3} \bullet_{[i_4 i_5]} + 2 \bullet_{[i_6 i_7]}.$$

Since  $\mathcal{S}$  acts only at the level of decoration it is straightforward to verify

$$\mathcal{S}(\tau_1 \tau_2) = \mathcal{S}(\tau_1) \mathcal{S}(\tau_2), \quad (\mathcal{S} \otimes \mathcal{S}) \Delta \tau_1 = \Delta \mathcal{S}(\tau_1), \quad (3.4.19)$$

for any  $\tau_1, \tau_2 \in \mathcal{H}(\mathcal{A}_N)$ . Moreover  $|\mathcal{S}(\sigma)|_\Gamma = |\sigma|_\Gamma$  for any  $\sigma \in \mathcal{F}(\mathcal{A}_N)$  as a consequence of the previous compatibility between the related weights. Therefore the application  $\mathcal{S}$  is a surjective Hopf algebra morphism sending  $\mathcal{H}^\Gamma(\mathcal{A}_N)$  to  $\mathcal{H}^\Gamma(\mathbb{A}_N)$ . By duality we can define its algebraic adjoint  $\mathcal{S}^*: \mathcal{H}(\mathbb{A}_N) \rightarrow \mathcal{H}(\mathcal{A}_N)$ . For any  $\mathbb{X} \in \mathcal{R}_b^\Gamma(\mathbb{A}_N)$  the properties of  $\mathcal{S}$  in (3.4.19) and the related weights imply that the function  $(t, s) \rightarrow \mathcal{S}^* \mathbb{X}_{st}$  is again a branched rough path, we will denote by the symbol  $\mathcal{S}^* \mathbb{X}$ .

**Proposition 3.4.16.** *Let  $\gamma \in (0, 1)$ . Then  $\mathbf{X} \in \mathcal{R}_b^\Gamma(\mathcal{A}_N)$  is symmetrical if and only if there exists a unique  $\mathbb{X} \in \mathcal{R}_b^\Gamma(\mathbb{A}_N)$  such that  $\mathbf{X} = \mathcal{S}^* \mathbb{X}$ .*

*Proof.* Applying the Lemma 3.4.14 at each decoration of every forest in  $\mathcal{H}(\mathcal{A}_N)$  it is trivial to show that for any couple of forests  $\tau, \sigma \in \mathcal{H}(\mathcal{A}_N)$  one has  $\tau \sim \sigma$  if and only if  $\mathcal{S}(\tau) = \mathcal{S}(\sigma)$ . Thus the proposition follows straightforwardly.  $\square$

By means of this proposition we can easily rewrite the Theorem 3.4.10 using  $\mathbb{A}_N$ . For any  $i_1, \dots, i_n \in \{1, \dots, d\}$  we introduce the notion of *symmetrical bracket polynomial*

$$[[i_1 i_2 \dots i_n]] := \mathcal{S}(\langle\langle i_1 i_2 \dots i_n \rangle\rangle). \quad (3.4.20)$$

Thus we can replace the Definition 3.4.8 with a new one.

**Definition 3.4.17.** For any  $\gamma \in (0, 1)$  an element  $\widehat{\mathbf{X}} \in \mathcal{R}_b^\Gamma(\mathbb{A}_N)$  is said to be a *symmetrical bracket extension* if it satisfies for all  $i_1, \dots, i_n \in \{1, \dots, d\}$ ,  $1 \leq n \leq N$  the identity

$$\langle \widehat{\mathbf{X}}_{st}, \bullet_{\mathcal{S}((i_1 \dots i_n))} \rangle = \langle \widehat{\mathbf{X}}_{st}, [[i_1 \dots i_n]] \rangle. \quad (3.4.21)$$

*Remark 3.4.18.* Using the same notation of the  $(i_1 \dots i_n)$ -variation for some path  $X$  in the Remark 3.4.9, if  $\widehat{\mathbf{X}}$  is a rough path over  $\widehat{X} \in \mathcal{C}^\Gamma(\mathbb{A}_N)$  and  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$  is the projection of  $\widehat{X}$  on  $\mathbb{R}^d$  then for any  $\alpha \in \mathbb{A}_N$  we call the path  $\widehat{X}^\alpha$  the  $\alpha$ -variation of  $X$ .

Thus we can reformulate the change of variable formula for a path  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$  in the branched case only in terms of a symmetrical bracket extension.

**Proposition 3.4.19.** *Let  $\gamma \in (0, 1)$  and  $N = \lfloor \gamma^{-1} \rfloor$  and  $\mathbf{X} \in \mathcal{R}_b^\gamma(\mathbb{R}^d)$  over  $X$ . Then for any  $\varphi \in C_b^{N+1}(\mathbb{R}^d, \mathbb{R})$  and any symmetrical bracket extension  $\widehat{\mathbf{X}}$  over  $\mathbf{X}$  one has*

$$\varphi(X_t) = \varphi(X_s) \sum_{i=1}^d \int_s^t \partial_i \Phi(X_r) d\mathbf{X}_r^i + \sum_{n=2}^N \sum_{\beta \in \mathbb{A}_N: |\beta|=n} \frac{1}{\beta!} \int_s^t \partial^\beta \Phi(X_r) d\widehat{\mathbf{X}}_r^\beta, \quad (3.4.22)$$

where  $\partial^\beta \Phi$  is the lift of the function  $\partial^\beta \varphi: \mathbb{R}^d \rightarrow \mathbb{R}$ .

*Proof.* Let us fix  $\varphi$  and  $\widehat{\mathbb{X}}$ . Using the Proposition 3.4.16 we apply the Kelly's Change of variable formula with the bracket extension  $\widehat{\mathbf{X}} = \mathcal{S}^* \widehat{\mathbb{X}}$  over the well defined path  $\widehat{X} \in \mathcal{C}^\Gamma(\mathcal{A}_N)$  defined for any  $(i_1 \cdots i_n) \in \mathcal{A}_N$  as  $\widehat{X}_t^{(i_1 \cdots i_n)} = \mathbb{X}_t^{\mathcal{S}((i_1 \cdots i_n))}$ . Thus we obtain

$$\varphi(X_t) - \varphi(X_s) = \sum_{i=1}^d \int_s^t \partial_i \Phi(X_r) d\mathbf{X}_r^i + \sum_{n=2}^N \frac{1}{n!} \sum_{i_1, \dots, i_n=1}^d \int_s^t \partial_{i_1} \cdots \partial_{i_n} \Phi(X_r) d\widehat{\mathbf{X}}_r^{(i_1 \cdots i_n)}.$$

For any  $i_1, \dots, i_n \in \{1, \dots, d\}$  one has the straightforward identity

$$\int_s^t \partial_{i_1} \cdots \partial_{i_n} \Phi(X_r) d\widehat{\mathbf{X}}_r^{(i_1 \cdots i_n)} = \int_s^t \partial^{\mathcal{S}((i_1 \cdots i_n))} \Phi(X_r) d\widehat{\mathbb{X}}_r^{\mathcal{S}((i_1 \cdots i_n))},$$

because the approximating sums defining both integrals are the same. Therefore we can apply the identity (3.4.18) and we obtain the desired formula.  $\square$

### 3.5 Quasi-geometric rough paths

The Proposition 3.4.1 and the Theorem 3.4.10 show us a deterministic change of variable formula which works with every path  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$ . These formulae are extremely general but they rely in the first case on a particular choice of  $\mathbb{X} \in \mathcal{R}_g^\gamma(\mathbb{R}^d)$  over  $X$  and in the second case on particular choice of a branched rough path  $\mathbf{X} \in \mathcal{R}_b^\gamma(\mathbb{R}^d)$  over  $X$  and the additional choice of a bracket extension. In what follows we introduce a new notion of rough paths defined on  $\mathbb{A}_N$ . This new object will be called *quasi-geometric rough path* and it will provide a sufficient simpler condition to build a bracket extension over  $X$ , as well as an alternative change of variable formula.

#### 3.5.1 Quasi-shuffle algebras and quasi-geometric rough paths

Looking at the Definition 3.4.17, we have rewritten the notion of a bracket extension in terms of a branched rough path with values in  $\mathbb{A}_N \subset \mathbb{A}$ . The alphabet  $\mathbb{A}$  comes naturally with an intrinsic operation of sum between multi-indices  $+: \mathbb{A} \times \mathbb{A} \rightarrow \mathbb{A}$  which is trivially associative and commutative. Using this additional operation we can introduce the notion of *quasi-shuffle algebra*, another Hopf algebra very similar to the tensor algebra. We recall its construction following the framework of [HI17].

To define a quasi shuffle algebra we start from an alphabet  $A$  endowed with a commutative and an associative bilinear map  $[\cdot, \cdot]: \mathbb{R}^A \widetilde{\otimes} \mathbb{R}^A \rightarrow \mathbb{R}^A$ . For any word  $w \in W(A)$ ,  $w = a_1 \cdots a_n$  we use the shorthand notation

$$[a_1 \cdots a_n] := [a_1[\cdots[a_{n-1}, a_n]] \cdots] \in \mathbb{R}^A$$

and the hypotheses on  $[\cdot, \cdot]$  allow to write unambiguously  $[a_1 \cdots a_n]$  independently on the order of the letters and the parenthesis. Using  $[\cdot, \cdot]$  we can define a new product on  $T(A)$ , the *quasi-shuffle product*  $\widehat{\sqcup}: T(A) \widetilde{\otimes} T(A) \rightarrow T(A)$ , given recursively as follows:

- for any word  $v \in W(A)$

$$\mathbf{1} \widehat{\sqcup} v := v \widehat{\sqcup} \mathbf{1} := v,$$

- for any couple of words  $v, w \in W(A)$  and any couple of letters  $a, b \in A$

$$av \hat{\sqcup} bw := (v \hat{\sqcup} bw)a + (av \hat{\sqcup} w)b + (v \hat{\sqcup} w)[ab]. \tag{3.5.1}$$

The couple  $(T(A), \hat{\sqcup})$  is called *quasi-shuffle algebra* and it has been shown in [Hof00, Thm. 2.1.] that it is always a commutative algebra for any choice of  $[\cdot, \cdot]$ . Similarly to the definition of the shuffle product in Section 3.2.1, it is possible to represent explicitly the quasi shuffle product in terms of surjections (see e.g. [EFMPW15, Pag. 7]). For any word  $w = a_1 \cdots a_n$  and all surjective map  $f: \{1, \dots, n\} \rightarrow \{1, \dots, p\}$  for  $p \leq n$  we set

$$f(a_1 \cdots a_n) := \left[ \prod_{j \in f^{-1}(1)} a_j \right] \cdots \left[ \prod_{j \in f^{-1}(p)} a_j \right].$$

Then we have the following identity

$$a_1 \cdots a_n \hat{\sqcup} b_1 \cdots b_m = \sum_{f \in \mathbb{S}_{n,m}} f(a_1 \cdots a_n b_1 \cdots b_m), \tag{3.5.2}$$

where  $\mathbb{S}_{n,m}$  is the set of all surjections  $f: \{1, \dots, n+m\} \rightarrow \{1, \dots, k\}$  satisfying  $f(1) < \dots < f(n), f(n+1) < \dots < f(n+m)$  for every integer  $k$  such that  $\max(m, n) \leq k \leq n+m$ . The main important property of the quasi shuffle property (see [Hof00, Thm. 3.1.]) is that the triple  $(T(A), \hat{\sqcup}, \nabla)$  with the same deconcatenation coproduct  $\nabla$  given in (3.2.2) is again a bialgebra with the same unity and counity of  $T(A)$ . We will denote it by the symbol  $\hat{T}(A)$  to stress the presence of a different product. The bialgebra  $\hat{T}(A)$  is naturally graduated but the natural graduation with respect the word length is not always compatible with  $\hat{\sqcup}$ . However it is always possible to find a sequence of vector spaces  $\{\mathcal{W}_k\}_{k \geq 0}$  depending on  $[\cdot, \cdot]$  such that  $\mathcal{W}_0 = \langle \mathbf{1} \rangle$  and  $\hat{T}(A)$  becomes a graduated bialgebra with respect to the subspaces  $\{\mathcal{W}_k\}_{k \geq 0}$ . Therefore  $\hat{T}(A)$  may also be viewed as a Hopf algebra.

Looking at the algebraic dual  $T((A))$  (which we denote by  $\hat{T}((A))$  in the quasi shuffle case) the choice of a different product for  $T(A)$  does not change the convolution product which is still the dual of the concatenation product  $\otimes$  but in this case the set of characters (which we denote by  $\hat{\mathbb{G}}(A)$ ) is given by

$$\hat{\mathbb{G}}(A) := \{h \in \hat{T}((A)) : \text{for every } a, b \in \hat{T}(A) \quad \langle h, a \hat{\sqcup} b \rangle = \langle h, a \rangle \langle h, b \rangle\}.$$

This set is still a group with the concatenation operation  $\otimes$  because the coproduct  $\nabla$  does not change. The same reasoning applies also to  $\hat{\mathbb{G}}_N(A) := \pi_N(\hat{\mathbb{G}}(A))$  and to  $\hat{\mathbb{G}}_\Gamma(A) := \pi_\Gamma(\hat{\mathbb{G}}(A))$  for any weighted alphabet  $(A, \Gamma)$ . In this last case we recall that it is not necessary to specify any behaviour between the product  $[\cdot, \cdot]$  and  $\Gamma$  to define  $\hat{\mathbb{G}}_\Gamma(A)$ , however for any given weighted alphabet  $(A, \Gamma)$  it will useful to introduce the following compatibility condition. Denoting for any  $a \in A$

$$\mathbb{R}^{\gamma_a} := \text{span}\{b \in A : \gamma_b = \gamma_a\} \subset \mathbb{R}^A,$$

we say that the product  $[\cdot, \cdot]$  is compatible with  $\Gamma$  if for any  $a, b \in A$  the product  $[\cdot, \cdot]$  satisfies on the subspaces  $\{\mathbb{R}^{\gamma_a}\}_{a \in A}$  the following property

$$[\cdot, \cdot]: \mathbb{R}^{\gamma_a} \tilde{\otimes} \mathbb{R}^{\gamma_b} \rightarrow \mathbb{R}^{\gamma_a + \gamma_b}.$$

Shuffle and quasi shuffle structures are intimately related. Indeed if  $[\cdot, \cdot]$  is the trivial product which is always zero, from (3.5.1) we obtain trivially (3.2.1), the recursive definition of the shuffle product, and we recover the same algebraic structure. On the other hand there exists also an explicit homomorphism between  $\hat{T}(A)$  and  $T(A)$ . We recall its construction. For any integer  $n > 0$  we say that the multi-index  $I = (i_1, \dots, i_m)$  with all strictly positive components is a composition of  $n$  if  $i_1 + \dots + i_m = n$ . The set of all composition of  $n$  is denoted by  $C(n)$  and its cardinality is  $2^{n-1}$  using some elementary combinatorial arguments. Using the same multi-index notation, for any composition  $I = (i_1, \dots, i_m)$  we set

$$I! = i_1! \cdots i_m!, \quad p(I) = i_1 \cdots i_m, \quad l(I) = m.$$

Moreover for any word  $w \in T(A)$ ,  $w = a_1 \cdots a_n$  and any  $I \in C(n)$ ,  $I = (i_1, \dots, i_m)$  we define the *contracted word*  $[w]_I$  as

$$[w]_I := [w_1 \cdots w_{i_1}][w_{i_1+1} \cdots w_{i_2}] \cdots [w_{i_{m-1}+1} \cdots w_n].$$

Using this notation we define two fundamental applications.

**Definition 3.5.1.** We define the maps  $\exp, \log: T(A) \rightarrow T(A)$  as the unique linear maps such that for any word  $w \in W(A)$  one has

$$\exp(w) := \sum_{I \in C(|w|)} \frac{1}{I!} [w]_I, \quad \log(w) := \sum_{I \in C(|w|)} \frac{(-1)^{|w|-l(I)}}{p(I)} [w]_I, \quad (3.5.3)$$

We will refer to these application as the *exponential* and the *logarithm* of words.

For instance for any triple of words  $a, b, c \in A$  the definitions in (3.5.3) become

$$\begin{aligned} \exp(ab) &= ab + \frac{1}{2}[ab], & \exp(abc) &= abc + \frac{1}{2!}[ab]c + \frac{1}{2!}a[bc] + \frac{1}{3!}[abc]; \\ \log(ab) &= ab - \frac{1}{2}[ab], & \log(abc) &= abc - \frac{1}{2}[ab]c - \frac{1}{2}a[bc] + \frac{1}{3}[abc]. \end{aligned}$$

A fundamental result of quasi-shuffle algebra consists in the relation between these functions and the tensor algebra  $T(A)$ . For its proof see [Hof00, Thm. 2.5].

**Theorem 3.5.2.** *The exponential of words defined in (3.5.3) is the unique Hopf algebra isomorphism between  $T(A)$  and  $\hat{T}(A)$  whose inverse is the logarithm of words.*

*Remark 3.5.3.* The main consequence of Theorem 3.5.2 implies that many definitions of  $\hat{T}(A)$  can be expressed using the exponential and logarithm of words on  $T(A)$ . For instance considering the adjoint maps  $\exp^*: \hat{T}((A)) \rightarrow T((A))$  and  $\log^*: T((A)) \rightarrow \hat{T}((A))$ , we deduce immediately the identities

$$\log^*(\mathbb{G}(A)) = \hat{\mathbb{G}}(A), \quad \exp^*(\hat{\mathbb{G}}(A)) = \mathbb{G}(A). \quad (3.5.4)$$

Moreover the identities (3.5.4) hold also for  $\hat{\mathbb{G}}_N(A)$  and  $\mathbb{G}_N(A)$  because  $\exp(w)$  and  $\log(w)$  are linear combinations of words with at most length  $|w|$ . In case  $[\cdot, \cdot]$  is compatible with  $\Gamma$  we deduce from the explicit definitions in (3.5.3) that for any  $w \in W(A)$

$\exp(w)$  and  $\log(w)$  are a linear combination of words where  $|\cdot|_\Gamma$  is constant on each word. Supposing this compatibility condition we have

$$\log^*(\mathbb{G}_\Gamma(A)) = \hat{\mathbb{G}}_\Gamma(A), \quad \exp^*(\hat{\mathbb{G}}_\Gamma(A)) = \mathbb{G}_\Gamma(A), \quad (3.5.5)$$

and the applications  $\exp^*$  and  $\log^*$  are group isomorphism.

As already explained in [BCEF18], we can explicitly link the quasi shuffle algebra  $\hat{T}(A)$  with  $\mathcal{H}(A)$  by means of an explicit linear map  $\psi: \mathcal{H}(A) \rightarrow \hat{T}(A)$ .

**Definition 3.5.4.** Let  $\psi: \mathcal{H}(A) \rightarrow \hat{T}(A)$  be the unique linear map defined recursively on the forests by the conditions:

- 1)  $\psi(\mathbf{1}) := \mathbf{1}$ ,
- 2) for every  $a \in A$  and every forest  $f = \tau_1 \cdots \tau_m \in \mathcal{F}(A)$

$$\psi(B_+^a(f)) := (\psi(\tau_1) \hat{\sqcup} \cdots \hat{\sqcup} \psi(\tau_m))a, \quad (3.5.6)$$

- 3) for every couple of forests  $\sigma, \tau \in \mathcal{F}(A)$   $\psi(\sigma\tau) := \psi(\sigma) \hat{\sqcup} \psi(\tau)$ .

We call  $\psi$  the *contracting arborification*.

We will denote by  $\psi^*: \hat{T}(A) \rightarrow \mathcal{H}(A)$  the adjoint of  $\psi$ . By construction  $\psi$  is a surjective algebra, but in general it is also an Hopf algebra morphism.

**Proposition 3.5.5.** For any  $h \in \mathcal{H}(A)$  one has

$$\nabla\psi(h) = (\psi \tilde{\otimes} \psi)\Delta h. \quad (3.5.7)$$

*Proof.* The proof is essentially the same as [HK15, Lem. 4.8] in the context of quasi shuffle product. We repeat here for the sake of completeness. Writing  $\mathcal{H}(A) = \bigoplus_{n=0}^{+\infty} \mathcal{H}_n(A)$  we will prove (3.5.7) on  $\bigoplus_{k=0}^m \mathcal{H}_k(A)$  for every  $m$  by induction. When  $m = 0$  the identity holds trivially. Supposing the identity true on  $\bigoplus_{k=0}^m \mathcal{H}_k(A)$  we prove it on  $\bigoplus_{k=0}^{m+1} \mathcal{H}_k(A)$  by proving the identity (3.5.7) for all forests  $h$  such that  $|h| = m + 1$ . Indeed any such forest  $h$  could be under form  $h = B_+^a(\sigma)$  or  $h = h_1 h_2$  for some forests  $\sigma, h_1, h_2$  such that  $|\sigma| \vee |h_1| \vee |h_2| \leq m$  and  $a \in A$ . In the first case one has

$$\nabla\psi(B_+^a(\sigma)) = \nabla(\psi(\sigma)a) = \psi(\sigma)a \tilde{\otimes} \mathbf{1} + \nabla\phi(\sigma)(\mathbf{1} \tilde{\otimes} a). \quad (3.5.8)$$

By induction  $\nabla\phi(\sigma) = (\psi \tilde{\otimes} \psi)\Delta\sigma$  and the right hand side of (3.5.8) becomes

$$\psi(\sigma)a \tilde{\otimes} \mathbf{1} + ((\psi \tilde{\otimes} \psi)\Delta\sigma)(\mathbf{1} \tilde{\otimes} a) = (\psi \tilde{\otimes} \psi)(B_+^a(\sigma) \tilde{\otimes} \mathbf{1} + (\text{id} \tilde{\otimes} B_+^a)\Delta\sigma) = (\psi \tilde{\otimes} \psi)\Delta B_+^a(\sigma).$$

In the other case, using the bialgebra properties of  $\nabla$  and  $\hat{\sqcup}$ , one has

$$\nabla\psi(h_1 h_2) = \nabla(\psi(h_1) \hat{\sqcup} \psi(h_2)) = \nabla\psi(h_1) \hat{\sqcup} \nabla\psi(h_2), \quad (3.5.9)$$

where we denote again by  $\hat{\sqcup}$  the tensorisation of the quasi shuffle product. The recursive hypothesis tells us that  $\nabla(\psi(h_i)) = (\psi \tilde{\otimes} \psi)\Delta h_i$  for  $i = 1, 2$ , therefore the definition of  $\psi$  as well as the bialgebra property of  $\Delta$  allows to write the right hand side of (3.5.9) as

$$(\psi \tilde{\otimes} \psi)\Delta h_1 \hat{\sqcup} (\psi \tilde{\otimes} \psi)\Delta h_2 = (\psi \tilde{\otimes} \psi)\Delta h_1 \cdot \Delta h_2 = (\psi \tilde{\otimes} \psi)\Delta h_1 h_2,$$

and the proposition is proved.  $\square$

*Remark 3.5.6.* The name of  $\psi$  comes as a tribute to Jean Lecalle’s arborification apparatus in the context of Hopf algebras of trees, as explained in [FM17]. We recall that in the shuffle case (that is  $\psi$  satisfies the properties 2) and 3) above with the shuffle product) the same function was introduced in [HK15] and it was denoted by  $\phi_g$ . Actually it is also possible to define a right inverse of both  $\psi$  and  $\phi_g$ ,  $\iota: T(A) \rightarrow \mathcal{H}(A)$  which maps any word  $w = a_1 \cdots a_n$  into  $\iota(a)$ , the associated *ladder tree* whose decoration on the leaf is  $a_1$ , the decoration on the vertex next to the leaf is  $a_2$  and so on until  $a_n$  is the decoration of the root. The map  $\iota$  is not an algebra morphism but using the recursive definition of  $\Delta$  it is possible to prove that it is a coalgebra morphism.

The algebraic structure of  $\hat{T}(A)$  leads naturally to the notion of *quasi-geometric rough path* for any fixed finite weighted alphabet  $(A, \Gamma)$  endowed with a commutative and associative bilinear map  $[\cdot, \cdot]: \mathbb{R}^A \tilde{\otimes} \mathbb{R}^A \rightarrow \mathbb{R}^A$ .

**Definition 3.5.7.** A  $\Gamma$ -regular quasi-geometric rough path  $\mathbb{X}$  is a function  $\mathbb{X}: [0, T]^2 \rightarrow T((A))$  satisfying for every  $w_1, w_2 \in T(A)$

$$\langle \mathbb{X}_{st}, w_1 \hat{\sqcup} w_2 \rangle = \langle \mathbb{X}_{st}, w_1 \rangle \langle \mathbb{X}_{st}, w_2 \rangle, \tag{3.5.10}$$

and the conditions 2g) and 3g) in the Definition 3.2.1

We denote by  $\hat{\mathcal{R}}_g^\Gamma(A)$  the set of all quasi-geometric rough paths. The Theorem 3.5.2 implies that we can equivalently rewrite the quasi-geometric rough path as a geometric rough paths and vice versa. As before we denote by  $\exp^*: \hat{T}((A)) \rightarrow T((A))$  and  $\log^*: T((A)) \rightarrow \hat{T}((A))$  the dual maps of  $\exp$  and  $\log$ .

**Proposition 3.5.8.** *Supposing that the product  $[\cdot, \cdot]$  is compatible with  $\Gamma$ , then the map  $\mathbb{X} \rightarrow \exp^* \mathbb{X}$  is a bijective application between  $\hat{\mathcal{R}}_g^\Gamma(A)$  and  $\mathcal{R}_g^\Gamma(A)$  whose inverse is given by the application  $\mathbf{X} \rightarrow \log^* \mathbf{X}$ .*

*Proof.* The proof follows from the Theorem 3.5.2. We explain in short how the algebraic properties of  $\exp$  allows to prove that  $\exp^* \mathbb{X} \in \mathcal{R}_g^\Gamma(A)$  for any  $\mathbb{X} \in \hat{\mathcal{R}}_g^\Gamma(A)$ . For any fixed couple of words  $w, v \in T(A)$  the morphism property of  $\exp$  and the equation (3.5.10) imply

$$\begin{aligned} \langle \exp^* \mathbb{X}_{st}, w \sqcup v \rangle &= \langle \mathbb{X}_{st}, \exp(w \sqcup v) \rangle = \langle \mathbb{X}_{st}, \exp(w) \hat{\sqcup} \exp(v) \rangle \\ &= \langle \mathbb{X}_{st}, \exp(w) \rangle \langle \mathbb{X}_{st}, \exp(v) \rangle = \langle \exp^* \mathbb{X}_{st}, w \rangle \langle \exp^* \mathbb{X}_{st}, v \rangle. \end{aligned}$$

At the same time the comorphism property of  $\exp$  implies for any  $s, u, t \in [0, T]^3$

$$(\exp^* \mathbb{X}_{su}) \otimes (\exp^* \mathbb{X}_{ut}) = (\mathbb{X}_{su} \tilde{\otimes} \mathbb{X}_{ut})(\exp \tilde{\otimes} \exp) \nabla = (\mathbb{X}_{su} \tilde{\otimes} \mathbb{X}_{ut}) \nabla \exp = \exp^* \mathbb{X}_{st}.$$

Finally using the hypothesis of compatibility between  $[\cdot, \cdot]$  and  $\Gamma$ , for any fixed word  $w \in W(A)$  there exist  $N \geq 1$  and a family of words  $\{w_i\}_{i=1, \dots, N}$  and  $\beta_i \geq 0$  such that

$$\exp(w) = \sum_{i=1}^N \beta_i w_i$$

and  $|w_i|_\Gamma = |w|_\Gamma$  for all  $i$  (see the Remark 3.5.3). Thus we obtain

$$\frac{|\langle \exp^* \mathbb{X}_{st}, w \rangle|}{|t - s|^{\hat{\gamma}|w|_\Gamma}} < \infty$$

Using the same reasoning we obtain that  $\mathbf{X} \rightarrow \log^* \mathbf{X}$  is a well defined function from  $\mathcal{R}_g(A)$  to  $\hat{\mathcal{R}}_g(A)$ . Since  $\log \exp = \exp \log = \text{id}$ , we conclude.  $\square$

Looking at the Definition 3.5.7, we can easily restate the notion of  $\Gamma$ -regular quasi-geometric rough path in terms of a path with values in  $\hat{\mathcal{G}}_\Gamma(A)$  and we can prove the equivalence by means of an explicit extension theorem. Thanks to Theorem 3.5.2 this result is already a consequence of the Theorem 3.2.7.

**Corollary 3.5.9** (Extension Theorem for quasi-geometric rough path). *Supposing that the product  $[\cdot, \cdot]$  is compatible with  $\Gamma$ , for any function  $\tilde{\mathbb{X}}: [0, T]^2 \rightarrow \hat{\mathcal{G}}_\Gamma(A)$  satisfying the same properties 1) 2) 3) of Definition 3.2.5, there exists a unique  $\mathbb{X} \in \hat{\mathcal{R}}_g^\Gamma(A)$  such that  $\pi_\Gamma \circ \mathbb{X} = \tilde{\mathbb{X}}$ . The function  $\tilde{\mathbb{X}}$  is called a quasi-geometric  $\Gamma$ -rough path.*

*Proof.* By means of the algebraic properties of the map  $\exp$  it is straightforward to prove that the function  $\tilde{\mathbf{X}}_{st} := \exp^* \tilde{\mathbb{X}}_{st}$  is a geometric  $\Gamma$ -rough path. Using the Theorem 3.2.7 on  $\tilde{\mathbf{X}}$  there exists a unique  $\mathbf{X} \in \mathcal{R}_g^\Gamma(A)$  such that  $\pi_\Gamma \circ \mathbf{X} = \tilde{\mathbf{X}}$ . Applying to both sides the map  $\log^*$ , the compatibility condition of  $[\cdot, \cdot]$  with  $\Gamma$  implies that the map  $\pi_\Gamma$  and  $\log^*$  commute. We conclude that  $\mathbb{X} = \log^* \mathbf{X}$  satisfies the properties of the statement.  $\square$

The same strategy allows to prove the following version of the Lyons-Victoir extension Theorem.

**Corollary 3.5.10** (Lyons-Victoir extension Theorem for quasi-geometric). *Supposing that the weights  $\{\gamma_a\}_{a \in A}$  satisfy the condition  $1 \notin \sum_{a \in A} \gamma_a \mathbb{N}$ , for any  $X \in \mathcal{C}^\Gamma(A)$  there exists  $\mathbb{X} \in \hat{\mathcal{R}}_g^\Gamma(A)$  over  $X$ .*

Another deep property of quasi-geometric rough paths is then given by the contracting arborification  $\psi$ . Indeed the existence of such Hopf algebra morphism allows to write down a quasi-geometric rough path as a specific family of branched rough path. In what follows we will denote by  $\psi^*: \hat{T}((A)) \rightarrow \mathcal{H}((A))$  the dual map of  $\psi$  and  $\iota^*: \mathcal{H}((A)) \rightarrow T((A))$ , the dual map of  $\iota: T(A) \rightarrow \mathcal{H}(A)$  introduced in the Remark 3.5.6. If the product  $[\cdot, \cdot]$  is compatible with  $\Gamma$  it is straightforward to show that the function  $\iota^* \pi_\Gamma$  is a well defined function from  $\mathcal{H}((A))$  to  $T^\Gamma((A))$ .

**Proposition 3.5.11.** *Supposing that the product  $[\cdot, \cdot]$  is compatible with  $\Gamma$ , then for any  $\mathbb{X} \in \hat{\mathcal{R}}_g^\Gamma(A)$  the function  $\psi^* \mathbb{X} \in \mathcal{R}_b(A)$ . Moreover an element  $\mathbf{X} \in \mathcal{R}_b^\Gamma(A)$  satisfies  $\mathbf{X} = \psi^* \mathbb{X}$  for some unique  $\mathbb{X} \in \hat{\mathcal{R}}_g(A)$  if and only if*

$$\langle \mathbf{X}_{st}, \sigma \rangle = \langle \mathbf{X}_{st}, \iota \psi(\sigma) \rangle \tag{3.5.11}$$

for every forest  $\sigma \in \mathcal{F}^\Gamma(A)$ .

*Proof.* We fix  $\mathbb{X} \in \widehat{\mathcal{R}}_g^\Gamma(A)$ . In order to check the properties (1b) (2b) (3b) for  $\psi^*\mathbb{X}$ , we use the same argument as in the Proposition 3.5.8 because when the product  $[\cdot, \cdot]$  is compatible with  $\Gamma$  then for any forest  $\sigma$   $\psi(\sigma)$  is a linear combination of words where  $|\cdot|_\Gamma$  does not change. Moreover if we impose  $\mathbf{X} = \psi^*\mathbb{X}$  one has trivially (3.5.11). On the other hand if one element  $\mathbf{X}$  satisfies (3.5.11), we introduce the function  $\mathbf{X}' := \iota^*\pi_\Gamma\mathbf{X}: [0, T]^2 \rightarrow T^\Gamma((A))$ . Using the hypothesis (3.5.11) and the fact that  $\mathbf{X} \in \mathcal{R}_b^\Gamma(A)$  we have the the following chain of identities

$$\begin{aligned} \langle \mathbf{X}'_{st}, v \widehat{\sqcap} w \rangle &= \langle \mathbf{X}'_{st}, \psi(\iota v \iota w) \rangle = \langle \pi_\Gamma \mathbf{X}_{st}, \iota \psi(\iota v \iota w) \rangle = \langle \pi_\Gamma \mathbf{X}_{st}, \iota v \iota w \rangle \\ &= \langle \pi_\Gamma \mathbf{X}_{st}, \iota v \rangle \langle \pi_\Gamma \mathbf{X}_{st}, \iota w \rangle = \langle \mathbf{X}'_{st}, v \rangle \langle \mathbf{X}'_{st}, w \rangle, \end{aligned}$$

for any couple of words  $v, w \in W(A)$  such that  $v \widehat{\sqcap} w \in T^\Gamma(A)$ . Using also the comorphism property of  $\iota$ , we can straightforward look at  $\mathbf{X}'$  as a function satisfying the hypothesis of the Corollary 3.5.9. Therefore there exists a unique  $\mathbb{X} \in \widehat{\mathcal{R}}_g^\Gamma(A)$  such that  $\pi_\Gamma\mathbb{X} = \iota^*\pi_\Gamma\mathbf{X}$ . Since  $\pi_\Gamma$  and  $\psi$  commute and  $\psi^*\iota^* = \text{id}$  we have  $\pi_\Gamma\psi^*\mathbb{X} = \pi_\Gamma\mathbf{X}$ . Thus we obtain  $\psi^*\mathbb{X} = \mathbf{X}$ .  $\square$

*Remark 3.5.12.* An equivalent version of Proposition 3.5.11 in the shuffle case is given in [HK15, Prop. 4.1], where the map  $\psi$  is replaced by  $\phi_g$  (see Remark 3.5.6).

### 3.5.2 Building a symmetrical bracket extensions

Let us apply the notion of quasi-geometric rough path in order to build a symmetrical bracket extension. Recalling some notation of the Section 3.4, for any  $\gamma \in (0, 1)$  we can endow the weighted alphabet  $(\mathbb{A}_N, \Gamma)$  of a commutative and associative bilinear product  $[\cdot, \cdot]: \mathbb{R}^{\mathbb{A}_N} \widehat{\otimes} \mathbb{R}^{\mathbb{A}_N} \rightarrow \mathbb{R}^{\mathbb{A}_N}$  defined for every  $\alpha, \beta \in \mathbb{A}_N$  as

$$[\alpha \widehat{\otimes} \beta] := \begin{cases} \alpha + \beta & \text{if } \alpha + \beta \in \mathbb{A}_N, \\ 0 & \text{otherwise} \end{cases} \quad (3.5.12)$$

This product is clearly commutative, associative and compatible with the weight  $\Gamma$ , therefore it is possible to consider the quasi shuffle algebra  $\widehat{T}(\mathbb{A}_N)$ . The presence of the product  $[\cdot, \cdot]$  allows to rewrite for any  $i_1, \dots, i_k \in \{1, \dots, d\}$  and  $k \leq N = \lfloor \gamma^{-1} \rfloor$  the identity

$$\mathcal{S}((i_1 \cdots i_k)) = [i_1 \cdots i_k] \quad (3.5.13)$$

where in the right hand side the letters  $i_1, \dots, i_d$  are identified with the corresponding vector of the canonical basis. Therefore we can graphically write the map  $\mathcal{S}$  by simply turning the rounded brackets of the decorations in  $\mathcal{A}_N$  into square brackets. There exists a deep connection between the notion of symmetrical bracket polynomial and the quasi-shuffle arborification. A central result in our analysis will be the following characterisation of the elements of  $\widehat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  in terms of bracket extensions.

**Theorem 3.5.13.** *Let  $\gamma \in (0, 1)$ . For any  $\mathbb{X} \in \widehat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  the rough path  $\psi^*\mathbb{X} \in \mathcal{R}_b^\Gamma(\mathbb{A}_N)$  is a symmetrical bracket extension.*

In order to prove this theorem we need the following algebraic lemma which links symmetrical bracket polynomials defined in (3.4.20) and the contracting arborification.



**Lemma 3.5.14.** *Let  $\gamma \in (0, 1)$ . Then we have*

$$\psi(\llbracket i_1 \cdots i_n \rrbracket) = [i_1 \cdots i_n], \quad (3.5.14)$$

for all  $1 \leq n \leq N = \lfloor \gamma^{-1} \rfloor$  and all  $i_1, \dots, i_n \in \{1, \dots, d\}$ .

*Proof.* We fix  $N$  and we will prove the theorem by induction on  $n$  up to the value  $N$ . In the case  $n = 1$  the identity (3.5.14) is verified by definition. We suppose that (3.5.14) holds for some  $1 \leq n < N$  and we want to prove

$$\psi(\llbracket i_1 \cdots i_{n+1} \rrbracket) = [i_1 \cdots i_{n+1}].$$

In order to do that we establish a recursive identity over the symmetrical bracket polynomials. Using the formula (3.4.7) and the following straightforward identity

$$\begin{aligned} \Delta'(\bullet_{i_1} \cdots \bullet_{i_{n+1}}) &= (\bullet_{i_{n+1}} \otimes \bullet_{i_1} \cdots \bullet_{i_n}) + (\bullet_{i_{n+1}} \otimes \mathbf{1})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) \\ &\quad + (\bullet_{i_1} \cdots \bullet_{i_n} \otimes \bullet_{i_{n+1}}) + (\mathbf{1} \otimes \bullet_{i_{n+1}})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n}), \end{aligned}$$

we have immediately

$$\begin{aligned} \llbracket i_1 \cdots i_{n+1} \rrbracket &= \psi(\mathcal{S}(\langle\langle i_1 i_2 \cdots i_n \rangle\rangle)) = \bullet_{i_1} \cdots \bullet_{i_{n+1}} - B_+^{[i_1 \cdots i_n]}(\bullet_{i_{n+1}}) - B_+^{i_{n+1}}(\bullet_{i_1} \cdots \bullet_{i_n}) \\ &\quad - \mathcal{SB}_+(\bullet_{i_{n+1}} \otimes \mathbf{1})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) - \mathcal{SB}_+(\mathbf{1} \otimes \bullet_{i_{n+1}})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n}). \end{aligned}$$

Then by definition of  $\psi$  we obtain

$$\begin{aligned} \psi(\llbracket i_1 \cdots i_{n+1} \rrbracket) &= \psi(\bullet_{i_1} \cdots \bullet_{i_n}) \widehat{\sqcup} i_{n+1} - \psi(\bullet_{i_1} \cdots \bullet_{i_n}) i_{n+1} - i_{n+1} [i_1 \cdots i_n] \\ &\quad - \psi(\mathcal{SB}_+(\bullet_{i_{n+1}} \otimes \mathbf{1})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) - \psi(\mathcal{SB}_+(\mathbf{1} \otimes \bullet_{i_{n+1}})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})). \end{aligned} \quad (3.5.15)$$

Writing the induction hypothesis using (3.4.7) one has

$$\psi(\bullet_{i_1} \cdots \bullet_{i_n}) = \psi(\mathcal{SB}_+\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) + [i_1 \cdots i_n],$$

Then the right hand side of (3.5.15) becomes

$$\begin{aligned} [i_1 \cdots i_n] \widehat{\sqcup} i_{n+1} - [i_1 \cdots i_n] i_{n+1} - i_{n+1} [i_1 \cdots i_n] &+ \psi(\mathcal{SB}_+\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) \widehat{\sqcup} i_{n+1} \\ &- \psi(\mathcal{SB}_+\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) i_{n+1} - \psi(\mathcal{SB}_+(\bullet_{i_{n+1}} \otimes \mathbf{1})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) \\ &- \psi(\mathcal{SB}_+(\mathbf{1} \otimes \bullet_{i_{n+1}})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})). \end{aligned}$$

After a simplification of the first three terms this expression becomes

$$\begin{aligned} [i_1 \cdots i_{n+1}] + \psi(\mathcal{SB}_+\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) \widehat{\sqcup} i_{n+1} &- \psi(\mathcal{SB}_+\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) i_{n+1} \\ &- \psi(\mathcal{SB}_+(\mathbf{1} \otimes \bullet_{i_{n+1}})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) - \psi(\mathcal{SB}_+(\bullet_{i_{n+1}} \otimes \mathbf{1})\Delta'(\bullet_{i_1} \cdots \bullet_{i_n})). \end{aligned} \quad (3.5.16)$$

Let us show that the remaining terms after  $[i_1 \cdots i_{n+1}]$  in (3.5.16) are zero. Using again the expression

$$\Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) = \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \bullet_{a_1} \cdots \bullet_{a_k} \otimes \bullet_{b_1} \cdots \bullet_{b_{n-k}},$$

where the sum is done over all the partition of  $\{i_1, \dots, i_n\}$  into two non empty blocks  $a = \{a_1, \dots, a_k\}$  and  $b = \{b_1, \dots, b_{n-k}\}$ , one has

$$\begin{aligned} \mathbf{SB}_+ \Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) &= \sum_{\{a,b\}=\{i_1, \dots, i_n\}} B_+^{[b_1 \cdots b_{n-k}]}(\bullet_{a_1} \cdots \bullet_{a_k}), \\ \mathbf{SB}_+(\bullet_{i_{n+1}} \otimes \mathbf{1}) \Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) &= \sum_{\{a,b\}=\{i_1, \dots, i_n\}} B_+^{[b_1 \cdots b_{n-k}]}(\bullet_{a_1} \cdots \bullet_{a_k} \bullet_{i_{n+1}}), \\ \mathbf{SB}_+(\mathbf{1} \otimes \bullet_{i_{n+1}}) \Delta'(\bullet_{i_1} \cdots \bullet_{i_n}) &= \sum_{\{a,b\}=\{i_1, \dots, i_n\}} B_+^{[b_1 \cdots b_{n-k} i_{n+1}]}(\bullet_{a_1} \cdots \bullet_{a_k}). \end{aligned}$$

Applying the recursive definition of  $\widehat{\sqcup}$  (see (3.2.1)), we finally obtain that

$$\begin{aligned} \psi(\mathbf{SB}_+ \Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) \widehat{\sqcup} i_{n+1} &= \sum_{\{a,b\}=\{i_1, \dots, i_n\}} (\psi(\bullet_{a_1} \cdots \bullet_{a_k}) [b_1 \cdots b_{n-k}]) \widehat{\sqcup} i_{n+1} \\ &= \sum_{\{a,b\}=\{i_1, \dots, i_n\}} \left( \psi(\bullet_{a_1} \cdots \bullet_{a_k}) [b_1 \cdots b_{n-k}] i_{n+1} + (\psi(\bullet_{a_1} \cdots \bullet_{a_k}) \widehat{\sqcup} i_{n+1}) [b_1 \cdots b_{n-k}] \right. \\ &\quad \left. + \psi(\bullet_a) [i_{n+1} b_1 \cdots b_{n-k}] \right) \\ &= \psi(\mathbf{SB}_+ \Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) i_{n+1} + \psi(\mathbf{SB}_+(\bullet_{i_{n+1}} \otimes \mathbf{1}) \Delta'(\bullet_{i_1} \cdots \bullet_{i_n})) \\ &\quad + \psi(\mathbf{SB}_+(\mathbf{1} \otimes \bullet_{i_{n+1}}) \Delta'(\bullet_{i_1} \cdots \bullet_{i_n})). \end{aligned}$$

Plugging this last equality in (3.5.16) we obtain the result.  $\square$

*Proof of Theorem 3.5.13.* Applying Proposition 3.5.11 to the alphabet  $\mathbb{A}_N$  one has immediately  $\psi^* \mathbb{X} \in \mathcal{R}_b^\Gamma(\mathbb{A}_N)$ . Moreover formula (3.5.14) implies directly

$$\langle \psi^* \mathbb{X}_{st}, [i_1 \cdots i_n] \rangle = \langle \mathbb{X}_{st}, \psi([i_1 \cdots i_n]) \rangle = \langle \mathbb{X}_{st}, [i_1 \cdots i_n] \rangle = \langle \psi^* \mathbb{X}_{st}, \bullet_{[i_1 \cdots i_n]} \rangle$$

and using the identification (3.5.13)  $\psi^* \mathbb{X}$  becomes a symmetrical bracket extension.  $\square$

*Remark 3.5.15.* As remarked before, Theorem 3.5.13 is only a sufficient criterion to show the existence of a bracket extension. Indeed the hypothesis  $\mathbf{X} = \psi^* \mathbb{X}$  implies more equations than (3.4.21). For instance when  $d = 1$  and  $\gamma \in (1/4, 1/3]$  the set  $\mathbb{A}_3$  becomes equal to  $\{1, 2, 3\}$  and a symmetrical bracket extension  $\mathbf{X}$  satisfies

$$\langle \mathbf{X}_{st}, \bullet_1 \bullet_1 \rangle = \langle \mathbf{X}_{st}, 2 \mathbf{!}_1^1 + \bullet_2 \rangle, \quad \langle \mathbf{X}_{st}, \bullet_1 \bullet_1 \bullet_1 \rangle = \langle \mathbf{X}_{st}, 3 \mathbf{!}_1^1 + 3 \mathbf{!}_1^2 + \bullet_3 \rangle.$$

(We decided to write the identities (3.4.21) as in (3.4.11)) However a branched rough path under the form  $\psi^* \mathbb{X}$  for some  $\mathbb{X} \in \widehat{T}(\mathbb{A}_3)$  satisfies also the identity

$$\langle \phi^* \mathbb{X}_{st}, \bullet_2 \bullet_1 \rangle = \langle \mathbb{X}_{st}, 2 \widehat{\sqcup} 1 \rangle = \langle \mathbb{X}_{st}, 21 + 12 + 3 \rangle = \langle \phi^* \mathbb{X}_{st}, \mathbf{!}_2^1 + \mathbf{!}_1^2 + \bullet_3 \rangle.$$

Using the property  $\mathbf{X}_{st} = \mathbf{X}_{su} * \mathbf{X}_{ut}$ , it is straightforward to prove that the function

$$(s, t) \rightarrow \langle \mathbf{X}_{st}, \bullet_2 \bullet_1 - \mathbf{!}_2^1 - \mathbf{!}_1^2 \rangle$$

for any  $\mathbf{X} \in \mathcal{R}_b^\Gamma(\mathbb{A}_3)$  can be written as the increment of a path and the additional hypothesis  $\mathbf{X} = \psi^*\mathbb{X}$  implies that this increment must coincide with  $\langle \mathbf{X}_{st}, \bullet_3 \rangle$ . This condition is not contained apriori in the definition of a symmetrical bracket extension. Thus the notion of symmetrical bracket extension is more general than the quasi-geometric rough path. Since in the case of  $\gamma \in (1/3, 1/2]$  there exists a unique bracket extension above any  $\mathbf{X} \in \mathcal{R}_b^\gamma(\mathbb{R}^d)$  (see Remark 3.4.11) and this bracket extension  $\widehat{\mathbf{X}}$  satisfies the identity 3.5.11 by construction, then we have  $\widehat{\mathbf{X}} = \psi^*\mathbb{X}$  for some unique  $\mathbb{X} \in \widehat{\mathcal{R}}_g^\Gamma(\mathbb{A}_2)$ .

### 3.6 Quasi-geometric change of variable formula

Since the notion of quasi-geometric rough path implies the notion of symmetrical bracket extension, for any  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$  and any  $\mathbb{X} \in \widehat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  such that  $\langle \mathbb{X}_{st}, e_i \rangle = X_t^i - X_s^i$  where  $e_i$  is the canonical basis of  $\mathbb{R}^d$  interpreted as a family of letters of  $\mathbb{A}_N$  (we resume this condition by saying that  $\mathbb{X}$  is over  $X$ ), we can easily deduce a change of variable formula on  $X$  by using Theorem 3.4.10 on  $\psi^*\mathbb{X}$ . However the resulting formula can be expressed and proved autonomously, without any reference to this theorem but following the same strategy of Proposition 3.4.1. Thus the two notions of quasi-geometric rough path and symmetrical bracket extension will result absolutely independent to obtain a formula for  $\varphi(X_t)$ .

#### 3.6.1 An alternative formula

In order to prove this new identity we should in principle define a new notion of controlled rough path with respect to  $\mathbb{X}$  (a so called quasi-geometric controlled rough path) but since the Definition 3.3.1 depends only on the property 2g) of a geometric rough path, we can keep the same definition. Starting from  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$  we can introduce for any  $m \geq 1$  the coordinate path of  $X: [0, T] \rightarrow (T_m^\Gamma(\mathbb{A}_N))^d$  (denoted by the same notation), which is given for any  $i = 1, \dots, d$  by

$$\langle k, (X_t)^i \rangle = \begin{cases} X_t^i & \text{if } k = \mathbf{1}^*, \\ 1 & \text{if } k = e_i \\ 0 & \text{for any other } k \in W_m^\Gamma(A), \end{cases} \tag{3.6.1}$$

Using the same argument of Section 3.2 we can trivially show that  $X \in (\mathcal{D}^{m\Gamma}(\mathbb{X}))^d$  for any  $m \geq 1$ . Passing to the rough integration and the lifting operation, we can easily check that the definition of the rough integral in Proposition 3.3.6 and the lift  $\Phi$  in (3.3.14a) and (3.3.14b) can be applied immediately to the quasi-shuffle case. The explicit structure of the alphabet  $\mathbb{A}_N$  and its associated commutative product  $[\cdot, \cdot]$  imply an explicit formula for  $\Phi(X_t)$  for any  $\varphi$  sufficiently smooth.

**Proposition 3.6.1.** *Let  $\gamma \in (0, 1)$  and  $\mathbb{X} \in \widehat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  over  $X$ . Then for any  $m \geq 1$  and  $\varphi \in C_b^m(\mathbb{R}^d, \mathbb{R})$  one has*

$$\Phi(X_t) = \sum_{u \in W_m^\Gamma(\mathbb{A}_N)} \frac{\partial^u \varphi(X_t)}{u!} u, \tag{3.6.2}$$

where the set  $W_m^\Gamma(\mathbb{A}_N)$  is given in (3.3.1) and for any word  $u \in W_m^\Gamma(\mathbb{A}_N)$  we define

$$\partial^u \varphi(X_t) := \begin{cases} \varphi(X_t) & \text{if } u = \mathbf{1}, \\ \partial^{\alpha_1 + \dots + \alpha_n} \varphi(X_t) & \text{if } u = \alpha_{i_1} \dots \alpha_{i_n}, \end{cases}$$

$$u! := \begin{cases} 1 & \text{if } u = \mathbf{1}, \\ \prod_{i=1}^n \alpha_i! & \text{if } u = \alpha_{i_1} \dots \alpha_{i_n}. \end{cases}$$

*Proof.* This proof has the same strategy of the proof given for the Proposition 3.3.10. Following the definition of  $\Phi$  in (3.3.14b) for the coordinate path  $X \in (\mathcal{D}^{m\Gamma}(\mathbf{X}))^d$  defined in (3.6.1) we obtain

$$\Phi(X_t) = \sum_{k \in N_m^d} \frac{\partial^k \varphi(X_t)}{k_1! \dots k_d!} e_1^{\hat{\sqcup} k_1} \hat{\sqcup} \dots \hat{\sqcup} e_d^{\hat{\sqcup} k_d},$$

where  $N_m^d = \{k \in \mathbb{N}^d : 0 \leq |k| < m\}$  and  $e_1, \dots, e_d$  is the canonical basis of  $\mathbb{R}^d$ . Thus the theorem is proved as long as we can verify the identity

$$\sum_{k \in N_m^d} \frac{\partial^k \varphi(X_t)}{k_1! \dots k_d!} e_1^{\hat{\sqcup} k_1} \hat{\sqcup} \dots \hat{\sqcup} e_d^{\hat{\sqcup} k_d} = \sum_{u \in W_m^\Gamma(\mathbb{A}_N)} \frac{\partial^u \varphi(X_t)}{u!} u, \quad (3.6.3)$$

We check this identity by induction on  $d$ . If  $d = 1$  by applying the identity  $\exp \log = \text{id}$  and the previous relation  $a^{\sqcup k} = k! a^{\otimes k}$  for any  $a \in \mathbb{A}_N = \{1, \dots, N\}$ , for any  $k \geq 0$  we have

$$e_1^{\hat{\sqcup} k} = \exp(\log(e_1^{\hat{\sqcup} k})) = \exp(\log(e_1)^{\sqcup k}) = k! \exp(e_1^{\otimes k}) = \sum_{I \in C(k)} \frac{k!}{I!} [e_1^{\otimes k}]_I. \quad (3.6.4)$$

Then the left hand side of (3.6.3) becomes

$$\sum_{k=0}^{m-1} \frac{\partial^k \varphi(X_t)}{k!} e_1^{\hat{\sqcup} k} = \sum_{k=0}^{m-1} \sum_{I \in C(k)} \frac{\partial^k \varphi(X_t)}{I!} [e_1^{\otimes k}]_I = \sum_{u \in W_m^\Gamma(\mathbb{A}_N)} \frac{\partial^u \varphi(X_t)}{u!} u,$$

because the couples  $(k, I)$  where  $I \in C(k)$  are in one-to-one correspondence with the elements  $W_m^\Gamma(\mathbb{A}_N)$ . Supposing the identity (3.6.3) true for  $d - 1$  we write the left hand side of (3.6.3) as

$$\sum_{n=0}^{m-1} \frac{1}{n!} \left( \sum_{k \in N_{m-n}^{d-1}} \frac{\partial^k \partial_d^n \varphi(X_t)}{k_1! \dots k_{d-1}!} e_1^{\hat{\sqcup} k_1} \hat{\sqcup} \dots \hat{\sqcup} e_{d-1}^{\hat{\sqcup} k_{d-1}} \right) \hat{\sqcup} e_d^{\hat{\sqcup} n},$$

where  $\mathbb{A}'_N = \{a \in \mathbb{A}_N, a = (a_1, \dots, a_d) : a_d = 0\}$ . Applying the induction hypothesis for every  $n$  and the identity (3.6.4) on  $e_d^{\hat{\sqcup} n}$  this sum becomes

$$\sum_{n=0}^{m-1} \sum_{I \in C(n)} \sum_{u' \in W_{m-n}^\Gamma(\mathbb{A}'_N)} \frac{\partial^u \partial_d^n \varphi(X_t)}{u'! I!} u \hat{\sqcup} [e_d^{\otimes n}]_I. \quad (3.6.5)$$

Using the direct definition of the quasi-shuffle product given in (3.5.2), the sum (3.6.5) becomes

$$\sum_{n=0}^{m-1} \sum_{I \in C(n)} \sum_{u' \in W_{m-n}^\Gamma(\mathbb{A}'_N)} \frac{\partial^u \partial_d^n \varphi(X_t)}{u'! I!} \sum_{f \in \mathbb{S}_{|u|, I(I)}} f(u[e_d^{\otimes n}]_I).$$

Since all the letters in the words  $u' \in W_{m-n}^\Gamma(\mathbb{A}'_N)$  have the  $d$ -th component equal to zero, the parameters  $(n, I, u', f)$  above are in one-to-one correspondence with all the words  $u \in W_m^\Gamma(\mathbb{A}_N)$  and  $u! = u'! I!$ , thereby obtaining the result.  $\square$

From the explicit formula of  $\Phi(X_t)$  in (3.6.2) we deduce immediately the consequent change of variable formula for  $\varphi(X_t)$ .

**Theorem 3.6.2.** *Let  $\gamma \in (0, 1)$  and  $N = \lfloor \gamma^{-1} \rfloor$ . Then for any  $\varphi \in C_b^{N+1}(\mathbb{R}^d, \mathbb{R})$  and any  $\widehat{\mathbb{X}} \in \widehat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  over  $X$  one has*

$$\varphi(X_t) = \varphi(X_s) + \sum_{\beta \in \mathbb{A}_N} \frac{1}{\beta!} \int_s^t \partial^\beta \Phi(X_r) d\widehat{\mathbb{X}}_r^\beta, \tag{3.6.6}$$

where  $\partial^\beta \Phi$  is the lift of  $\partial^\beta \varphi: \mathbb{R}^d \rightarrow \mathbb{R}$ .

*Proof.* This proof has the same strategy of the proof given for the Proposition 3.4.1. For any  $\beta \in \mathbb{A}_N$  the function  $\partial^\beta \varphi \in C_b^{N+1-|\beta|}(\mathbb{R}^d, \mathbb{R})$  thus the controlled rough path  $\partial^\beta \Phi(X_r) \in \mathcal{D}^{(N+1-|\beta|)\Gamma}(\mathbb{X})$  as a consequence of the Theorem 3.3.8 and the rough integrals in (3.6.6) are all well defined. Using the explicit definition of  $\partial^\beta \Phi(X_r)$  in (3.3.19) and the development (3.3.12) for the rough integral one has

$$\frac{1}{\beta!} \int_s^t \partial^\beta \Phi(X_r) d\widehat{\mathbb{X}}_r^\beta = \sum_{u \in W_\beta^\Gamma(\mathbb{A}_N)} \frac{\partial^u \partial^\beta \varphi(X_t)}{u! \beta!} \langle \widehat{\mathbb{X}}_{st}, u\beta \rangle + o(|t - s|),$$

where  $W_\beta^\Gamma(\mathbb{A}_N) = \{u \in W^\Gamma(\mathbb{A}_N) : 0 \leq |u|_\Gamma \leq N - |\beta|\}$ . Summing on  $\beta \in \mathbb{A}_N$  we obtain

$$\begin{aligned} \sum_{\beta \in \mathbb{A}_N} \frac{1}{\beta!} \int_s^t \partial^\beta \Phi(X_r) d\widehat{\mathbb{X}}_r^\beta &= \sum_{\beta \in \mathbb{A}_N} \sum_{u \in W_\beta^\Gamma(\mathbb{A}_N)} \frac{\partial^u \partial^\beta \varphi(X_t)}{u! \beta!} \langle \widehat{\mathbb{X}}_{st}, u\beta \rangle + o(|t - s|) \\ &= \sum_{w \in W_N^\Gamma(\mathbb{A}_N)} \frac{\partial^w \varphi(X_t)}{w!} \langle \widehat{\mathbb{X}}_{st}, w \rangle + o(|t - s|), \end{aligned} \tag{3.6.7}$$

where  $W_N^\Gamma(\mathbb{A}_N) = \{w \in W^\Gamma(\mathbb{A}_N) : 0 < |w|_\Gamma \leq N\}$ . On the other hand, interpreting  $X_t$  as a controlled rough path of regularity  $N + 1$  the hypothesis of  $\varphi$  implies that  $\Phi(X_t) \in \mathcal{D}^{(N+1)\Gamma}(\mathbb{X})$  by means of the Theorem 3.3.8. Thus we can apply the bound (3.3.2) on  $\langle \mathbf{1}^*, \Phi(X_t) \rangle$  to obtain

$$\varphi(X_t) - \varphi(X_s) = \sum_{u \in W_N^\Gamma(\mathbb{A}_N)} \frac{\partial^u \varphi(X_t)}{u!} \langle \widehat{\mathbb{X}}_{st}, u \rangle + o(|t - s|), \tag{3.6.8}$$

Since the right hand side of the equations (3.6.7) and (3.6.8) is equal to  $o(|t - s|)$ , we conclude using the same argument of the Proposition 3.4.1.  $\square$

The formula (3.6.6) can be obtained also as a trivial consequence of (3.4.22) after this identification of rough integrals.

**Proposition 3.6.3.** *Let  $\gamma \in (0, 1)$ ,  $\beta \in \mathbb{A}_N$  and  $f \in C_b^{N+1-|\beta|}(\mathbb{R}^d, \mathbb{R})$ . Then for any  $\widehat{\mathbb{X}} \in \widehat{\mathcal{R}}_g(\mathbb{A}_N)$  over  $X$  one has*

$$\int_s^t F_b(X_r) d(\psi^* \widehat{\mathbb{X}})_r^\beta = \int_s^t F_g(X_r) d\widehat{\mathbb{X}}_r^\beta, \quad (3.6.9)$$

where  $F_b$  and  $F_g$  are respectively the lift of  $f$  with respect to the forest product and the quasi-shuffle product.

*Proof.* By hypothesis both integrals are well defined and they satisfy the following identities modulo  $o(|t - s|)$

$$\begin{aligned} \int_s^t F_b(X_r) d(\psi^* \widehat{\mathbb{X}})_r^\beta &= \sum_{m=0}^{N-|\beta|} \frac{1}{m!} \sum_{i_1, \dots, i_m=1}^d \partial_{i_1} \cdots \partial_{i_m} f(X_s) \langle \psi^* \widehat{\mathbb{X}}_{st}, B_+^\beta(\bullet_{e_{i_1}} \cdots \bullet_{e_{i_m}}) \rangle, \\ \int_s^t F_g(X_r) d\widehat{\mathbb{X}}_r^\beta &= \sum_{u \in W_\beta^\Gamma(\mathbb{A}_N)} \frac{\partial^u f(X_t)}{u!} \langle \widehat{\mathbb{X}}_{st}, u\beta \rangle. \end{aligned}$$

By definition of  $\psi$ , for any  $i_1, \dots, i_m \in \{1, \dots, d\}$  one has

$$\langle \psi^* \widehat{\mathbb{X}}_{st}, B_+^\beta(\bullet_{e_{i_1}} \cdots \bullet_{e_{i_m}}) \rangle = \langle \widehat{\mathbb{X}}_{st}, (e_{i_1} \widehat{\sqcup} \cdots \widehat{\sqcup} e_{i_m})\beta \rangle.$$

Using this identity and the combinatorial identity (3.4.18) we can write

$$\begin{aligned} &\sum_{m=0}^{N-|\beta|} \frac{1}{m!} \sum_{i_1, \dots, i_m=1}^d \partial_{i_1} \cdots \partial_{i_m} f(X_s) \langle \widehat{\mathbb{X}}_{st}, (e_{i_1} \widehat{\sqcup} \cdots \widehat{\sqcup} e_{i_m})\beta \rangle \\ &= \sum_{0 \leq |k| \leq N-|\beta|} \frac{\partial^k f(X_t)}{k_1! \cdots k_d!} \langle \widehat{\mathbb{X}}_{st}, (e_1^{\widehat{\sqcup} k_1} \widehat{\sqcup} \cdots \widehat{\sqcup} e_d^{\widehat{\sqcup} k_d})\beta \rangle. \end{aligned} \quad (3.6.10)$$

Applying finally the identity (3.6.3), we deduce that the right hand side of (3.6.10) becomes

$$\sum_{u \in W_\beta^\Gamma(\mathbb{A}_N)} \frac{\partial^u f(X_t)}{u!} \langle \widehat{\mathbb{X}}_{st}, u\beta \rangle.$$

Thereby obtaining the desired result.  $\square$

*An alternative proof of Theorem 3.6.2.* Thanks to the Theorem 3.5.13, the rough path  $\psi^* \widehat{\mathbb{X}} \in \widehat{\mathcal{R}}_b^\Gamma(\mathbb{A}_N)$  is a symmetrical bracket extension over  $X$ . Thus the formula (3.4.22) reads as

$$\varphi(X_t) = \varphi(X_s) + \sum_{i=1}^d \int_s^t \partial_i \Phi(X_r) d\psi^* \widehat{\mathbb{X}}_r^i + \sum_{n=2}^N \sum_{\beta \in \mathbb{A}_N: |\beta|=n} \frac{1}{\beta!} \int_s^t \partial^\beta \Phi(X_r) d\psi^* \widehat{\mathbb{X}}_r^\beta,$$

where the integrals are calculated from  $\psi^* \widehat{\mathbb{X}}$ . Using the equality (3.6.9), we conclude immediately.  $\square$

*Remark 3.6.4.* The main simplification lying in the formula (3.6.6) consists in the nature of  $\mathbb{X}$ . Indeed differently from the notion of symmetrical bracket extension there is a bijection between  $\hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  and  $\mathcal{R}_g^\Gamma(\mathbb{A}_N)$ . Thus to obtain a non trivial change of variable formula for one path  $X$  it is sufficient to define only a geometric rough path over  $X$  without any double definition as described in Theorem 3.4.10.

### 3.6.2 Extension of coordinates

In order to make all the change of variable formula (3.6.6) effective, it is reasonable to study the existence of a quasi-geometric rough path  $\mathbb{X} \in \hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  over one path  $X \in \mathcal{C}^\gamma(\mathbb{R}^d)$ . Looking at the Corollary 3.5.10, for any fixed path  $X' \in \mathcal{C}^\Gamma(\mathbb{A}_N)$  there exists  $\mathbb{X} \in \hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  over it. Thus the existence of a change of variable formula on  $X$  will depend on the choice of a path  $X'$  such that  $X'_{\mathbb{R}^A} = X$ , an *extension of coordinates*, and a choice of a rough path over  $X'$ . As a further simplification we can use the Proposition 3.5.8 and we reduce ourselves to study the existence of a geometric rough path  $\mathbb{X}' \in \mathcal{R}_g^\Gamma(\mathbb{A}_N)$  over  $X'$ .

Let us recall two basic results related to the extension of coordinates in the geometric setting. We fix  $(A, \Gamma)$  and  $(B, H)$  two finite weighted alphabets where  $A \subset B$ . The first result is related to the behaviour of Theorem 3.2.9 with respect to the extension of coordinates. For its proof see [TZ18, Cor. 5.8].

**Proposition 3.6.5.** *Let  $X \in \mathcal{C}^\Gamma(A)$  and  $Y \in \mathcal{C}^H(B)$  be an extension of coordinates and let  $\mathbf{X} \in \mathcal{R}_g^\Gamma(A)$ ,  $\mathbf{Y} \in \mathcal{R}_g^H(B)$  be two rough paths obtained respectively from the application of the Theorem 3.2.9 to  $X$  and  $Y$  following the construction of [TZ18, Thm. 5.7]. Then the restriction of  $\mathbf{Y}$  on  $T(A)$  coincides with  $\mathbf{X}$ .*

The second result is more explicit and tells us an explicit way to deduce  $\mathbf{X} \in \mathcal{R}_g^\Gamma(A)$  when we extend  $X$  by simply putting to zero its extra coordinates.

**Proposition 3.6.6.** *Let  $X \in \mathcal{C}^\Gamma(A)$  and  $\mathbf{X} \in \mathcal{R}_g^\Gamma(A)$  over  $X$ . Considering the function*

$$(X^0)_t^a = \begin{cases} (X_t)^a & \text{if } a \in A, \\ 0 & \text{for any other } a \in B \setminus A, \end{cases}$$

*then the function  $\mathbf{X}^0: [0, T]^2 \rightarrow T((B))$  which is given for any  $w \in W(B)$*

$$\langle \mathbf{X}^0, w \rangle = \begin{cases} \langle \mathbf{X}_{st}, w \rangle & \text{if } w \in W(A), \\ 0 & \text{if } w \text{ contains at least one letter of } B \setminus A, \end{cases}$$

*is a geometric rough path over  $X^0$ . We call  $\mathbf{X}_0 \in \mathcal{R}_g^H(B)$  the zero extension of  $\mathbf{X}$ .*

*Proof.* Without loss of generality we can assume  $B = A \sqcup \{b\}$  since  $A$  and  $B$  are finite. Then the proposition will follow by proving that  $\mathbf{X}^0$  satisfies the properties 1g), 2g) in the Definition 3.2.1 because the property 3g) is trivial. To prove 2g) we remark that for any word  $w \in W(B)$  containing the letter  $b$ , writing  $\nabla w = w^{(1)} \tilde{\otimes} w^{(2)}$  we have that  $b$  is contained in  $w^{(1)}$  or in  $w^{(2)}$  thus the we have  $\langle \mathbf{X}_{st}^0, w \rangle = \langle \mathbf{X}_{su}^0 \otimes \mathbf{X}_{ut}^0, w \rangle$

for any  $w \in W(B)$  because both sides are zero when  $w$  contains  $b$ . Finally for any couple of words  $w$  and  $u$  of  $W(B)$  containing at least in one of them the letter  $b$ , the sum  $w \sqcup u$  contains the letter  $b$  in all of its terms by definition, thereby obtaining  $\langle \mathbf{X}_{st}^0, w \sqcup u \rangle = \langle \mathbf{X}_{st}^0, w \rangle \langle \mathbf{X}_{st}^0, u \rangle$  because both sides are trivially zero.  $\square$

*Remark 3.6.7.* The explicit structure of the zero extension, even if it is trivial to define, represents a useful tool to define for any  $\{1, \dots, d\} \subset A \subset \mathbb{A}_N$  and  $\mathbf{X} \in \mathcal{R}_g^\Gamma(A)$  over  $X \in \mathcal{C}^\Gamma(A)$  a *canonical* object  $\mathbf{X}' \in \mathcal{R}_g^\Gamma(\mathbb{A}_N)$  over  $X$  and by means of the Proposition 3.5.8 a suitable  $\mathbb{X}' \in \hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  over  $X$  where we can apply the formula (3.6.6). We stress the fact that an equivalent formulation of the Proposition 3.6.6 in the quasi-shuffle case might be false because for any couple of words  $w$  and  $u$  of  $W(B)$  containing at least in one of them the letter  $b$ , the the sum  $w \hat{\sqcup} u$  contains words where the letter  $b$  might not be present and the function  $\mathbf{X}^0$  might not be a character.





# Chapter 4

## Rough Itô formulae for the one dimensional heat equation

### 4.1 Introduction

We consider  $u: [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ , the solution of the stochastic heat equation with additive noise on

$$\begin{cases} \partial_t u = \frac{1}{2} \partial_{xx} u + \xi \\ v(0, x) = 0. \end{cases} \quad (4.1.1)$$

where  $\xi = \partial W / \partial t \partial x$  is the space-time white noise over  $[0, +\infty) \times \mathbb{R}$  associated to  $W$ , a Brownian sheet on  $[0, +\infty) \times \mathbb{R}$  and defined on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . The existence and uniqueness of a weak solution for (4.1.1) is a classical result in the SPDE theory (see e.g. [DPZ14] or [Wal84]). Moreover the solution  $v$  can be described explicitly as the random field

$$u(t, x) = \int_0^t \int_{\mathbb{R}} G(t-s, x-y) W(ds, dy), \quad (4.1.2)$$

where for any  $t > 0$ ,  $x \in \mathbb{R}$ ,  $G(t, x) = (2\pi t)^{-1/2} e^{-x^2/2t}$  is the heat kernel and the integration  $W(ds, dy)$  is in the sense of the Wiener integral on  $L^2([0, +\infty) \times \mathbb{R})$ . For any  $x \in \mathbb{R}$  we will consider the process  $X_t = u(t, x)$ .

Using the standard properties of the Wiener integral we will easily show in the section 4.3 that the process  $X$  is a centred Gaussian process whose covariance function is explicitly given by

$$\mathbb{E}[X_t X_s] = \frac{1}{\sqrt{2\pi}} (\sqrt{t+s} - \sqrt{|t-s|}). \quad (4.1.3)$$

Looking at function (4.1.3) we deduce immediately that the trajectories of  $X$  are almost 1/4-Hölder continuous but  $X$  is not a Markov process neither its increments are independent or homogeneous. As recalled before, this hypothesis prevents us to define an Itô calculus on this process. Moreover looking at the existing literature of rough

paths, the covariance (4.1.3) cannot be included neither in the framework of Gaussian rough paths, as studied in [FV10a].

In this chapter we apply the resulting theory of rough change of variable formulae on  $X$  in order to recover some new and old Itô formulae. All these formulae will be obtained by providing the existence of some explicit one dimensional quasi-geometric rough paths over this process. As explained in the previous chapter, roughly speaking a one dimensional quasi-geometric rough path is a geometric rough path  $\mathbb{X}$  defined over an extended path of *variations*, that is a process  $\tilde{X} = (\tilde{X}^1, \dots, \tilde{X}^N): [0, T] \rightarrow \mathbb{R}^N$  for some  $N \geq 1$  such that  $\tilde{X}^1 = X$  and the others components  $\{\tilde{X}^i\}_{2 \leq i \leq N}$  (the  $i$ -th order variations) satisfy some algebraic and analytical properties of the standard quadratic variation up to the order  $N$ . For any  $\mathbb{X}$  satisfying these properties, the resulting rough integrals will satisfy some non trivial change of variable formula. Specifically for any quasi-geometric rough path  $\mathbb{X}$  over  $X$  and any sufficiently smooth function  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  we will obtain the general identity

$$\begin{aligned} \varphi(X_t) = \varphi(X_s) &+ \int_s^t \Phi'(X_r) d\mathbb{X}_r^1 + \frac{1}{2!} \int_s^t \Phi''(X_r) d\mathbb{X}_r^2 \\ &+ \frac{1}{3!} \int_s^t \Phi^{(3)}(X_r) d\mathbb{X}_r^3 + \frac{1}{4!} \int_s^t \Phi^{(4)}(X_r) d\mathbb{X}_r^4, \end{aligned} \quad (4.1.4)$$

where the terms in the right hand side of (4.1.4) are the rough integrals of the controlled rough paths lifting  $\{\varphi^{(i)}(X_t)\}_{1 \leq i \leq 4}$  with respect to  $\mathbb{X}^i$  the iterated integrals of the  $i$ -th order variation. Different choices of  $\mathbb{X}$  will lead to different values of these rough integrals and they will coincide with new notions of stochastic calculus for  $X$ .

The approaches we will develop in this chapter are four and for each one we will also provide a probabilistic description of the resulting rough integrals. First we can use the one dimensional nature of the process  $X$  to identify  $\mathbf{X}$ , the unique geometric rough path over  $\mathbf{X}$ , obtaining the so called *geometric formula*

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \Phi'(X_r) d\mathbf{X}_r, \quad (4.1.5)$$

where the rough integrals involving the higher order derivatives of  $\varphi$  in (4.1.4) are all zero. Moreover we can also describe the nature of the rough integral in (4.1.5) in terms of a Stratonovich-type sums, using some results contained in [BS10].

Passing to the other examples, we will construct two quasi-geometric rough paths  $\mathfrak{X}$  and  $\hat{\mathfrak{X}}$  such that when we plug them in (4.1.4) the third and the fourth order rough integral cancel out and the second integral can be described explicitly. These are called respectively the *Burdzy-Swanson formula* and the *Cheridito-Nualart formula*

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \Phi'(X_r) d\mathfrak{X}_r^1 + \frac{\kappa}{2} \int_s^t \varphi''(X_r) dB_r, \quad (4.1.6)$$

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \Phi'(X_r) d\hat{\mathfrak{X}}_r^1 + \frac{1}{2} \int_s^t \varphi''(X_r) d\sigma_r^2, \quad (4.1.7)$$

where  $B$  is a Brownian motion independent of  $X$ ,  $\sigma^2$  is the variance function  $\sigma_r^2 = \mathbb{E}[X_r^2]$  and the integrals in the formulae are an Itô integral and a Stieltjes one. The name of the identity (4.1.6) comes from the paper [BS10], where the same formula was obtained using different techniques but in that article the resulting equality holds only in law and not almost surely. On the other hand the formula (4.1.7) is named after the paper [CN05], where the authors obtain an analogous change of variable formula starting from the fractional Brownian motion with Hurst parameter  $H \in (0, 1/2)$ . Finally we recall some basic properties of the process  $X$  from [Swa07] to deduce the existence of a quasi-geometric rough path  $\mathcal{X}$  where there is the fourth order rough integral which becomes a standard Lebesgue integral. We call this identity the *quartic variation formula*

$$\begin{aligned} \varphi(X_t) = & \varphi(X_s) + \int_s^t \Phi'(X_r) d\mathcal{X}_r^1 + \frac{1}{2!} \int_s^t \Phi''(X_r) d\mathcal{X}_r^2 \\ & + \frac{1}{3!} \int_s^t \Phi^{(3)}(X_r) d\mathcal{X}_r^3 + \frac{6}{4!\pi} \int_s^t \varphi^{(4)}(X_r) dr. \end{aligned} \quad (4.1.8)$$

The sum of the remaining rough integrals in the right hand side of (4.1.8) can be also associated to the limit of a compensated Riemann sum introduced in the recent paper [CP19], where the authors generalise the usual Föllmer calculus [Föl81] in a generalised context for processes admitting higher order variation.

The present chapter will be organised as follows. After recalling in the section 2 the explicit rough change of variable formula in the geometric and the quasi-geometric setting we provide a general criterion to prove the existence of a geometric rough path. Moreover the section 3 will be devoted to provide all the explicit constructions reported above. The last section will finally study an interesting identity between some generalised Stratonovich integrals and again the rough integration with respect to the canonical geometric extension of a one dimensional rough path (see Proposition 4.2.7). In this way we can give a simplified proof of the Stratonovich type formula obtained in [BS10] by means of this general criterion.

All techniques presented here are based on the Gaussian structure of  $X$ . Thus we can be easily adapt the propositions presented here to obtain the same identities on a fractional Brownian motion with Hurst parameter  $1/4$  or, looking at  $X$  as a bifractional motion (see [RT06]), we can extend this results on a larger class of Gaussian process. However, we chose to interpret  $X$  as the time evolution of the equation (4.1.1) because is the simplest time evolution of generic family of stochastic partial differential equations

$$\begin{cases} \partial_t u = \frac{1}{2} \partial_{xx} u + f(u) + \sigma(u) \xi \\ u(0, x) = 0 \end{cases}$$

where  $f, \sigma: \mathbb{R} \rightarrow \mathbb{R}$  are Lipschitz functions. It is a well known fact (see [Wal84]) that these equations admits a unique solution  $u$  which is almost surely continuous in  $t$  but a general theory to study the properties of the process  $t \rightarrow u(t, x)$  for fixed  $x$  is still missing and the techniques developed here could be also the starting point to study this interesting family of stochastic processes.

## 4.2 One dimensional change of variable formula

The scope of this section is to recall the change of variable formula on continuous paths  $Y: [0, T] \rightarrow \mathbb{R}^d$  when  $d = 1$  for quasi-geometric rough path. This hypothesis simplifies drastically many general expressions contained in the previous chapter so we will restate briefly the previous results to improve the readability of this chapter. An explicit criterion of the existence for rough paths will allow to define rigorously some random rough path  $\mathbf{X}$  over the process  $X$ .

### 4.2.1 General results

Let us fix some notation from the previous chapter. In what follows we start from a continuous function  $Y: [0, T] \rightarrow \mathbb{R}$  belonging to  $\mathcal{C}^\gamma(\mathbb{R})$ , the space of  $\gamma$ -Hölder continuous function for some  $\gamma \in (0, 1)$ . Then for any choice of the parameter  $\gamma$  we introduce the weighted alphabet  $(\mathbb{A}_N, \Gamma)$  where  $N = \lfloor \gamma^{-1} \rfloor$ ,  $\mathbb{A}_N = \{1, \dots, N\}$  and for any  $m \in \mathbb{A}_N$  we set  $\Gamma_m = m\gamma$ . Analogously for any subset  $A \subset \mathbb{A}_N$  we will henceforth keep the same weight  $\Gamma$  to define the weighted alphabet  $(A, \Gamma)$ . Subsequently we can consider  $W(\mathbb{A}_N)$  the set of words generated from the alphabet  $\mathbb{A}_N$  union the empty word  $\mathbf{1}$  and the the tensor algebra  $T(\mathbb{A}_N)$ , interpreted as the free vector space generated from the  $W(\mathbb{A}_N)$ . Four operations can be defined on the set  $T(\mathbb{A}_N)$ :

- The concatenation product  $\diamond: T(\mathbb{A}_N) \times T(\mathbb{A}_N) \rightarrow T(\mathbb{A}_N)$ , defined by extending linearly the juxtaposition of words, that is for any couple of words  $v = a_1 \cdots a_n$ ,  $w = b_1 \cdots b_m \in W(\mathbb{A}_N)$  one has

$$v \diamond w := a_1 \cdots a_n b_1 \cdots b_m.$$

- The deconcatenation coproduct  $\nabla: T(\mathbb{A}_N) \rightarrow T(\mathbb{A}_N) \otimes T(\mathbb{A}_N)$  defined by the relations

$$\begin{aligned} \nabla \mathbf{1} &:= \mathbf{1} \otimes \mathbf{1}, \\ \nabla w &:= \sum_{i=0}^m a_1 \cdots a_i \otimes a_{i+1} \cdots a_m, \end{aligned}$$

for any word  $w = a_1 \cdots a_m \in W(\mathbb{A}_N)$  and extended linearly.

- The shuffle product  $\sqcup: T(\mathbb{A}_N) \times T(\mathbb{A}_N) \rightarrow T(\mathbb{A}_N)$ , defined by extending linearly the recursive relations

$$\begin{aligned} \mathbf{1} \sqcup v &:= v \sqcup \mathbf{1} := v, \\ av \sqcup bw &:= (v \sqcup bw)a + (av \sqcup w)b. \end{aligned}$$

- The quasi-shuffle product  $\hat{\sqcup}: T(\mathbb{A}_N) \times T(\mathbb{A}_N) \rightarrow T(\mathbb{A}_N)$ , defined on the other hand by extending linearly the recursive relations

$$\begin{aligned} \mathbf{1} \hat{\sqcup} v &:= v \hat{\sqcup} \mathbf{1} := v, \\ av \hat{\sqcup} bw &:= (v \hat{\sqcup} bw)a + (av \hat{\sqcup} w)b + (v \hat{\sqcup} w)[a, b], \end{aligned}$$

where the operation  $[\cdot, \cdot]: \mathbb{R}^{\mathbb{A}_N} \times \mathbb{R}^{\mathbb{A}_N} \rightarrow \mathbb{R}^{\mathbb{A}_N}$  is the unique bilinear function such that for any  $i, j \in \{1, \dots, N\}$  we have

$$[i, j] = [j, i] := \begin{cases} i + j & \text{if } i + j \leq N, \\ 0 & \text{otherwise.} \end{cases} \quad (4.2.1)$$

Looking at the definition of these operations we remark that  $\nabla$  and  $\sqcup$  can be defined on  $T(A)$ , the tensor algebra generated from  $W(A)$ , the set of word built from a generic subset  $A \subset \mathbb{A}_N$ , whereas the product  $\hat{\sqcup}$  can be only defined on  $T(\mathbb{A}_N)$  due to the definition of the additional operation  $[\cdot, \cdot]$  defined in (4.2.1). Furthermore for any  $w \in W(\mathbb{A}_N)$  of the form  $w = i_1 \cdots i_n$ , we introduce several quantities associated to  $w$ : the word length  $|w| = n$ , the word homogeneity  $|w|_\Gamma = \sum_{k=1}^n i_k$  and the the word factorial  $w! = \prod_{l=1}^n i_l!$ . Denoting by  $T((A))$  and  $W((A))$  respectively the dual vector space of  $T(A)$ , we give a simplified version of the Definitions 3.2.1 and 3.5.10 in this context to recall the notion of rough path:

**Definition 4.2.1.** A function  $\mathbf{Y}: [0, T]^2 \rightarrow T((A))$  is called a geometric rough path if it satisfies the following conditions:

- 1) for every  $w_1, w_2 \in T(A)$   $\langle \mathbf{Y}_{st}, w_1 \sqcup w_2 \rangle = \langle \mathbf{Y}_{st}, w_1 \rangle \langle \mathbf{Y}_{st}, w_2 \rangle$ ,
- 2) for every  $s, t, u \in [0, T]^3$  and  $w \in T(A)$   $\mathbf{Y}_{st} = \langle \mathbf{Y}_{su} \otimes \mathbf{Y}_{ut}, \nabla w \rangle$ ,
- 3) for every word  $w \in W(A)$   $\sup_{s \neq t} |\langle \mathbf{Y}_{st}, w \rangle| / |t - s|^{\gamma|w|_\Gamma} < +\infty$ .

**Definition 4.2.2.** A function  $\mathbb{Y}: [0, T]^2 \rightarrow T((\mathbb{A}_N))$  is called a one dimensional quasi-geometric rough path if it satisfies for every  $w_1, w_2 \in T(\mathbb{A}_N)$  the identity

$$\langle \mathbb{Y}_{st}, w_1 \hat{\sqcup} w_2 \rangle = \langle \mathbb{Y}_{st}, w_1 \rangle \langle \mathbb{Y}_{st}, w_2 \rangle$$

and the conditions 2), 3) in the Definition 4.2.1.

We will denote respectively the sets of geometric rough path and the one dimensional quasi-geometric rough paths by  $\mathcal{R}_g^\Gamma(A)$  and  $\hat{\mathcal{R}}_g(\mathbb{A}_N)$  and the elements of both sets are denoted by  $\mathcal{R}_g(\mathbb{A}_N)$  when  $A = \mathbb{A}_N$ . Recalling our initial function  $Y$ , we say that a geometric or quasi-geometric rough path  $\mathbf{Y}$  is over  $Y$  if one has the identity  $\langle \mathbf{Y}_{st}, 1 \rangle = Y_t - Y_s$ . Moreover if and a path  $Z: [0, T] \rightarrow \mathbb{R}$  satisfies  $\langle \mathbf{Y}_{st}, n \rangle = Z_t - Z_s$  for some  $2 \leq n \leq N$  we call  $Z$  the  $n$ -th order variation of  $Y$ .

The choice of a specific rough path  $\mathbf{Y} \in \mathcal{R}_g(\mathbb{A}_N)$  over  $Y$  allows us to consider form any  $m \in \mathbb{A}_N$  and any sufficiently smooth function  $\psi: \mathbb{R} \rightarrow \mathbb{R}$  the rough integral

$$\int_s^t \Psi(Y_r) d\mathbf{Y}_r^m,$$

where  $\Psi$  is the lifting of the function  $\psi$  (this operation will be always denoted with a capital letter) defined in (3.3.14a), (3.3.14b). The general expression defining this function is given in the Proposition 3.3.6 and we will recall it in this particular case. In what follows we will denote by  $C_b^m(\mathbb{R}, \mathbb{R})$ , the space of  $C^m$  functions whose all derivatives are bounded up to the order  $m$ .

**Proposition 4.2.3.** *Let  $\mathbf{Y} \in \mathcal{R}_g(\mathbb{A}_N)$  be a quasi-geometric rough path over  $Y$ ,  $m \in \mathbb{A}_N$  and  $\psi \in C_b^{N-m+1}(\mathbb{R}, \mathbb{R})$ . Then for any sequence of partitions  $\mathcal{P}_n$  of  $[0, t]$  such that its mesh-size  $|\mathcal{P}_n| \rightarrow 0$  one has the identity*

$$\int_0^t \Psi(Y_r) d\mathbf{Y}_r^m = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \sum_{u \in W_m^\Gamma(\mathbb{A}_N)} \frac{\psi^{(|u|_\Gamma)}(Y_u)}{u!} \langle \mathbf{Y}_{uv}, u m \rangle, \quad (4.2.2)$$

where  $W_m^\Gamma(\mathbb{A}_N) := \{u \in W^\Gamma(\mathbb{A}_N) : |u|_\Gamma \leq N - m\}$ . The rough integral is a  $\gamma$ -Hölder function whose value in 0 is 0. Moreover there exists  $\varepsilon > 0$  such that for any  $t, s \in [0, T]$  we have the approximated equality

$$\int_s^t \Psi(Y_r) d\mathbf{Y}_r^m = \sum_{u \in W_m^\Gamma(\mathbb{A}_N)} \frac{\psi^{(|u|_\Gamma)}(Y_s)}{u!} \langle \mathbf{Y}_{st}, u m \rangle + O(|t - s|^{1+\varepsilon}). \quad (4.2.3)$$

Using the notion of rough integral, we can easily write the two change of variable formulae obtained in Proposition 3.1.1 and Theorem 3.6.2 in this simplified setting.

**Proposition 4.2.4.** *Let  $N = \lfloor \gamma^{-1} \rfloor$  and  $\varphi \in C_b^{N+1}(\mathbb{R}, \mathbb{R})$ . For any one dimensional  $\mathbf{Y} \in \mathcal{R}_g^\Gamma(\{1\})$  over  $Y$  one has*

$$\varphi(Y_t) = \varphi(Y_s) + \int_s^t \Phi'(Y_r) d\mathbf{Y}_r^1. \quad (4.2.4)$$

**Proposition 4.2.5.** *Let  $N = \lfloor \gamma^{-1} \rfloor$  and  $\varphi \in C_b^{N+1}(\mathbb{R}, \mathbb{R})$ . For any  $\mathbb{Y} \in \hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  over the path  $Y$  one has*

$$\varphi(Y_t) = \varphi(Y_s) + \sum_{n=1}^N \frac{1}{n!} \int_s^t \Phi^{(n)}(Y_r) d\mathbb{Y}_r^n. \quad (4.2.5)$$

*Remark 4.2.6.* Unlike in the previous chapter, we will consider only the change of variable formulae for geometric and quasi-geometric rough paths without using Theorem 3.4.10 and the symmetrical bracket extension. This choice is made because all our examples are covered by this hypothesis. More generally the existence of an explicit example of a bracket extension over a continuous path  $Y$  which is not a quasi-geometric rough path is still unknown in the literature.

In order to apply the formulae (4.2.4) and (4.2.5) it is necessary to start from some rough path over  $Y$ . The hypothesis  $d = 1$  really simplifies also this problem in the geometric setting, yielding a unique candidate.

**Proposition 4.2.7.** *For any  $Y \in \mathcal{C}^\gamma(\mathbb{R})$ , there exists a unique  $\mathbf{Y} \in \mathcal{R}_g^\Gamma(\{1\})$  over  $Y$ , which is given explicitly by the unique linear function  $\mathbf{Y}: [0, T]^2 \rightarrow T((\mathbb{R}))$  defined for any  $n \geq 0$*

$$\langle \mathbf{Y}_{st}, 1^{\circ n} \rangle = \frac{(Y_t - Y_s)^n}{n!}. \quad (4.2.6)$$

We call  $\mathbf{Y}$  the canonical geometric extension of  $Y$ .

*Proof.* The function  $\mathbf{Y}$  is well defined because  $T(\mathbb{R})$  contains only words with the letter 1. We can check directly that  $\mathbf{Y}$  is a geometric rough path as an elementary exercise. We give here the proof for the sake of completeness. Since the Hölder property

$$\sup_{s \neq t} \frac{|\langle \mathbf{Y}_{st}, 1^{\diamond n} \rangle|}{|t - s|^{n\gamma}} < \infty$$

is trivial, we simply check the Chen relations and the character property. We fix  $n \geq 0$  and applying the definition of the coproduct  $\nabla$  we obtain for any  $s, u, t \in [0, T]$

$$\langle \mathbf{Y}_{su} \otimes \mathbf{Y}_{ut}, \nabla 1^{\diamond n} \rangle = \sum_{k=0}^n \frac{(Y_u - Y_s)^k}{k!} \frac{(Y_t - Y_u)^{n-k}}{(n-k)!} = \frac{(Y_t - Y_s)^n}{n!},$$

as a trivial consequence of Newton's binomial theorem. Choosing two integers  $m$  and  $n$  by definition of the shuffle product we have the trivial identity

$$1^{\diamond n} \sqcup 1^{\diamond m} = \frac{(m+n)!}{m!n!} 1^{\diamond(n+m)} \quad (4.2.7)$$

Therefore for any  $s, t \in [0, T]$  we have

$$\langle \mathbf{Y}_{st}, 1^{\diamond n} \sqcup 1^{\diamond m} \rangle = \langle \mathbf{Y}_{st}, 1^{\diamond n} \rangle \langle \mathbf{Y}_{st}, 1^{\diamond m} \rangle.$$

The function  $\mathbf{Y}$  is also unique because for any  $\mathbf{Y}' \in \mathcal{R}_g^\gamma(\mathbb{R})$  over  $Y$  we can use the multiplicative property of  $\mathbf{Y}'$  to obtain the definition (4.2.6) from (4.2.7).  $\square$

Combining the equation (4.2.6) with the explicit expression of the rough integral (4.2.2), we deduce immediately that the rough integral in the right hand side of (4.2.4) is given by

$$\int_0^t \Phi'(Y_r) d\mathbf{Y}_r^1 = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \sum_{l=1}^N \varphi^{(l)}(Y_u) \frac{(Y_v - Y_u)^l}{l!}, \quad (4.2.8)$$

where  $\mathcal{P}_n$  is an arbitrary sequence of partition of  $[0, t]$  such that its mesh-size  $|\mathcal{P}_n| \rightarrow 0$ . We recall finally two useful deterministic propositions to prove the existence of some rough path  $\mathbf{Y}' \in \mathcal{R}_g^\Gamma(\mathbb{A}_N)$  (defined on a wider alphabet) over  $Y$ . The first one is given by simplifying the Proposition 3.6.6.

**Proposition 4.2.8.** *Let  $(A, \Gamma)$  be a weighted alphabet such that  $\{1\} \subset A \subset \mathbb{A}_N$  and let  $\mathbf{Y} \in \mathcal{R}_g^\Gamma(A)$  be over  $Y$ . Then the function  $\mathbf{Y}^0: [0, T]^2 \rightarrow T((\mathbb{A}_N))$  defined as*

$$\langle \mathbf{Y}_{st}^0, w \rangle = \begin{cases} \langle \mathbf{Y}_{st}, w \rangle & \text{if } w \in W(A), \\ 0 & \text{if } w \text{ contains at least one letter of } \mathbb{A}_N \setminus A, \end{cases} \quad (4.2.9)$$

belongs to  $\mathcal{R}_g^\Gamma(\mathbb{A}_N)$ . We call  $\mathbf{Y}^0$  the zero extension of  $\mathbf{Y}$ .



We usually apply the Proposition 4.2.8 when we want to extend a non trivial  $\mathbf{Y} \in \mathcal{R}_g^\Gamma(A)$  up to  $\mathcal{R}_g^\Gamma(\mathbb{A}_N)$ . Finally in order to pass from elements of  $\mathcal{R}_g^\Gamma(\mathbb{A}_N)$  to elements of  $\hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$ , we recall the isomorphisms  $\exp, \log: T(\mathbb{A}_N) \rightarrow T(\mathbb{A}_N)$ , which are given explicitly for any word  $w \in W(\mathbb{A}_N)$  by

$$\exp(w) := \sum_{I \in C(|w|)} \frac{1}{I!} [w]_I, \quad \log(w) := \sum_{I \in C(|w|)} \frac{(-1)^{|w|-l(I)}}{p(I)} [w]_I, \quad (4.2.10)$$

where the sum runs over  $C(|w|)$ , the set of compositions  $I$  of the word's length  $|w|$ . We denote by  $\log^*, \exp^*: T(\mathbb{A}_N) \rightarrow T(\mathbb{A}_N)$  the adjoints of  $\exp, \log$ . Looking at Proposition 3.5.8 in this case we obtain:

**Proposition 4.2.9.** *For any  $\mathbf{Y} \in \mathcal{R}_g^\Gamma(\mathbb{A}_N)$  over  $Y$  the function  $\log^* \mathbf{Y}$  is a one dimensional quasi-geometric rough path over  $Y$ .*

Thanks to the Proposition (4.2.9) and 4.2.8, we can reduce the construction of a one-dimensional quasi-geometric rough path  $\mathbb{Y} \in \hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_N)$  to the construction of a geometric rough path  $\mathbf{Y} \in \mathcal{R}_g^\Gamma(A)$  on a smaller alphabet  $A \subset \mathbb{A}_N$  and later defining  $\mathbb{Y} = \log^* \mathbf{Y}^0$ .

## 4.2.2 An existence criterion for rough paths

In order to show the existence of some specific geometric rough paths  $\mathbf{Y} \in \mathcal{R}_g^\Gamma(A)$  over  $Y$  we will need to generalise the deterministic result obtained in Proposition 4.2.7. As explained in its proof, the algebraic identity between  $1^{\circ n}$  and  $1^{\sqcup n}$  in (4.2.7) allows to fix uniquely the value of the canonical geometric extension. Equivalently we could ask if there exists a “minimal” proper subset of words  $B \subset W(\mathbb{A}_N)$  such that every word  $w \in W(\mathbb{A}_N)$  can be written as a *polynomial* using the shuffle product. This is exactly the property of the *Lyndon words*. In what follows for all subsets  $A \subset \mathbb{A}_N$  we shall denote by  $(A, \preceq)$  the set  $A$  endowed with the total order of  $\mathbb{A}_N = \{1, \dots, N\}$ . By means of the standard lexicographic order induced by  $(A, \preceq)$  we can define a total order on  $W(A)$ . We will denote it by the same notation  $(W(A), \preceq)$ .

**Definition 4.2.10.** Let  $A \subset \mathbb{A}_N$ . A non-empty word  $w \in W(A)$  is said to be a *Lyndon word* if for any decomposition  $w = uv$  where  $v$  and  $u$  are non empty words one has  $w \prec vu$ . We denote by  $L(A)$  the set of all Lyndon words upon the alphabet  $A$ .

For instance if  $A = \{1, 2\}$  the word  $w = 12$  is a Lyndon word because we have easily  $w \prec 21$  but the word  $v = 211$  is not a Lyndon word because we can decompose  $v = (2)(11)$  and we have  $112 \prec v$ . Since the alphabet  $A$  has also a weight  $\Gamma$ , we introduce the sets

$$W^\Gamma(\mathbb{A}_N) := \{v \in W(\mathbb{A}_N) : |v|_\Gamma \leq \hat{\gamma}^{-1}\}, \quad W^m(\mathbb{A}_N) := \{v \in W(\mathbb{A}_N) : |v| \leq m\}, \quad (4.2.11)$$

$$L^\Gamma(A) := \{w \in L(A) : |w|_\Gamma \leq \gamma^{-1}\}, \quad L^m(A) := \{w \in L(A) : |w| \leq m\}. \quad (4.2.12)$$

By definition of  $|\cdot|_\Gamma$  we have trivially  $L^\Gamma(A) \subset L^N(A)$  where  $N = \lfloor \gamma^{-1} \rfloor$ . The study of Lyndon words is a classical topic in combinatorics (see e.g. [Lot97]). One of their main properties is then related to the possibility of generating the shuffle algebra  $(T(A), \sqcup)$ . We resume this property in a single statement. For its proof see [Lyn54, Thm 5.1.5] and [MR89].

**Theorem 4.2.11.** *For any word  $w \in W(A)$  there exists a unique writing of  $w$  as*

$$w = l_1^{\circ i_1} \cdots l_n^{\circ i_n}, \tag{4.2.13}$$

where  $l_1, \dots, l_n \in L(A)$  satisfy  $l_1 \succ \cdots \succ l_n$  and  $i_1, \dots, i_n \geq 1$ . We call this writing the Lyndon decomposition of  $w$ . Moreover for any word  $w$  written as (4.2.13) one has

$$w = \frac{l_1^{\sqcup i_1} \sqcup \cdots \sqcup l_n^{\sqcup i_n}}{i_1! \cdots i_n!} - \sum_{u \prec w} \alpha_u u \tag{4.2.14}$$

for some sequence of natural numbers  $\alpha_u$  depending on  $w$ . Therefore every word of  $W(A)$  can expressed as a polynomial in terms of the Lyndon words  $L(A)$ .

*Remark 4.2.12.* Theorem 4.2.11 does not only prove that  $T(A)$  is generated by the polynomials of the words in  $L(A)$  but it shows also an iterative method for writing the coefficients. Since the function  $|\cdot|_\Gamma$  is additive on the concatenation product we deduce from (4.2.13) that every word belonging to  $W^\Gamma(A)$  can be written as a concatenation of Lyndon words belonging only to  $L^\Gamma(A)$ . Therefore we deduce recursively from (4.2.14) that  $T^\Gamma(A)$ , the linear space generated by  $W^\Gamma(A)$ , coincides with the polynomials using the shuffle product whose indeterminates are the Lyndon words of  $L^\Gamma(A)$ . Using the definition of the shuffle product we deduce that every  $u \prec w$  in the relation 4.2.14 satisfies also  $|u|_\Gamma = |w|_\Gamma$ . These considerations allow to rewrite (4.2.14) by saying that for any  $w \in W^\Gamma(A)$ , writing  $L^\Gamma(A) = \{l_1, \dots, l_M\}$  and considering  $K$ , the set of all multi-indices  $J = (j_1, \dots, j_M) \in K$  such that  $\sum_{k=1}^M j_k |l_k|_\Gamma = |w|_\Gamma$ , there exist a unique sequence  $\{\beta_J\}_{J \in K}$  such that

$$w = \sum_{J \in K} \beta_J l_1^{\sqcup j_1} \sqcup \cdots \sqcup l_M^{\sqcup j_M}. \tag{4.2.15}$$

As a trivial example of the identity (4.2.15) when  $A = \{a\}$ , we have the identity

$$a^{\circ n} = a^{\sqcup n} / n!,$$

which was used in Proposition 4.2.7.

The properties of the Lyndon words allow to prove the following existence theorem. In what follows we will denote by  $\mathcal{L}^\Gamma(A)$  the linear space of  $L^\Gamma(A)$ ,  $\mathcal{L}^\Gamma((A))$  the dual of  $\mathcal{L}^\Gamma(A)$ . Moreover we denote by  $T^\Gamma(A)$  the linear space of  $W^\Gamma(A)$  and  $T^\Gamma((A))$  the dual of  $T^\Gamma(A)$ .

**Lemma 4.2.13.** *Let  $(A, \Gamma)$  be a weighted alphabet such that  $A \subset \mathbb{A}_N$ . For any function  $\tilde{\mathbb{Y}}: [0, T]^2 \rightarrow \mathcal{L}^\Gamma((A))$  there exists a unique  $\mathbb{Y}: [0, T]^2 \rightarrow T^\Gamma((A))$  extending  $\tilde{\mathbb{Y}}$  such that*

$$\langle \mathbb{Y}_{st}, v \sqcup w \rangle = \langle \mathbb{Y}_{st}, v \rangle \langle \mathbb{Y}_{st}, w \rangle, \tag{4.2.16}$$

for any  $w, v \in W^\Gamma(A)$  satisfying  $v \sqcup w \in T^\Gamma(A)$ . Moreover if for every  $s, t, u \in [0, T]^3$  and  $u \in L_\Gamma(A)$  one has  $\langle \mathbb{Y}_{st}, u \rangle = \langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla u \rangle$ , then for any  $w \in W^\Gamma(A)$  we have

$$\langle \mathbb{Y}_{st}, w \rangle = \langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla w \rangle. \quad (4.2.17)$$

*Proof.* Writing  $L^\Gamma(A) = \{l_1, \dots, l_M\}$  this result is a direct consequence of the identity (4.2.15) which holds for any  $w \in W^\Gamma(A)$ . Thus if  $\mathbb{Y}: [0, T]^2 \rightarrow \mathcal{L}^\Gamma((A))$  satisfies (4.2.16) and  $\langle \mathbb{Y}_{st}, l_i \rangle = \langle \tilde{\mathbb{Y}}_{st}, l_i \rangle$  for any  $i = 1, \dots, M$  we deduce that  $\mathbb{Y}$  must satisfy for any  $s, t \in [0, T]^2$

$$\langle \mathbb{Y}_{st}, w \rangle = \sum_{J \in K} \beta_J \langle \tilde{\mathbb{Y}}_{st}, l_1 \rangle^{j_1} \dots \langle \tilde{\mathbb{Y}}_{st}, l_M \rangle^{j_M}. \quad (4.2.18)$$

Since the shuffle product is commutative, for any  $w, v \in W^\Gamma(A)$  such that  $|v|_\Gamma + |w|_\Gamma \leq \gamma^{-1}$  defining  $\mathbb{Y}$  as (4.2.18), we obtain by construction the property (4.2.16). Let us prove the property (4.2.17). Using the decomposition (4.2.15) given before one has for any  $w \in W^\Gamma(A)$

$$\langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla w \rangle = \sum_{J \in K} \beta_J \langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla(l_1^{\sqcup j_1} \sqcup \dots \sqcup l_M^{\sqcup j_M}) \rangle. \quad (4.2.19)$$

By construction the deconcatenation coproduct  $\nabla$  is an algebra morphism. Therefore for any  $j_1, \dots, j_m \geq 1$  we have

$$\nabla(l_1^{\sqcup j_1} \sqcup \dots \sqcup l_M^{\sqcup j_M}) = (\nabla l_1)^{\sqcup' j_1} \sqcup' \dots \sqcup' (\nabla l_M)^{\sqcup' j_M},$$

where the product  $\sqcup': (T(A) \otimes T(A)) \otimes (T(A) \otimes T(A)) \rightarrow T(A) \otimes T(A)$  is given by

$$(a \otimes b) \sqcup' (c \otimes d) := a \sqcup c \otimes b \sqcup d.$$

Since the function  $\mathbb{Y}$  satisfies the equation (4.2.16), then for any  $j_1, \dots, j_m \geq 1$  we obtain by hypothesis

$$\begin{aligned} \langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla(l_1^{\sqcup j_1} \sqcup \dots \sqcup l_M^{\sqcup j_M}) \rangle &= \langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla l_1 \rangle^{j_1} \dots \langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla l_M \rangle^{j_M} \\ &= \langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla l_1 \rangle^{j_1} \dots \langle \mathbb{Y}_{su} \otimes \mathbb{Y}_{ut}, \nabla l_M \rangle^{j_M} = \langle \mathbb{Y}_{st}, l_1 \rangle^{j_1} \dots \langle \mathbb{Y}_{st}, l_1 \rangle^{j_M}. \end{aligned}$$

Then the right hand side of (4.2.19) becomes equal to  $\langle \mathbb{Y}_{st}, w \rangle$  and we conclude.  $\square$

We conclude the section by applying this criterion to a random process  $\mathbf{X}: [0, T]^2 \rightarrow T^\Gamma((A))$  in order to deduce the existence of a *random rough paths*, that is a random variable whose are a.s. in  $\mathcal{R}_g(A)$ . The probabilistic estimates we will present are a generalisation of Kolmogorov's continuity Theorem (see [RY04, Thm 2.1]), translated in the context of rough paths. This result can be also seen as an extension of [FH14, Thm 3.1] for any Hölder regularity.

**Proposition 4.2.14** (Kolmogorov's criterion on Lyndon words). *Let  $\gamma \in (0, 1)$ ,  $(A, \Gamma)$  a weighted alphabet  $A \subset \mathbb{A}_N$  and  $\mathbf{X}: [0, T]^2 \rightarrow T^\Gamma((A))$  a given process satisfying a.s. the properties (4.2.16) and (4.2.17). Supposing that there exists  $q > N$  and  $C > 0$  such that*

$$\frac{1}{N} - \frac{1}{q} > \frac{1}{N+1}, \quad \mathbb{E}|\langle \mathbf{X}_{st}, u \rangle|^q \leq C|t - s|^{(q|u|_\Gamma)/N}. \quad (4.2.20)$$

for any  $u \in L^\Gamma(A)$ . Then there exists a modification of  $\mathbf{X}$  (also denoted by  $\mathbf{X}$ ) such that for any  $w \in W^\Gamma(A)$  the function  $(s, t) \rightarrow \langle \mathbf{X}_{st}, w \rangle$  is continuous and it satisfies

$$\mathbb{E} \left| \sup_{s \neq t} \frac{|\langle \mathbf{X}_{st}, w \rangle|}{|t - s|^{\gamma|w|_\Gamma}} \right|^{q/|w|_\Gamma} < \infty. \quad (4.2.21)$$

Therefore by means of the extension Theorem (see Theorem 3.2.7) there exists a unique random rough path  $\mathbf{X}' \in \mathcal{R}_g^\Gamma(A)$  over  $\mathbf{X}$ .

*Proof.* We will prove the result by induction on  $|w|$ , the word's length. Moreover we can fix  $T = 1$  without loss of generality. In case  $|w| = 1$  the word  $w$  is a letter  $a \in A \subset \mathbb{A}_N$  and since every letter  $a$  belongs to  $L^\Gamma(A)$  by hypothesis we deduce

$$\mathbb{E} |\langle \mathbf{X}_{st}, a \rangle|^q \leq C |t - s|^{(q|a|_\Gamma)/N}. \quad (4.2.22)$$

Then writing  $\langle \mathbf{X}_{st}, a \rangle$  as the increment of a path  $X^a: [0, 1] \rightarrow \mathbb{R}$  we can apply directly the classical Kolmogorov's continuity criterion (see [RY04, Thm 2.1]) to obtain the property (4.2.21) because  $1/(N + 1) < \gamma < 1/N$ . Supposing the result true for any word of length equal to  $n$ , we will first show that the result holds for any  $u \in L^\Gamma(A)$  such that  $|u| = n + 1$  by adapting the proof of the Kolmogorov's continuity criterion in this context. Let  $D_n$  be the set of integer multiples of  $2^{-n}$  in  $[0, 1]$ . The cardinality of  $D_n$  is  $2^n$  and the smallest distance between its elements is  $2^{-n} =: |D_n|$ . In order to show the existence of a modification of  $F(s, t) := \langle \mathbf{X}_{st}, u \rangle$  satisfying (4.2.21) it is sufficient to define  $F$  on every  $s, t \in \bigcup_{n \geq 1} D_n$  filling the remaining values by continuity. Introducing the random variable  $K_n := \sup_{t \in D_n} |F(t, t + 2^{-n})|$ , by hypothesis (4.2.20) we have

$$\mathbb{E} |K_n|^q \leq \sum_{t \in D_n} \mathbb{E} |F(t, t + 2^{-n})|^q \leq \#D_n |D_n|^{(q|u|_\Gamma)/N} = C |D_n|^{(q|u|_\Gamma)/(N-1)}.$$

Let  $s < t \in \bigcup_{n \geq 1} D_n$  such that  $|D_{m+1}| < t - s < |D_m|$ . Using the properties of the partition  $D_n$  we can express the interval  $[s, t]$  as a partition

$$s = \tau_0 < \dots < \tau_M = t, \quad (4.2.23)$$

where each  $\tau_i \in D_n$  for  $n \geq m + 1$  and for each fixed  $n \geq m + 1$  there are at most two intervals with values  $\tau_i \in D_n$ . Using the property (4.2.17) and the notation  $\nabla u = u \otimes \mathbf{1} + \mathbf{1} \otimes u + u^{(1)} \otimes u^{(2)}$  we obtain

$$\begin{aligned} |F(s, t)| &= \left| \sum_{i=0}^{M-1} F(\tau_i, \tau_{i+1}) + \langle \mathbf{X}_{s\tau_i}, u^{(1)} \rangle \langle \mathbf{X}_{\tau_i\tau_{i+1}}, u^{(2)} \rangle \right| \\ &\leq \sum_{i=0}^{M-1} |F(\tau_i, \tau_{i+1})| + \sum_{i=0}^{M-1} |\langle \mathbf{X}_{s\tau_i}, u^{(1)} \rangle \langle \mathbf{X}_{\tau_i\tau_{i+1}}, u^{(2)} \rangle|. \end{aligned} \quad (4.2.24)$$

Since every word  $u^{(1)}$  and  $u^{(2)}$  has strictly smaller length than  $u$  and  $|u^{(1)}|_\Gamma + |u^{(2)}|_\Gamma$ , we can use the recursive hypothesis and for any couple of terms  $u^{(1)}, u^{(2)}$  there exist two positive random variables  $K^{u^{(1)}}, K^{u^{(2)}}$  such that for any  $i = 1, 2$

$$|\langle \mathbf{X}_{st}, u^{(i)} \rangle| \leq K^{u^{(i)}} |t - s|^{\gamma|u^{(i)}|_\Gamma}, \quad \mathbb{E} |K^{u^{(i)}}|^{q/|u^{(i)}|_\Gamma} < +\infty. \quad (4.2.25)$$

Plugging the estimates (4.2.25) into the right hand side of (4.2.24) and using the properties of the partition (4.2.23) we obtain

$$\begin{aligned}
|F(s, t)| &\leq 2 \sum_{n \geq m+1} K_n + K^{u^{(1)}} K^{u^{(2)}} \sum_{i=0}^{M-1} |\tau_{i+1} - \tau_i|^{\gamma|u^{(2)}|_\Gamma} |\tau_i - s|^{\gamma|u^{(1)}|_\Gamma} \\
&\leq 2 \sum_{n \geq m+1} K_n + \left( 2 \sum_{n \geq m+1} |D_n|^{\gamma|u^{(1)}|_\Gamma} \right) \left( 2 \sum_{n \geq m+1} |D_n|^{\gamma|u^{(2)}|_\Gamma} \right) K^{u^{(1)}} K^{u^{(2)}} \\
&\leq 2 \sum_{n \geq m+1} K_n + C' |D_{m+1}|^{\gamma|u|_\Gamma} K^{u^{(1)}} K^{u^{(2)}},
\end{aligned}$$

for some deterministic constant  $C' > 0$ . Therefore we obtain

$$\frac{|F(s, t)|}{|t - s|^{\gamma|u|_\Gamma}} \leq 2 \sum_{n \geq m+1} \frac{K_n}{|D_{m+1}|^{\gamma|u|_\Gamma}} + C K^{u^{(1)}} K^{u^{(2)}} \leq K + C' K^{u^{(1)}} K^{u^{(2)}},$$

where we consider the random variable  $K := 2 \sum_{n \geq 0} K_n / |D_{m+1}|^{\gamma|u|_\Gamma}$ . The bound (4.2.22) implies immediately

$$\mathbb{E}|K|^{q/|u|_\Gamma} \leq 2^{q/|u|_\Gamma} \sum_{n \geq 0} \frac{\mathbb{E}|K_n|^{q/|u|_\Gamma}}{|D_n|^{\gamma q}} \leq C 2^{q/|u|_\Gamma} \sum_{n \geq 0} |D_n|^{q(\frac{1}{N} - \gamma) - 1} < +\infty$$

because by hypothesis  $q(\frac{1}{N} - \gamma) < q(\frac{1}{N} - \frac{1}{N+1}) > 1$ . Moreover the product  $K^{u^{(1)}} K^{u^{(2)}}$  belongs to  $L^{q/|u|_\Gamma}(\mathbb{P})$  as a trivial consequence of (4.2.25) and the Hölder inequality. Thus the process  $F(s, t)$  satisfies (4.2.21) and up to modifications can be extended to a Hölder continuous process with the right exponent for any  $u \in L^\Gamma(A)$  of length  $n + 1$ . Writing any  $w \in W^\Gamma(A)$  such that  $|w| = n + 1$  as the sum (4.2.15) where  $|l_i| \leq n + 1$  for any  $i$  we have automatically (4.2.21) and the continuous trajectories because  $\sum_{k=1}^M j_k |l_k|_\Gamma = |w|_\Gamma$ , thereby obtaining the result.  $\square$

### 4.3 Stochastic heat equation

We apply the results in the previous section to  $X$  the time evolution of the stochastic heat equation (4.1.1), which is given explicitly by

$$X_t = \int_0^t \int_{\mathbb{R}} G(t - s, x - y) W(ds, dy),$$

for some fixed  $x \in \mathbb{R}$ . We recall from the general definition of the Wiener integration [NZ98] that for every  $f \in L^2([0, +\infty) \times \mathbb{R})$  the random variable

$$I(f) = \int_0^{+\infty} \int_{\mathbb{R}} f(s, y) W(ds, dy)$$

is a centred Gaussian variable such that for any  $f, g \in L^2([0, +\infty) \times \mathbb{R})$  we have the identity

$$\mathbb{E}[I(f)I(g)] = \int_0^{+\infty} \int_{\mathbb{R}} f(s, y)g(s, y) dsdy, \quad (4.3.1)$$

Let us resume the main properties of the process  $X$ .

**Proposition 4.3.1.** *The process  $X$  is a centred Gaussian process whose covariance function is explicitly given by (4.1.3). Moreover there exists a continuous modification of  $X$  belonging to  $\mathcal{C}^\gamma(\mathbb{R})$  for some fixed  $\gamma \in (1/5, 1/4)$  such that for any  $p \geq 1$*

$$\mathbb{E} \left| \sup_{s \neq t} \frac{|X_t - X_s|}{|t - s|^\gamma} \right|^p < \infty \quad (4.3.2)$$

*Proof.* This result is a classical consequence of the properties associated to the Wiener integration (see [Swa07]). We recall it here for sake of completeness. By applying the property (4.3.1) and the semigroup property of the heat kernel  $G$  we obtain

$$\begin{aligned} \mathbb{E}[X_t X_s] &= \mathbb{E}[v(t, x)v(s, x)] = \int_0^t \int_{\mathbb{R}} G(t-r, x-y)G(s-r, x-y)dydr \\ &= \int_0^t G(t+s-2r, 0)dr = \frac{1}{\sqrt{2\pi}}(\sqrt{t+s} - \sqrt{|t-s|}), \end{aligned}$$

thereby obtaining the identity (4.1.3). The explicit form of the covariance implies

$$\mathbb{E}|X_t - X_s|^2 = \frac{1}{\sqrt{\pi}}(\sqrt{t} + \sqrt{s} - \sqrt{2(t+s)} + \sqrt{2|t-s|}),$$

which implies immediately for any  $0 \leq s < t$

$$\begin{aligned} \mathbb{E}|X_t - X_s|^2 - \frac{2}{\sqrt{\pi}}\sqrt{|t-s|} &= \frac{1}{\sqrt{\pi}} \frac{(\sqrt{t} - \sqrt{s})^2}{(\sqrt{t} + \sqrt{s} + \sqrt{2(t+s)})} \leq \frac{(\sqrt{t} - \sqrt{s})^2}{\sqrt{\pi}(1 + \sqrt{2})\sqrt{t}} \\ &= \frac{|t-s|^2}{\sqrt{\pi}(1 + \sqrt{2})\sqrt{t}(\sqrt{t} + \sqrt{s})^2} \leq \frac{|t-s|^2}{\sqrt{\pi}(1 + \sqrt{2})t^{3/2}}. \end{aligned} \quad (4.3.3)$$

Hence by symmetry we deduce that there exists a constant  $C > 0$  such that

$$\mathbb{E}|X_t - X_s|^2 \leq C\sqrt{|t-s|}.$$

Combining this inequality with the Gaussian nature of  $X$  allows to deduce for every integer  $p \geq 2$  the existence of a constant  $C_p > 0$  such that

$$\mathbb{E}|X_t - X_s|^p \leq C_p |t-s|^{p/4} \quad (4.3.4)$$

Choosing  $p$  sufficiently large we can apply the classical tool of Kolmogorov's continuity criterion (see [RY04, Thm 2.1]) to obtain that there exists a continuous modification of  $X$  belonging to the space of  $\gamma$  Hölder functions for  $\gamma \in (1/5, 1/4)$  satisfying (4.3.2).  $\square$

From the Proposition 4.3.1, we can apply immediately the identities (4.2.4), (4.2.5) to the a.s. realisations of  $X$ , obtaining the following rough change of variable formulae

**Proposition 4.3.2.** *For any  $\varphi \in C_b^5(\mathbb{R}, \mathbb{R})$  we have the a.s. identity*

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \Phi'(X_r) d\mathbf{X}_r^1. \quad (4.3.5)$$

Moreover the rough integral  $\int_s^t \Phi'(Y_r) d\mathbf{X}_r^1$  is defined explicitly as the limit

$$\int_s^t \Phi'(X_r) d\mathbf{X}_r^1 = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \sum_{l=1}^4 \varphi^{(l)}(X_u) \frac{(X_v - X_u)^l}{l!}, \quad (4.3.6)$$

where  $\mathcal{P}_n$  is an arbitrary sequence of partition of  $[s, t]$  such that its mesh-size  $|\mathcal{P}_n| \rightarrow 0$ .

**Proposition 4.3.3.** *For any  $\varphi \in C_b^5(\mathbb{R}, \mathbb{R})$  and any random  $\mathbb{X} \in \hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_4)$  over the path  $X$  one has the a.s. identity*

$$\varphi(X_t) = \varphi(X_s) + \sum_{n=1}^4 \frac{1}{n!} \int_s^t \Phi^{(n)}(X_r) d\mathbb{X}_r^n. \quad (4.3.7)$$

We will discuss in the following some possible choices of random one dimensional quasi-geometric rough paths  $\mathbb{X}$  in (4.3.7) and some probabilistic properties.

### 4.3.1 Burdzy-Swanson formula

The problem of showing a change of variable formula on the same process  $X$  for functions  $g \in C^m(\mathbb{R}, \mathbb{R})$  was already studied in the paper [BS10] by means of some different probabilistic tools. We recall here the main formula obtained in this work. For its proof we refer to [BS10, Thm 6.2].

**Theorem 4.3.4.** *Let  $g \in C^9(\mathbb{R}, \mathbb{R})$ . We consider the sequence of processes given by the midpoint-Riemann approximations*

$$I_t^n(g') := \sum_{j=1}^{\lfloor nt/2 \rfloor} g'(X_{t_{2j-1}})(X_{t_{2j}} - X_{t_{2j-2}}), \quad t_j = \frac{j}{n}, \quad t \in [0, T]$$

and for some explicit constant  $\kappa > 0$  the sequence of alternating squares

$$B_t^n := \kappa^{-1} \sum_{j=1}^{2\lfloor nt/2 \rfloor} (-1)^j (X_{t_j} - X_{t_{j-1}})^2, \quad t_j = \frac{j}{n}, \quad t \in [0, T].$$

Then the triple  $(X, B^n, I^n(g'))$  converges in law with respect to the Skorohod topology on  $D([0, T], \mathbb{R}^3)$ , the space of càdlàg functions, to the triple  $(X, B, I)$  where  $B$  is a Brownian motion independent of  $X$  and writing the process  $I$  as  $\int_0^t g'(X_s) d^M X_s$  we have the identity in law between processes

$$g(X_t) - g(X_0) - \frac{\kappa}{2} \int_0^t g''(X_s) dB_s \stackrel{d}{=} \int_0^t g'(X_s) d^M X_s, \quad (4.3.8)$$

where the integral on the left hand side (4.3.8) is an Itô integral.

*Remark 4.3.5.* The original formulation of (4.3.8) is given on a specific class of regular functions  $h: \mathbb{R} \times [0, +\infty) \rightarrow \mathbb{R}$  and it is formulated on the process  $h(X_t, t)$ . Then we have the presence also of the additional term  $\int_0^t \partial_t h(X_s, s) ds$ . The choice of considering this simpler formulation is made because we can easily obtain this correction using the general theory developed in the previous chapter.

In what follows we will show how this theorems can be included in the context of the path-wise change of variable formulae given in the introduction. we remark immediately that the formula (4.3.8) is weaker than the path-wise identities (4.3.7). However we can replace the left hand side of (4.3.8) with a rough integral with respect to a specific quasi-geometric rough path  $\mathfrak{X} \in \hat{\mathcal{R}}_g^\Gamma(\mathbb{A}_4)$  over  $X$ . In order to build  $\mathfrak{X}$  we will apply the Proposition 4.2.8 and the Proposition 4.2.9 to show the existence of geometric rough path  $\mathbf{X} \in \mathcal{R}_g^\Gamma(\{1, 2\})$  over  $X$ .

**Proposition 4.3.6.** *Let  $B$  be a Brownian motion independent of  $X$  and  $\kappa > 0$  the constant given in Theorem 4.3.4. Then there exists a unique  $\mathbb{X} \in \mathcal{R}_g^\Gamma(\{1, 2\})$  such that*

$$\begin{aligned} \langle \mathbb{X}_{st}, 1 \rangle &:= X_t - X_s, & \langle \mathbb{X}_{st}, 2 \rangle &:= \kappa(B_t - B_s), \\ \langle \mathbb{X}_{st}, 12 \rangle &:= \kappa \int_s^t (X_r - X_s) dB_r, & \langle \mathbb{X}_{st}, 112 \rangle &:= \kappa \int_s^t \frac{(X_r - X_s)^2}{2} dB_r. \end{aligned} \quad (4.3.9)$$

Considering  $\mathbb{X}^0 \in \mathcal{R}_g^\Gamma(\mathbb{A}_4)$ , the zero extension of  $\mathbb{X}$ , we call the quasi-geometric rough path  $\mathfrak{X} = \log^* \mathbb{X}^0$  the Burdzy-Swanson extension of  $X$ .

*Remark 4.3.7.* We observe that the conditions (4.3.9) imposed to define  $\mathbb{X}$  are chosen to define  $\mathbb{X}$  on  $L^\Gamma(\{1, 2\})$ , the set of Lyndon words defined in (4.2.12). Indeed we recall that for any alphabet  $A = \{a, b\} \subset \mathbb{A}_N$  where  $a \prec b$  the set  $L^4(A)$  containing the Lyndon words whose length is smaller than 4 is explicitly given by

$$L^4(A) = \{a, b, ab, aab, abb, aaab, aabb, abbb\}. \quad (4.3.10)$$

Then by choosing  $\gamma \in (1/5, 1/4)$  we deduce that

$$L^\Gamma(\{1, 2\}) = \{1, 2, 12, 112\}.$$

Thus the hypothesis (4.3.9) identifies a random process  $\tilde{\mathbb{X}}: [0, T]^2 \rightarrow T^\Gamma(\{\{1, 2\}\})$  which satisfies the property (4.2.16) by means of Proposition (4.2.13).

*Proof.* Considering the random process  $\tilde{\mathbb{X}}: [0, T]^2 \rightarrow T^\Gamma(\{\{1, 2\}\})$  defined from the conditions (4.3.9), we will prove the existence and the uniqueness of  $\mathbb{X}$  by showing that there exists a modification of  $\tilde{\mathbb{X}}$  satisfying a.s. the property (4.2.17) and the bound (4.2.21). Starting with the Chen's property we see that for the values  $\langle \tilde{\mathbb{X}}_{st}, 2 \rangle$  and  $\langle \tilde{\mathbb{X}}_{st}, 1 \rangle$  the identity is trivial and we only need to verify for any  $s, u, t \in [0, T]^3$  the identities

$$\begin{aligned} \langle \tilde{\mathbb{X}}_{st}, 12 \rangle &= \langle \tilde{\mathbb{X}}_{su}, 12 \rangle + \langle \tilde{\mathbb{X}}_{ut}, 12 \rangle + \langle \tilde{\mathbb{X}}_{su}, 1 \rangle \langle \tilde{\mathbb{X}}_{ut}, 2 \rangle, \\ \langle \tilde{\mathbb{X}}_{st}, 112 \rangle &= \langle \tilde{\mathbb{X}}_{su}, 112 \rangle + \langle \tilde{\mathbb{X}}_{ut}, 112 \rangle + \langle \tilde{\mathbb{X}}_{ut}, 11 \rangle \langle \tilde{\mathbb{X}}_{ut}, 2 \rangle + \langle \tilde{\mathbb{X}}_{ut}, 1 \rangle \langle \tilde{\mathbb{X}}_{ut}, 12 \rangle, \end{aligned} \quad (4.3.11)$$



which hold trivially because of the classical properties of Itô integration and the identity  $\langle \tilde{\mathfrak{X}}_{st}, 11 \rangle = 1/2 \langle \tilde{\mathfrak{X}}_{st}, 1 \rangle^2$ . Passing to the bound (4.2.21), Recalling the properties of  $X$  and the Brownian motion  $B$ , for any  $q > 4$  there exists a constant  $C_q \geq 0$  depending on  $q$  such that

$$\mathbb{E}|\langle \tilde{\mathfrak{X}}_{st}, 1 \rangle|^q \leq C_p |t - s|^{q/4}, \quad \mathbb{E}|\langle \tilde{\mathfrak{X}}_{st}, 2 \rangle|^q \leq C_p |t - s|^{q/2}.$$

Moreover we can apply the Burkholder-Davis-Gundy inequality ([RY04, Thm IV.4.1]) and the previous estimates to obtain that for any  $q > 4$  there exist  $C'_q \geq 0$  and  $D \geq 0$  such that

$$\begin{aligned} \mathbb{E}|\langle \tilde{\mathfrak{X}}_{st}, 12 \rangle|^q &\leq C'_q \left( \int_s^t \mathbb{E}|X_r - X_s|^2 dr \right)^{q/2} \leq C'_q C_q \left( \int_s^t \sqrt{|r - s|} dr \right)^{q/2} \\ &\leq D |t - s|^{(q^3)/4}, \\ \mathbb{E}|\langle \tilde{\mathfrak{X}}_{st}, 112 \rangle|^q &\leq C'_q \left( \int_s^t \mathbb{E}|X_r - X_s|^4 dr \right)^{q/2} \leq C'_q C_q \left( \int_s^t |r - s| dr \right)^{q/2} \\ &\leq D |t - s|^q. \end{aligned}$$

Therefore we can apply the Proposition 4.2.14 to conclude.  $\square$

*Remark 4.3.8.* Looking at the definition of the quasi-geometric rough path  $\mathfrak{X}$ , the first line of the conditions (4.3.9) are chosen to obtain the identity

$$\langle \mathfrak{X}_{0t}, 11 \rangle = \langle \mathfrak{X}_{0t}^0, \log(11) \rangle = \langle \mathfrak{X}_{0t}, 11 \rangle - \frac{1}{2} \langle \mathfrak{X}_{0t}, 2 \rangle = \frac{X_t^2}{2} - \frac{\kappa}{2} B_t.$$

Thus the function  $t \rightarrow \langle \mathfrak{X}_{0t}, 11 \rangle$  must coincide in law with the process  $\int_0^t X_r d^M X_r$  obtained from the Theorem 4.3.4 when  $g(x) = x^2$  (see [Swa07, Cor 4.8]). Hence the name Burdzy-Swanson extension for  $\mathfrak{X}$ .

Let us show how does the general change of variable formula (4.2.5) becomes when we consider the Burdzy-Swanson extension of  $X$ .

**Theorem 4.3.9.** *Let  $\mathfrak{X}$  be the Burdzy-Swanson extension of  $X$  and  $\psi \in C_b^3(\mathbb{R}, \mathbb{R})$ . Then we have the a.s. equality*

$$\int_s^t \Psi(X_r) d\mathfrak{X}_r^2 = \kappa \int_s^t \psi(X_r) dB_r. \quad (4.3.12)$$

Therefore for any  $\varphi \in C_b^5(\mathbb{R}, \mathbb{R})$  we have the identity

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \Phi'(X_r) d\mathfrak{X}_r^1 + \frac{\kappa}{2} \int_s^t \varphi''(X_r) dB_r. \quad (4.3.13)$$

*Proof.* By definition of the rough integral given in the equation (4.2.2) we have

$$\begin{aligned} \int_s^t \Psi(X_r) d\mathfrak{X}_r^2 = \\ \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \psi(X_u) \langle \mathfrak{X}_{uv}, 2 \rangle + \psi'(X_u) \langle \mathfrak{X}_{uv}, 12 \rangle + \psi''(X_u) \langle \mathfrak{X}_{uv}, 112 \rangle + \frac{\psi''(X_u)}{2} \langle \mathfrak{X}_{uv}, 22 \rangle, \end{aligned}$$

for any sequence of partition  $\mathcal{P}_n$  of  $[s, t]$  such that  $|\mathcal{P}_n| \rightarrow 0$ . Moreover the convergence holds a.s. uniformly on  $s < t$  such that  $s, t \in [0, T]$ . Let us rewrite the terms of the sum using the rough path  $\mathbb{X}$  defined by (4.3.9)

$$\langle \mathfrak{X}_{uv}, 12 \rangle = \langle \mathbb{X}_{uv}^0, \log(12) \rangle = \langle \mathbb{X}_{uv}^0, 12 \rangle - \frac{1}{2} \langle \mathbb{X}_{uv}^0, 3 \rangle = \langle \mathbb{X}_{uv}, 12 \rangle = \kappa \int_u^v (X_r - X_u) dB_r,$$

$$\begin{aligned} \langle \mathfrak{X}_{uv}, 112 \rangle &= \langle \mathbb{X}_{uv}^0, \log(112) \rangle = \langle \mathbb{X}_{uv}^0, 112 \rangle - \frac{1}{2} \langle \mathbb{X}_{uv}^0, 22 \rangle - \frac{1}{2} \langle \mathbb{X}_{uv}^0, 13 \rangle + \frac{1}{3} \langle \mathbb{X}_{uv}^0, 4 \rangle \\ &= \langle \mathbb{X}_{uv}, 112 \rangle - \frac{1}{2} \langle \mathbb{X}_{uv}, 22 \rangle = \kappa \int_u^v \frac{(X_r - X_u)^2}{2} dB_r - \frac{\kappa^2}{4} (B_v - B_u)^2, \end{aligned}$$

$$\frac{1}{2} \langle \mathfrak{X}_{uv}, 22 \rangle = \frac{1}{2} \langle \mathbb{X}_{uv}^0, \log(22) \rangle = \frac{1}{2} \langle \mathbb{X}_{uv}^0, 22 \rangle - \frac{1}{4} \langle \mathbb{X}_{uv}^0, 4 \rangle = \frac{1}{2} \langle \mathbb{X}_{uv}, 22 \rangle = \frac{\kappa^2}{4} (B_v - B_u)^2.$$

Thus the converging sequence defining the rough integral simplifies and it becomes

$$\int_s^t \Psi(X_r) d\mathfrak{X}_r^2 = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \psi(X_u) \langle \mathbb{X}_{uv}, 2 \rangle + \psi'(X_u) \langle \mathbb{X}_{uv}, 12 \rangle + \psi''(X_u) \langle \mathbb{X}_{uv}, 112 \rangle. \quad (4.3.14)$$

On the other hand, using a classical argument of stochastic calculus (see [RY04, Prop IV.2.13]), for any fixed sequence of partition  $\mathcal{P}_n$  of  $[s, t]$  such that  $|\mathcal{P}_n| \rightarrow 0$  one has

$$\int_s^t \psi(X_r) dB_r = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \psi(X_u) (B_v - B_u), \quad (4.3.15)$$

where the convergence is in probability uniformly on  $s, t$  as before. Passing to a subsequence we can suppose that the convergence (4.3.15) holds a.s. and using the identity  $\langle \mathfrak{X}_{uv}, 2 \rangle = \kappa(B_v - B_u)$  we subtract the sequence (4.3.15) from (4.3.14) to obtain that the sequence

$$X_n := \sum_{[u,v] \in \mathcal{P}_n} \psi'(X_u) \langle \mathbb{X}_{uv}, 12 \rangle + \psi''(X_u) \langle \mathbb{X}_{uv}, 112 \rangle$$

converges a.s. to some random variable  $\chi_{st}$  uniformly on  $s, t$ . In order to prove the thesis it is sufficient to show that this limit is zero a.s. By fixing a sequence of partitions  $\mathcal{P}_n = \{s = t_0^n < \dots < t_{p(n)+1}^n = t\}$  and introducing for  $i = 0, \dots, p(n) - 1$  the family of square integrable random variables  $M_i^n, N_i^n$

$$\begin{aligned} M_i^n &:= \int_{t_i^n}^{t_{i+1}^n} (X_r - X_{t_i}) dB_r = \int_0^T \mathbf{1}_{[t_i^n, t_{i+1}^n)}(r) (X_r - X_{t_i}) dB_r, \\ N_i^n &:= \int_{t_i^n}^{t_{i+1}^n} \frac{(X_r - X_{t_i})^2}{2} dB_r = \int_0^T \mathbf{1}_{[t_i^n, t_{i+1}^n)}(r) \frac{(X_r - X_{t_i})^2}{2} dB_r. \end{aligned}$$

we can rewrite the sequence  $X_n$  as

$$X_n = \kappa \sum_{i=0}^{p(n)-1} \psi'(X_{t_i^n}) M_i^n + \psi''(X_{t_i^n}) N_i^n.$$

We will prove the convergence of  $X_n$  to 0 by estimating the second moment of  $X_n$ . Writing  $X_n^2$  as a double sum we have

$$X_n^2 = \kappa^2 \sum_{i,j=0}^{p(n)-1} \left( \psi'(X_{t_j^n})\psi'(X_{t_i^n})M_i^n M_j^n + \psi''(X_{t_i^n})\psi''(X_{t_j^n})N_j^n N_i^n \right. \\ \left. + \psi''(X_{t_i^n})\psi'(X_{t_j^n})N_i^n M_j^n + \psi''(X_{t_j^n})\psi'(X_{t_i^n})N_j^n M_i^n \right).$$

Applying the hypothesis of boundedness of the higher derivatives of  $\psi$  there exists a constant  $C > 0$  such that

$$\mathbb{E}X_n^2 \leq C \sum_{i,j=0}^{p(n)-1} \mathbb{E}[M_i^n M_j^n] + \mathbb{E}[N_i^n N_j^n] + \mathbb{E}[N_i^n M_j^n] + \mathbb{E}[M_i^n N_j^n]. \quad (4.3.16)$$

For any couple of indexes  $i, j$ , the Itô isometry and the Gaussian nature of  $X$  provides the identities

$$\mathbb{E}[M_i^n M_j^n] = \delta_{ij} \kappa^2 \int_{t_i^n}^{t_{i+1}^n} \mathbb{E}(X_r - X_{t_i^n})^2 dr, \quad \mathbb{E}[N_i^n N_j^n] = \delta_{ij} \frac{\kappa^2}{4} \int_{t_i^n}^{t_{i+1}^n} \mathbb{E}(X_r - X_{t_i^n})^4 dr, \\ \mathbb{E}[N_i^n M_j^n] = \delta_{ij} \frac{\kappa^2}{2} \int_{t_i^n}^{t_{i+1}^n} \mathbb{E}(X_r - X_{t_i^n})^3 dr = 0.$$

Plugging them into (4.3.16) and using the general estimate (4.3.4), there exist a couple of constant  $C', C'' > 0$  such that

$$\mathbb{E}X_n^2 \leq C' \sum_{i=0}^{p(n)-1} \int_{t_i^n}^{t_{i+1}^n} (r - t_i^n)^{1/2} + (r - t_i^n) dr \leq C'' \sum_{i=0}^{p(n)-1} (t_{i+1}^n - t_i^n)^{3/2} + (t_{i+1}^n - t_i^n)^2 \\ \leq 2TC''(|\mathcal{P}_n|^{1/2} + |\mathcal{P}_n|)$$

Thereby obtaining the convergence  $X_n \rightarrow 0$  uniformly on  $s, t$  as  $|\mathcal{P}_n| \rightarrow 0$ . Using Fatou's lemma and Jensen inequality we have

$$\mathbb{E} \sup_{s,t} |\chi_{st}| \leq \lim_{n \rightarrow \infty} \mathbb{E}|X_n| \leq \lim_{n \rightarrow \infty} \mathbb{E}|X_n|^2 = 0,$$

concluding that  $\chi_{st} = 0$  a.s. The formula (4.3.13) is obtained by combining Proposition 4.2.5, the identity (4.3.12) and observing trivially that the third of fourth order rough integrals are zero.  $\square$

Combining Theorem 4.3.4 with Theorem 4.3.9 we obtain the following alternative expression of the first order rough integral with respect to  $\mathfrak{X}$ .

**Corollary 4.3.10.** *For any  $\psi \in C_b^5(\mathbb{R}, \mathbb{R})$  we have the identity in law between the processes*

$$\int_0^t \Psi(X_r) d\mathfrak{X}_r^1 \stackrel{d}{=} \int_0^t \psi(X_r) d^M X_r. \quad (4.3.17)$$

*Remark 4.3.11.* Looking at the identity (4.3.17), we stress the fact that the right hand side of the identity is a limit defined on a specific sequence partition of  $[0, t]$ , whereas the rough integral is independent of this choice. Moreover this identity suggests that the convergence of the triplet  $(X, B^n, I^n(g'))$  explained in Theorem 4.3.4 could be restated by showing the existence of a sequence of discrete rough paths converging in law to  $\mathfrak{X}$ .

### 4.3.2 Cheridito-Nualart formula

The Theorem 4.3.9 is not the only possible applications of the general formula (4.3.7). Using the same strategy of the construction of the Burdzy-Swanson rough path, we can deduce a different quasi geometric rough path above  $X$  with a non trivial quadratic variation. In order to define it, instead of starting from a Brownian motion  $B$  independent of  $X$  we consider the deterministic function of the variance of  $X$ ,  $\sigma^2: [0, T] \rightarrow \mathbb{R}$

$$\sigma_r^2 := \mathbb{E}(X_r^2) = \frac{1}{\sqrt{\pi}} r^{1/2}$$

The function  $\sigma^2$  is clearly an absolutely continuous function  $1/2$  Hölder continuous. Thus we can integrate the continuous trajectories of  $X$  with respect to  $\sigma^2$ . Replacing the stochastic integrals in 4.3.9 with these deterministic integrals, we have an alternative construction.

**Proposition 4.3.12.** *There exists a unique  $\mathbf{X} \in \mathcal{R}_g^\Gamma(\{1, 2\})$  such that*

$$\begin{aligned} \langle \mathbf{X}_{st}, 1 \rangle &:= X_t - X_s, & \langle \mathbf{X}_{st}, 2 \rangle &:= \sigma_t^2 - \sigma_s^2, \\ \langle \mathbf{X}_{st}, 12 \rangle &:= \int_s^t (X_r - X_s) d\sigma_r^2, & \langle \mathbf{X}_{st}, 112 \rangle &:= \int_s^t \frac{(X_r - X_s)^2}{2} d\sigma_r^2. \end{aligned} \quad (4.3.18)$$

Considering  $\mathbf{X}^0$ , the zero extension of  $\mathbf{X}$ , we call the quasi geometric rough path  $\hat{\mathfrak{X}} = \log^* \mathbf{X}^0$  the Cheridito-Nualart extension of  $X$ .

*Proof.* As explained before we check that the random process  $\mathbf{Z}: [0, T]^2 \rightarrow T^\Gamma(\{1, 2\})$  defined from the conditions (4.3.18), satisfies the properties (4.2.17) and the bounds (4.2.21). Starting with the Chen's property we only need to verify for any  $s, u, t \in [0, T]^3$  the same identities of the equation (4.3.11) with  $\mathbf{Z}$  instead of  $\tilde{\mathfrak{X}}$ . However the same identities hold trivially because of the standard properties of Stieltjes integral and the identity  $\langle \mathbf{Z}_{st}, 11 \rangle = 1/2 \langle \mathbf{Z}_{st}, 1 \rangle^2$ . Then we conclude by showing the property (4.2.21) for  $\mathbf{Z}$ . Using the Jensen inequality for the Stieltjes integrals and the Hölder continuity of  $\sigma^2$  there exists a constant  $C > 0$  such that for any  $q > 4$  we have the a.s. estimates

$$\begin{aligned} |\langle \mathbf{Z}_{st}, 12 \rangle|^q &\leq |\sigma_t^2 - \sigma_s^2|^{q-1} \int_s^t |X_r - X_s|^q d\sigma_r^2 \leq C^{p-1} |t - s|^{(q-1)/2} \int_s^t |X_r - X_s|^q d\sigma_r^2, \\ |\langle \mathbf{Z}_{st}, 112 \rangle|^q &\leq |\sigma_t^2 - \sigma_s^2|^{q-1} \int_s^t |X_r - X_s|^{2q} d\sigma_r^2 \leq C^{p-1} |t - s|^{(q-1)/2} \int_s^t |X_r - X_s|^{2q} d\sigma_r^2. \end{aligned}$$

Combining them with the bound (4.3.4), for any  $q > 4$  there exists a positive constant  $D_q$  such that

$$\begin{aligned}\mathbb{E}|\langle \mathbf{Z}_{st}, 12 \rangle|^q &\leq D_q |t-s|^{(q-1)/2} \int_s^t |r-s|^{q/4} d\sigma_r^2 \leq CD_q |t-s|^{(q^3)/4}, \\ \mathbb{E}|\langle \mathbf{Z}_{st}, 112 \rangle|^q &\leq D_q |t-s|^{(q-1)/2} \int_s^t |r-s|^{q/2} d\sigma_r^2 \leq CD_q |t-s|^q.\end{aligned}$$

Therefore we can apply again the Proposition 4.2.14 and we conclude.  $\square$

Let us show how does the general change of variable formula (4.3.7) modifies in case of the variance extension. In order to prove this formula we need to recall a specific property of the process  $X$  involving the *rectangular increments* and a sharp estimate on some deterministic integrals

**Lemma 4.3.13.** *For all  $0 \leq v < u \leq s < t$  one has*

$$|\mathbb{E}[(X_u - X_v)(X_t - X_s)]| \leq \sqrt{\frac{2}{\pi}} \frac{|t-s||u-v|}{|s-v|\sqrt{t-u}} \quad (4.3.19)$$

*Proof.* This result was already obtained in [Swa07, Lem. 2.2]. We repeat here the proof for sake of completeness. Using the explicit form of the covariance given in (4.1.3) for any  $r > u$  one has the identities

$$\begin{aligned}g(r) &:= \mathbb{E}[(X_u - X_v)X_r] = \frac{1}{\sqrt{2\pi}}(\sqrt{r+u} - \sqrt{r-u} + \sqrt{r+v} - \sqrt{r-v}), \\ g'(r) &= \frac{1}{2\sqrt{2\pi}} \left( \frac{\sqrt{r+v} - \sqrt{r+u}}{\sqrt{(r+u)(r+v)}} - \frac{\sqrt{r-v} - \sqrt{r-u}}{\sqrt{(r-u)(r-v)}} \right).\end{aligned}$$

Using the trivial estimate

$$\left| \frac{\sqrt{r \pm v} - \sqrt{r \pm u}}{\sqrt{(r \pm u)(r \pm v)}} \right| \leq \frac{|u-v|}{\sqrt{r-u}|r-v|},$$

we obtain immediately

$$\begin{aligned}|g(t) - g(s)| &\leq \int_s^t |g'(r)| dr \leq \frac{1}{\sqrt{2\pi}} \int_s^t \frac{|u-v|}{\sqrt{r-u}|r-v|} dr \leq \frac{1}{\sqrt{2\pi}} \frac{|u-v|}{|s-v|} \int_s^t \frac{1}{\sqrt{r-u}} dr \\ &= \sqrt{\frac{2}{\pi}} \frac{|u-v|}{|s-v|} (\sqrt{t-u} - \sqrt{s-u}) \leq \sqrt{\frac{2}{\pi}} \frac{|t-s||u-v|}{|s-v|\sqrt{t-u}}.\end{aligned}$$

Thereby obtaining the thesis.  $\square$

**Lemma 4.3.14.** *There exists a constant  $D > 0$  such that for any  $\varepsilon > 0$  there exists  $\delta > 0$  such that for all  $|t-s| \leq \delta$  one has*

$$\int_s^t \mathbb{E}[(X_r - X_s)^2] d\sigma_r^2 \leq D|t-s|\varepsilon. \quad (4.3.20)$$

*Proof.* We suppose  $s < t$ . Applying the bound (4.3.4) and the definition of  $\sigma^2$  we obtain easily that there exists a constant  $C > 0$  such that

$$\begin{aligned} \int_s^t \mathbb{E}[(X_r - X_s)^2] d\sigma_r^2 &\leq C \int_s^t \frac{\sqrt{r-s}}{\sqrt{r}} dr = C[\sqrt{r-s}\sqrt{r} - s \log(\sqrt{r-s} + \sqrt{r})]_{r=s}^{r=t} \\ &= C \left( \sqrt{t-s}\sqrt{t} - s \ln \left( \frac{\sqrt{t-s} + \sqrt{t}}{\sqrt{s}} \right) \right) \\ &= C \left( \sqrt{t-s}\sqrt{t} - s \ln \left( 1 + \frac{\sqrt{t}-\sqrt{s}}{\sqrt{s}} \right) - s \ln \left( 1 + \frac{\sqrt{t-s}}{\sqrt{t}} \right) \right). \end{aligned}$$

Using the classical estimate  $x - \ln(1+x) \leq x^2/2$  holding for any  $x \in (-1, 1)$ , by sending  $s \rightarrow t$  we obtain

$$\begin{aligned} \int_s^t \mathbb{E}[(X_r - X_s)^2] d\sigma_r^2 &\leq \\ &C \left( \sqrt{t-s}\sqrt{t} - \sqrt{s}(\sqrt{t}-\sqrt{s}) + \frac{1}{2}(\sqrt{t}-\sqrt{s})^2 - s \frac{\sqrt{t-s}}{\sqrt{t}} + \frac{s}{2t}(t-s) \right) \\ &= C(t-s) \left( \frac{\sqrt{t-s}}{\sqrt{t}} + \frac{s}{2t} + \frac{1}{2} \frac{(\sqrt{t}-\sqrt{s})^2}{t-s} - \sqrt{s} \frac{(\sqrt{t}-\sqrt{s})}{(t-s)} \right) \\ &= C(t-s) \left( \frac{\sqrt{t-s}}{\sqrt{t}} + \frac{1}{2} \frac{(\sqrt{t}-\sqrt{s})}{\sqrt{t}+\sqrt{s}} + \frac{s}{2t} - \sqrt{s} \frac{1}{\sqrt{t}+\sqrt{s}} \right). \end{aligned}$$

Since the function between round brackets converges to 0 uniformly on  $s \rightarrow t$ , we obtain the thesis.  $\square$

**Theorem 4.3.15.** *Let  $\hat{\mathfrak{X}}$  be the Cheridito-Nualart extension of  $X$  and  $\psi \in C_b^3(\mathbb{R}, \mathbb{R})$ . Then we have the a.s. equality*

$$\int_s^t \Psi(X_r) d\hat{\mathfrak{X}}_r^2 = \int_s^t \psi(X_r) d\sigma_r^2, \quad (4.3.21)$$

Therefore for any  $\varphi \in C_b^5(\mathbb{R}, \mathbb{R})$  we have the identity

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \Phi'(X_r) d\hat{\mathfrak{X}}_r^1 + \frac{1}{2} \int_s^t \varphi''(X_r) d\sigma_r^2. \quad (4.3.22)$$

*Proof.* Since we want to calculate the same rough integral of the Theorem 4.3.9 and the rough path  $\mathbf{X} \in \mathcal{R}_g^\Gamma(\{1, 2\})$  defined in 4.3.12 has the same form of  $\mathbb{X} \in \mathcal{R}_g^\Gamma(\{1, 2\})$  given in Proposition 4.3.6, we can repeat the same algebraic calculations in the proof of Theorem 4.3.9 and we obtain immediately the identity

$$\int_s^t \Psi(X_r) d\hat{\mathfrak{X}}_r^2 = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \psi(X_u) \langle \mathbf{X}_{uv}, 2 \rangle + \psi'(X_u) \langle \mathbf{X}_{uv}, 12 \rangle + \psi''(X_u) \langle \mathbf{X}_{uv}, 112 \rangle,$$

for any sequence of partition  $\mathcal{P}_n$  of  $[s, t]$  such that  $|\mathcal{P}_n| \rightarrow 0$ . Moreover the convergence holds a.s. uniformly on  $s < t$  such that  $s, t \in [0, T]$  (see the equation(4.3.14)). On the

other hand by definition of Stieltjes integral, for any fixed sequence of partition  $\mathcal{P}_n$  of  $[s, t]$  such that  $|\mathcal{P}_n| \rightarrow 0$  one has

$$\int_s^t \psi(X_r) d\sigma_r^2 = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \psi(X_u) (\sigma_v^2 - \sigma_u^2),$$

where the convergence is a.s. uniformly on  $s, t$  as before. Using the identity  $\langle \hat{\mathbf{X}}_{uv}, 2 \rangle = \sigma_v^2 - \sigma_u^2$  we can subtract again both identities as in the proof of Theorem 4.3.9 to obtain that the random variable

$$\hat{X}_n := \sum_{[u,v] \in \mathcal{P}_n} \psi'(X_u) \langle \mathbf{X}_{uv}, 12 \rangle + \psi''(X_u) \langle \mathbf{X}_{uv}, 112 \rangle$$

converges a.s. to some random variable  $\hat{\chi}_{st}$  uniformly on  $s, t$ . We will show that this limit is zero a.s. Choosing a sequence of partitions  $\mathcal{P}_n = \{s = t_0 < \dots < t_{p(n)+1}^n = t\}$  and introducing for  $i = 0, \dots, p(n) - 1$  the family of random variables  $\hat{M}_i^n, \hat{N}_i^n$

$$\hat{M}_i^n := \int_{t_i^n}^{t_{i+1}^n} (X_r - X_{t_i^n}) d\sigma_r^2, \quad \hat{N}_i^n := \int_{t_i^n}^{t_{i+1}^n} \frac{(X_r - X_{t_i^n})^2}{2} d\sigma_r^2,$$

we write the sequence  $\hat{X}_n$  as

$$\hat{X}_n = \kappa \sum_{i=0}^{p(n)-1} \psi'(X_{t_i^n}) \hat{M}_i^n + \psi''(X_{t_i^n}) \hat{N}_i^n. \quad (4.3.23)$$

Let us prove the convergence of  $\hat{X}_n$  to 0 by estimating its second moment. Writing again  $X_n^2$  as a double sum and applying the hypothesis of boundedness of the higher derivatives of  $\psi$  there exists a constant  $C > 0$  such that

$$\mathbb{E} \hat{X}_n^2 \leq C \sum_{i,j=0}^{p(n)-1} \mathbb{E}[\hat{M}_i^n \hat{M}_j^n] + \mathbb{E}[\hat{N}_i^n \hat{N}_j^n] + \mathbb{E}[\hat{M}_i^n \hat{N}_j^n] + \mathbb{E}[\hat{N}_i^n \hat{M}_j^n].$$

The Gaussian nature of the process  $X$  implies that the terms  $\mathbb{E}[\hat{M}_i^n \hat{N}_j^n]$  and  $\mathbb{E}[\hat{N}_j^n \hat{M}_i^n]$  are zero because they involve a product of an odd number of Gaussian random variables. Thus by means of the Fubini Theorem we need only to estimate for any  $i, j$  the quantities

$$\begin{aligned} \mathbb{E}[\hat{M}_i^n \hat{M}_j^n] &= \int_{t_i^n}^{t_{i+1}^n} \int_{t_j^n}^{t_{j+1}^n} \mathbb{E}[(X_r - X_{t_i^n})(X_s - X_{t_j^n})] d\sigma_s^2 d\sigma_r^2, \\ \mathbb{E}[\hat{N}_i^n \hat{N}_j^n] &= \int_{t_i^n}^{t_{i+1}^n} \int_{t_j^n}^{t_{j+1}^n} \mathbb{E}[(X_r - X_{t_i^n})^2 (X_s - X_{t_j^n})^2] d\sigma_s^2 d\sigma_r^2. \end{aligned} \quad (4.3.24)$$

In case  $i = j$  we can easily apply again the Jensen inequality and the Hölder property of the function  $\sigma^2$  to obtain that there exists a constant  $C > 0$  such that

$$(\hat{M}_i^n)^2 \leq |\sigma_{t_{i+1}^n}^2 - \sigma_{t_i^n}^2| \int_{t_i^n}^{t_{i+1}^n} |X_r - X_{t_i^n}|^2 d\sigma_r^2 \leq C |t_{i+1}^n - t_i^n|^{1/2} \int_{t_i^n}^{t_{i+1}^n} |X_r - X_{t_i^n}|^2 d\sigma_r^2,$$

$$(\hat{N}_i^n)^2 \leq |\sigma_{t_{i+1}^n}^2 - \sigma_{t_i^n}^2| \int_{t_i^n}^{t_{i+1}^n} |X_r - X_{t_i^n}|^4 d\sigma_r^2 \leq C |t_{i+1}^n - t_i^n|^{1/2} \int_{t_i^n}^{t_{i+1}^n} |X_r - X_{t_i^n}|^4 d\sigma_r^2.$$

By taking the expectation, using the rectangular bound (4.3.4) and applying the Hölder property of the function  $\sigma^2$  we obtain that there exist a couple of constants  $C', D > 0$  such that

$$\begin{aligned} \mathbb{E}[(\hat{M}_i^n)^2] &\leq C' |t_{i+1}^n - t_i^n|^{1/2} \int_{t_i^n}^{t_{i+1}^n} |r - t_i^n|^{1/2} d\sigma_r^2 \leq D |t_{i+1}^n - t_i^n|^{3/2}, \\ \mathbb{E}[(\hat{N}_i^n)^2] &\leq C' |t_{i+1}^n - t_i^n|^{1/2} \int_{t_i^n}^{t_{i+1}^n} |r - t_i^n| d\sigma_r^2 \leq D |t_{i+1}^n - t_i^n|^2. \end{aligned} \quad (4.3.25)$$

We pass to estimate the terms in (4.3.24) for two indices  $i < j$ . Using the estimates on the rectangular increments given in (4.3.19) and the Hölder property of  $\sigma^2$  there exists a constant  $C'' > 0$  such that for any  $j > i + 1$  one has

$$\begin{aligned} |\mathbb{E}[\hat{M}_i^n \hat{M}_j^n]| &\leq \int_{t_i^n}^{t_{i+1}^n} \int_{t_j^n}^{t_{j+1}^n} |\mathbb{E}[(X_r - X_{t_i^n})(X_s - X_{t_j^n})]| d\sigma_s d\sigma_r \\ &\leq C'' \int_{t_i^n}^{t_{i+1}^n} \int_{t_j^n}^{t_{j+1}^n} \frac{|r - t_i^n| |s - t_j^n|}{|t_j^n - t_i^n| \sqrt{r - s}} d\sigma_s d\sigma_r \leq C'' C \frac{|t_{i+1}^n - t_i^n|^{3/2} |t_{j+1}^n - t_j^n|^{3/2}}{|t_j^n - t_{i+1}^n|^{3/2}}. \end{aligned} \quad (4.3.26)$$

Using the same rectangular estimate we can bound also  $\mathbb{E}[\hat{M}_i^n \hat{M}_{i+1}^n]$  in the same way as the first line of (4.3.24). Passing to the terms  $\mathbb{E}[\hat{N}_i^n \hat{N}_j^n]$ , we can apply Wick's formula on the product of four Gaussian variables and the estimate (4.3.19), obtaining for any  $j > i + 1$

$$\begin{aligned} \mathbb{E}[\hat{N}_i^n \hat{N}_j^n] &= \mathbb{E}[\hat{N}_i^n] \mathbb{E}[\hat{N}_j^n] + 2 \int_{t_i^n}^{t_{i+1}^n} \int_{t_j^n}^{t_{j+1}^n} (\mathbb{E}[(X_r - X_{t_i^n})(X_s - X_{t_j^n})])^2 d\sigma_s^2 d\sigma_r^2 \\ &\leq \mathbb{E}[\hat{N}_i^n] \mathbb{E}[\hat{N}_j^n] + 2C'' \int_{t_i^n}^{t_{i+1}^n} \int_{t_j^n}^{t_{j+1}^n} \frac{|r - t_i^n|^2 |s - t_j^n|^2}{|t_j^n - t_i^n|^2 |r - s|} d\sigma_s^2 d\sigma_r^2 \\ &\leq \mathbb{E}[\hat{N}_i^n] \mathbb{E}[\hat{N}_j^n] + 2C'' C \frac{|t_{i+1}^n - t_i^n|^{5/2} |t_{j+1}^n - t_j^n|^{5/2}}{|t_j^n - t_{i+1}^n|^3} \\ &\leq \mathbb{E}[\hat{N}_i^n] \mathbb{E}[\hat{N}_j^n] + 2C'' C \frac{|t_{i+1}^n - t_i^n|^{11/8} |t_{j+1}^n - t_j^n|^{11/8}}{|t_j^n - t_{i+1}^n|^{3/4}}. \end{aligned} \quad (4.3.27)$$

By means of the same estimate we can bound also  $\mathbb{E}[\hat{N}_i^n \hat{N}_{i+1}^n]$  with a term satisfying the second line of (4.3.24) plus  $\mathbb{E}[\hat{N}_i^n] \mathbb{E}[\hat{N}_{i+1}^n]$ . Let us prove the convergence  $\hat{X}_n \rightarrow 0$  uniformly on  $s, t$  as  $|\mathcal{P}_n| \rightarrow 0$  thanks to the estimates (4.3.24), (4.3.26) and (4.3.27). Choosing an arbitrary  $\varepsilon > 0$ , we apply the Lemma 4.3.14 and we choose  $\delta(\varepsilon) > 0$  such that the property (4.3.20) holds. By fixing  $N$  such  $|\mathcal{P}_n| \leq \delta(\varepsilon)$  for every  $n \geq N$  we



have

$$\begin{aligned}
\mathbb{E}\hat{X}_n^2 &\leq \sum_{i=0}^{p(n)-1} \mathbb{E}[(\hat{M}_i^n)^2 + (\hat{N}_i^n)^2] + 2 \sum_{i<j}^{p(n)-1} \mathbb{E}[\hat{M}_i^n \hat{M}_j^n] + \mathbb{E}[\hat{N}_i^n \hat{N}_j^n] \\
&\leq 2D \sum_{i=0}^{p(n)-1} |t_{i+1}^n - t_i^n|^2 + |t_{i+1}^n - t_i^n|^{3/2} + C''C2 \sum_{i+1<j}^{p(n)-1} \frac{|t_{i+1}^n - t_i^n|^{3/2} |t_{j+1}^n - t_j^n|^{3/2}}{|t_j^n - t_{i+1}^n|^{3/2}} \\
&\quad + 2C''C \sum_{i+1<j}^{p(n)-1} \frac{|t_{i+1}^n - t_i^n|^{5/2} |t_{j+1}^n - t_j^n|^{5/2}}{|t_j^n - t_{i+1}^n|^3} + \varepsilon \sum_{i<j}^{p(n)-1} |t_{i+1}^n - t_i^n| |t_{j+1}^n - t_j^n| \\
&\leq +2C''C\delta(\varepsilon) \int_0^T \int_0^T \frac{1}{(x-y)^{3/2}} dx dy + 2C''C\delta(\varepsilon)^{3/4} \int_0^T \int_0^T \frac{1}{(x-y)^{3/4}} dx dy \\
&\quad + DT(\delta(\varepsilon) + \delta(\varepsilon)^{1/2}) + \varepsilon 2T.
\end{aligned}$$

Since every term with  $\delta(\varepsilon)$  and the integrals are finite, we obtain the convergence to 0. We conclude then as in the proof Theorem 4.3.9.  $\square$

From the explicit formula (4.3.22) we deduce an interesting property of the first order rough integral with respect to  $\hat{\mathfrak{X}}$ , by means of the Gaussian nature of  $X$ .

**Proposition 4.3.16.** *For any  $\psi \in C_b^5(\mathbb{R}, \mathbb{R})$  one has*

$$\mathbb{E} \int_s^t \Psi'(X_r) d\hat{\mathfrak{X}}_r^1 = 0,$$

*Proof.* Applying the formula (4.3.22) with  $\psi$  one has

$$\int_s^t \Psi'(X_r) d\hat{\mathfrak{X}}_r^1 = \psi(X_t) - \psi(X_s) - \frac{1}{2} \int_s^t \psi''(X_r) d\sigma_r^2. \quad (4.3.28)$$

Since  $\psi$  is bounded and continuous, we can use the explicit formula of the density function of the random variable  $X_r$  to obtain that the function  $r \rightarrow \mathbb{E}[\psi(X_r)]$  is differentiable and we can derivate under the sign of integral. Recalling the notation  $G(t, x) = (2\pi t)^{-1/2} e^{-x^2/2t}$  and the fact that  $G$  satisfies the heat equation for any  $r > 0$  we have

$$\begin{aligned}
\frac{d}{dr} \mathbb{E}[\psi(X_r)] &= \frac{d}{dr} \int_{\mathbb{R}} \psi(y) G(\sigma_r^2, y) dy = \int_{\mathbb{R}} \psi(y) \partial_t G(\sigma_r^2, y) \frac{d}{dr} \sigma_r^2 dy \\
&= \frac{1}{2} \frac{d}{dr} \sigma_r^2 \int_{\mathbb{R}} \psi(y) G_{xx}(\sigma_r^2, y) dy = \frac{1}{2} \frac{d}{dr} \sigma_r^2 \int_{\mathbb{R}} \psi''(y) G(\mathbb{E}(\sigma_r^2), y) dy \\
&= \frac{1}{2} \frac{d}{dr} \sigma_r^2 \mathbb{E}[\psi''(X_r)].
\end{aligned}$$

Integrating this identity from  $s$  to  $t$  and applying the expectation to (4.3.28) we obtain the desired identity.  $\square$

*Remark 4.3.17.* Looking at the resulting formula (4.3.22), this last identity tells us that the rough integral with respect to  $\hat{\mathfrak{X}}$  has the same zero expectation of a stochastic integral. Actually this property could be a guess to describe alternatively the law of this rough integral. Indeed following the literature of Itô formulae for fractional Brownian motion, we can find in [CN05, Lem. 4.3] a formula with exactly the same second order term of (4.3.22) and a first term given by a generalisation of the Skorohod integral (whose expectation is trivially zero). Hence the name Cheridito-Nualart. By analogy with the fractional Brownian motion case we could think the rough integral with respect to  $\hat{\mathfrak{X}}$  should coincide with a suitable definition of Skorohod integral built from the Gaussian process  $X$ . This identification result could help to study some situations where the test function is not sufficiently smooth and it would help to develop a Tanaka formula on  $X$ .

### 4.3.3 Quartic variation formula

We conclude the section by showing the existence of an another one dimensional quasi-geometric rough path over  $X$  with a non trivial fourth order rough integral in (4.3.7). The construction of this new object is based upon the existence of a so called “quartic variation” for the process  $X$  (see [Swa07, Thm 2.3]). We recall here this property

**Theorem 4.3.18.** *For any sequence of partitions  $\mathcal{P}_n$  of  $[0, t]$  whose mesh-size  $|\mathcal{P}_n| \rightarrow 0$  the following sequence of quartic increments*

$$V_t^n := \sum_{[u,v] \in \mathcal{P}_n} (X_v - X_u)^4,$$

*satisfies the property*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \left| V_t^n - \frac{6}{\pi} t \right|^2 \right] \rightarrow 0. \quad (4.3.29)$$

*Proof.* We repeat here the proof of this result from [Swa07] for sake of completeness. Since  $V_t^n$  is monotone in  $t$ , it is sufficient to prove the convergence in  $L^2(\mathbb{P})$   $V_t^n \rightarrow (6/\pi)t$  for each fixed  $t$ . Using the notation  $\delta_{st}X = X_t - X_s$  we can add and subtract for any subdivision  $[u, v] \in \mathcal{P}_n$  the quantity  $A_{uv} = \mathbb{E}(\delta_{uv}X)^4 = 3(\mathbb{E}(\delta_{uv}X)^2)^2$  and we obtain the decomposition:

$$V_t^n - \frac{6}{\pi}t = \sum_{[u,v] \in \mathcal{P}_n} ((\delta_{uv}X)^4 - A_{uv}) + \sum_{[u,v] \in \mathcal{P}_n} \left( A_{uv} - \frac{6}{\pi}(v - u) \right). \quad (4.3.30)$$

Thus the theorem will follow by showing that the two sums on the right hand side of (4.3.30) converge both to 0 in  $L^2(\mathbb{P})$ . By means of the estimate (4.3.3) obtained in the proof of the Proposition 4.3.1 and the moment estimate (4.3.4) there exist two

constant  $C, C' > 0$  such that for any subdivision  $[u, v] \in \mathcal{P}_n$  one has

$$\begin{aligned} |A_{uv} - \frac{6}{\pi}(v-u)| &= 3 \left( \mathbb{E}[(\delta_{uv}X)^2] - \sqrt{\frac{2(v-u)}{\pi}} \right) \left( \mathbb{E}[(\delta_{uv}X)^2] + \sqrt{\frac{2(v-u)}{\pi}} \right) \\ &\leq C \frac{|v-u|^{5/2}}{v^{3/2}} \leq C' \frac{|v-u|^{7/4}}{v^{3/4}} \leq C' |\mathcal{P}_n|^{3/4} \frac{|v-u|}{v^{3/4}}. \end{aligned}$$

Summing this identity over all  $[u, v] \in \mathcal{P}_n$  and sending  $n \rightarrow \infty$  one has

$$\left| \sum_{[u,v] \in \mathcal{P}_n} \left( A_{uv} - \frac{6}{\pi}(v-u) \right) \right| \leq C' |\mathcal{P}_n|^{3/4} \sum_{[u,v] \in \mathcal{P}_n} \frac{|v-u|}{v^{3/4}} \rightarrow 0,$$

because the remaining sum converges to the finite integral  $\int_0^t x^{-3/4} dx$ . We pass to the other sum Writing the square of the first sum as a double sum we obtain

$$\begin{aligned} &\mathbb{E} \left( \sum_{[u,v] \in \mathcal{P}_n} ((\delta_{uv}X)^4 - A_{uv}) \right)^2 = \\ &= \sum_{[u,v] \in \mathcal{P}_n} \sum_{[u',v'] \in \mathcal{P}_n} \mathbb{E} [((\delta_{uv}X)^4 - A_{uv})(\delta_{u'v'}X)^4 - A_{u'v'}] \\ &= \sum_{[u,v] \in \mathcal{P}_n} \sum_{[u',v'] \in \mathcal{P}_n} |\mathbb{E}[(\delta_{uv}X)^4(\delta_{u'v'}X)^4] - 9\mathbb{E}[(\delta_{uv}X)^2]^2\mathbb{E}[(\delta_{u'v'}X)^2]^2|. \end{aligned}$$

Using the Wick formula for Gaussian variables for any couple of subdivisions  $[u, v] \in \mathcal{P}_n$ ,  $[u', v'] \in \mathcal{P}_n$  and the Hölder inequality there exists a constant  $D > 0$  such that

$$\begin{aligned} &\mathbb{E}[(\delta_{uv}X)^4(\delta_{u'v'}X)^4] - 9\mathbb{E}[(\delta_{uv}X)^2]^2\mathbb{E}[(\delta_{u'v'}X)^2]^2 = \\ &= 24\mathbb{E}[(\delta_{uv}X)(\delta_{u'v'}X)]^4 + 72\mathbb{E}[(\delta_{u'v'}X)^2]\mathbb{E}[(\delta_{uv}X)^2]\mathbb{E}[(\delta_{uv}X)(\delta_{u'v'}X)]^2 \\ &\leq D \left( \mathbb{E}[(\delta_{u'v'}X)^2]\mathbb{E}[(\delta_{uv}X)^2]\mathbb{E}[(\delta_{uv}X)(\delta_{u'v'}X)]^2 \right). \end{aligned}$$

Summing this estimate with (4.3.4) there exists a constant  $D' > 0$  such that

$$\begin{aligned} &\mathbb{E} \left( \sum_{[u,v] \in \mathcal{P}_n} ((\delta_{uv}X)^4 - A_{uv}) \right)^2 \leq \\ &\leq D' \sum_{[u,v] \in \mathcal{P}_n} \sum_{[u',v'] \in \mathcal{P}_n} |v' - u'|^{1/2} |v - u|^{1/2} \mathbb{E}[(\delta_{uv}X)(\delta_{u'v'}X)]^2. \end{aligned}$$

Since the terms in the diagonal part of this sum converge trivially, the result will follow by showing that

$$I^n := \sum_{[u,v] \cap [u',v'] = \emptyset} |v' - u'|^{1/2} |v - u|^{1/2} \mathbb{E}[(\delta_{uv}X)(\delta_{u'v'}X)]^2 \rightarrow 0.$$

Using the estimates (4.3.19), for any couple  $[u, v] \cap [u', v'] = \emptyset$  we have

$$\mathbb{E}[(\delta_{uv}X)(\delta_{u'v'}X)]^2 \leq D'' \frac{|v' - u'|^2 |v - u|^2}{|u - u'|^2 |v - v'|} \leq M |\mathcal{P}_n|^{3/4} \frac{|v' - u'|^{1/2} |v - u|^{1/2}}{|u - v'|^{3/4}},$$

for some finite constants  $D'', M > 0$ . Summing up this last inequality we obtain

$$I^n \leq M |\mathcal{P}_n|^{3/4} \sum_{[u,v] \cap [u',v'] = \emptyset} \frac{|v' - u'| |v - u|}{|u - v'|^{3/4}} \rightarrow 0,$$

because the remaining sum converges to the finite integral  $\int_{[0,t]^2} |x - y|^{-3/4} dx dy$ .  $\square$

*Remark 4.3.19.* Starting from  $\mathcal{Q}_n$  a sequence of partitions of an interval  $[s, t]$  it follows trivially from the previous proof that  $V_{st}^n$ , the corresponding sequence of quartic increments defined on  $\mathcal{Q}_n$ , satisfies also the property

$$\mathbb{E} \left[ \sup_{s < t \in [0, T]} \left| V_{st}^n - \frac{6}{\pi} (t - s) \right|^2 \right] \rightarrow 0. \tag{4.3.31}$$

Since the usual quadratic variation of a process satisfies a property similar to (4.3.31), we call the function  $t \rightarrow (6/\pi)t$  the quartic variation of  $X$ .

Using the formalism of the one dimensional quasi-geometric rough paths, we can easily construct a rough path where the quartic variation of  $X$  becomes the 4-th order variation of the process  $X$  in the context of the rough change of variable formula.

**Proposition 4.3.20.** *There exists a unique  $\mathbb{X} \in \mathcal{R}_g^\Gamma(\{1, 4\})$  such that*

$$\langle \mathbb{X}_{st}, 1 \rangle := X_t - X_s, \quad \langle \mathbb{X}_{st}, 4 \rangle := \frac{6}{\pi} (t - s). \tag{4.3.32}$$

*Considering  $\mathbb{X}^0$ , the zero extension of  $\mathbb{X}$ , we call  $\mathcal{X} = \log^* \mathbb{X}^0$  the quartic variation extension of  $X$ .*

*Proof.* The result follows straightforwardly from the application of Lemma 4.2.15 and the general properties of the process  $X$ . Indeed we obtain immediately from the homogeneity of the words that  $L^\Gamma(\{1, 4\})$ , the set of Lyndon words with small homogeneity becomes simply the set  $\{1, 4\}$  and the conditions (4.3.32) identify uniquely a map  $\mathbb{Z}: [0, T]^2 \rightarrow T^\Gamma(\{1, 4\})$  satisfying a.s. the hypothesis of the extension Theorem (see Theorem 3.2.7). Thereby obtaining the thesis.  $\square$

Similarly with the previous sections, we look at the associated change of variable formula induced by the choice of  $\mathcal{X}$  in (4.3.7).

**Theorem 4.3.21.** *Let  $\mathcal{X}$  be the quartic variation extension of  $X$  and  $\psi \in C_b^1(\mathbb{R}, \mathbb{R})$ . Then we have the a.s. equality*

$$\int_s^t \Psi(X_r) d\mathcal{X}_r^4 = \frac{6}{\pi} \int_s^t \psi(X_r) dr, \tag{4.3.33}$$

Therefore for any  $\varphi \in C_b^5(\mathbb{R}, \mathbb{R})$  we have the identity

$$\begin{aligned} \varphi(X_t) = & \varphi(X_s) + \int_s^t \Phi'(X_r) d\mathcal{X}_r^1 + \frac{1}{2!} \int_s^t \Phi''(X_r) d\mathcal{X}_r^2 \\ & + \frac{1}{3!} \int_s^t \Phi^{(3)}(X_r) d\mathcal{X}_r^3 + \frac{6}{4!\pi} \int_s^t \varphi^{(4)}(X_r) dr. \end{aligned} \quad (4.3.34)$$

Moreover the sum of remaining rough integrals satisfies the identity

$$\sum_{l=1}^3 \frac{1}{l!} \int_s^t \Phi^{(l)}(X_r) d\mathcal{X}_r^l = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \sum_{l=1}^3 \varphi^{(l)}(X_u) \frac{(X_v - X_u)^l}{l!}, \quad (4.3.35)$$

where the right hand side of (4.3.35) converges in  $L^2(\mathbb{P})$  uniformly on compact sets for any sequence of partitions  $\mathcal{P}_n$  of  $[s, t]$  whose mesh-size  $|\mathcal{P}_n| \rightarrow 0$ .

*Proof.* The identity (4.3.33) follows easily from the general definition of the rough integral given in equation (4.2.2). Indeed for any sequence of partitions  $\mathcal{P}_n$  of  $[s, t]$  whose mesh-size  $|\mathcal{P}_n| \rightarrow 0$  we obtain immediately

$$\int_s^t \Psi(X_r) d\mathcal{X}_r^4 = \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \psi(X_u) \langle \mathcal{X}_{uv}, 4 \rangle = \lim_{n \rightarrow \infty} \frac{6}{\pi} \sum_{[u,v] \in \mathcal{P}_n} \psi(X_u) (v - u)$$

and the identity (4.3.33) follows trivially. In order to prove the relation (4.3.35) we will write down the approximating sum (4.2.2) defining each rough integral in the left hand side of (4.3.35). In particular, using the definition of the map  $\log$  and the zero extension operation we have the following algebraic identities

$$\begin{aligned} \frac{1}{3!} \sum_{u \in W_3^{\Gamma}(\mathbb{A}_4)} \frac{\varphi^{(|u|_{\Gamma}+3)}(X_u)}{u!} \langle \mathcal{X}_{uv}, u 3 \rangle &= \frac{1}{3!} (\varphi^{(3)}(X_u) \langle \mathbb{X}_{uv}, 3 \rangle + \varphi^{(4)}(X_u) \langle \mathbb{X}_{uv}^0, \log(13) \rangle) \\ &= -\frac{1}{12} \varphi^{(4)}(X_u) \langle \mathbb{X}_{uv}, 4 \rangle; \end{aligned}$$

$$\begin{aligned} \frac{1}{2!} \sum_{u \in W_2^{\Gamma}(\mathbb{A}_4)} \frac{\varphi^{(|u|_{\Gamma}+2)}(X_u)}{u!} \langle \mathcal{X}_{uv}, u 1 \rangle &= \frac{1}{2!} \left( \varphi''(X_u) \langle \mathbb{X}_{uv}, 2 \rangle + \varphi^{(3)}(X_u) \langle \mathbb{X}_{uv}^0, \log(12) \rangle \right. \\ &+ \left. \varphi^{(4)}(X_u) \frac{1}{2} \langle \mathbb{X}_{uv}^0, \log(22) \rangle + \varphi^{(4)}(X_u) \langle \mathbb{X}_{uv}^0, \log(112) \rangle \right) = \frac{1}{4!} \varphi^{(4)}(X_u) \langle \mathbb{X}_{uv}, 4 \rangle; \end{aligned}$$

$$\begin{aligned} \sum_{u \in W_1^{\Gamma}(\mathbb{A}_4)} \frac{\varphi^{(|u|_{\Gamma}+1)}(X_u)}{u!} \langle \mathcal{X}_{uv}, u 1 \rangle &= \varphi'(X_u) \langle \mathbb{X}_{uv}, 1 \rangle + \varphi''(X_u) \langle \mathbb{X}_{uv}^0, \log(11) \rangle \\ &+ \varphi'''(X_u) \langle \mathbb{X}_{uv}^0, \log(111) \rangle + \varphi^{(4)}(X_u) \frac{1}{2} \langle \mathbb{X}_{uv}^0, \log(211) \rangle + \varphi^{(4)}(X_u) \langle \mathbb{X}_{uv}^0, \log(1111) \rangle \\ &+ \varphi^{(4)}(X_u) \frac{1}{2} \langle \mathbb{X}_{uv}^0, \log(121) \rangle + \varphi^{(4)}(X_u) \frac{1}{3!} \langle \mathbb{X}_{uv}^0, \log(31) \rangle \\ &= \varphi'(X_u) \langle \mathbb{X}_{uv}, 1 \rangle + \varphi''(X_u) \langle \mathbb{X}_{uv}, 11 \rangle + \varphi^{(3)}(X_u) \langle \mathbb{X}_{uv}, 111 \rangle + \varphi^{(4)}(X_u) \langle \mathbb{X}_{uv}, 1111 \rangle. \end{aligned}$$

Summing up all these three identities and recalling the definition of  $\mathbb{X}$  we obtain

$$\begin{aligned} & \sum_{l=1}^3 \frac{1}{l!} \sum_{u \in W_l^\Gamma(\mathbb{A}_4)} \frac{\varphi^{(|u|_{\Gamma}+l)}(X_u)}{u!} \langle \mathcal{X}_{uv}, u l \rangle = \\ & \sum_{l=1}^3 \varphi^{(l)}(X_u) \frac{(X_v - X_u)^l}{l!} + \frac{1}{4!} \varphi^{(4)}(X_u) [(X_v - X_u)^4 - \frac{6}{\pi}(v - u)]. \end{aligned} \quad (4.3.36)$$

Since the derivatives of  $\varphi$  are bounded, for any for any sequence of partitions  $\mathcal{P}_n$  of  $[s, t]$  whose mesh-size  $|\mathcal{P}_n| \rightarrow 0$  the identity (4.3.36) implies

$$\begin{aligned} & \mathbb{E} \left[ \sum_{[u,v] \in \mathcal{P}_n} \sum_{l=1}^3 \frac{1}{l!} \sum_{u \in W_l^\Gamma(\mathbb{A}_4)} \frac{\varphi^{(|u|_{\Gamma}+l)}(X_u)}{u!} \langle \mathcal{X}_{uv}, u l \rangle - \sum_{l=1}^3 \varphi^{(l)}(X_u) \frac{(X_v - X_u)^l}{l!} \right]^2 \\ & \leq M \mathbb{E} \left[ \sum_{[u,v] \in \mathcal{P}_n} [(X_v - X_u)^4 - \frac{6}{\pi}(v - u)] \right]^2 \end{aligned}$$

for some constant  $M > 0$ . Thus we can easily apply the Theorem 4.3.18 and we deduce that the compensated Riemann sum in the right hand side of (4.3.35) converges uniformly in  $L^2(\mathbb{P})$  to the a.s. limit of

$$\sum_{[u,v] \in \mathcal{P}_n} \sum_{l=1}^3 \frac{1}{l!} \sum_{u \in W_l^\Gamma(\mathbb{A}_4)} \frac{\varphi^{(|u|_{\Gamma}+l)}(X_u)}{u!} \langle \mathcal{X}_{uv}, u l \rangle.$$

Thereby obtaining the thesis.  $\square$

*Remark 4.3.22.* Following the proof of the result we remark an interesting property of the second and third order rough integrals given in (4.3.34). Indeed looking at the sum defining them, it is straightforward to show that these integrals do not cancel even if the second and third order path are zero! This property, which is impossible in the usual integration theory, is perfectly coherent in the context of the rough integration, where the integrals are not depending only on the increments of a path. The compensated Riemann sum in the right hand side of (4.3.35) was already introduced in [CP19] in the context of the *path-wise* stochastic calculus. However differently from that context the limit of this sum does not depend on the sequence of partitions defining it.

## 4.4 Stratonovich type formula

We conclude the chapter by studying the rough change of variable formula (4.3.5) with respect to  $\mathbf{X}$ , the canonical geometric extension of the process  $X$ . Even if the rough integral  $\int \Phi'(X_r) d\mathbf{X}_r^1$  is defined as the a.s. limit an explicit corrected Riemann, the explicit form of the formula

$$\varphi(X_t) = \varphi(X_s) + \int \Phi'(X_r) d\mathbf{X}_r^1$$

reminds us the standard Stratonovich calculus. Actually an interesting statement contained in [BS10] shows that it is also possible to define an effective Stratonovich integral over  $X$ . For the proof of this result we refer to [BS10, Cor. 4.5].

**Theorem 4.4.1.** *For any  $g \in C^7(\mathbb{R}, \mathbb{R})$  we introduce the sequence of trapezoid-type approximations*

$$T_t^n(g') := \sum_{j=1}^{\lfloor nt \rfloor} \frac{g'(X_{t_{j-1}}) + g'(X_{t_j})}{2} (X_{t_j} - X_{t_{j-1}}), \quad t_j = \frac{j}{n}, \quad t \in [0, T]. \quad (4.4.1)$$

*Then the sequence  $T_t^n$  converges in probability uniformly on  $t \in [0, T]$  to some continuous process  $\int_0^t g'(X_r) d^\circ X_r$  which satisfies the identity*

$$g(X_t) = g(X_0) + \int_0^t g'(X_r) d^\circ X_r. \quad (4.4.2)$$

Comparing the Theorem 4.4.1 with the Proposition 4.3.2 we deduce immediately that for any sufficiently smooth function  $\psi: \mathbb{R} \rightarrow \mathbb{R}$  one has the a.s. identity

$$\int_s^t \Psi(X_r) d\mathbf{X}_r^1 = \int_s^t \psi(X_r) d^\circ X_r,$$

but the approximating sums of both integrals do not coincide at all. In what follows we will study the general relations between the compensated Riemann sums defining the rough integral in (4.3.6) and some generalisations of the trapezoid-type sums in a slightly general framework, in order to provide a simplified proof of the Theorem 4.4.1.

#### 4.4.1 A general identity between symmetric integrals

In order to state a general relation between rough integral and trapezoid sums, we put aside the specific process  $X$  and we consider a generic stochastic process  $Z$  admitting a modification  $\mathcal{C}^\gamma(\mathbb{R})$  for some  $\gamma \in (0, 1)$ , we recall the parameter  $N = \lfloor \gamma^{-1} \rfloor$  and we denote by  $\mathbf{Z} \in \mathcal{R}_g^\gamma(\{1\})$  the canonical geometric extension of  $Z$ . The operation we will consider in this subsection is the rough integration of a generic one dimensional random controlled rough path  $\mathbf{Y} \in \mathcal{D}^{N\gamma}(\mathbf{Z})$  (see the Definition 3.3.1) with respect to  $\mathbf{Z}$ . The general operation is given in the Proposition 3.3.6 but in this one dimensional case we have

$$\int_0^t \mathbf{Y}_r d\mathbf{Z}_r := \lim_{n \rightarrow \infty} \sum_{[u,v] \in \mathcal{P}_n} \sum_{k=0}^{N-1} \langle 1^{\circ k}, \mathbf{Y}_u \rangle \frac{(Z_v - Z_u)^k}{k!}, \quad (4.4.3)$$

defined for any sequence of partitions of  $[0, t]$ ,  $\mathcal{P}_n = \{0 = t_0^n < \dots < t_{p(n)}^n = t\}$  such that its mesh size  $|\mathcal{P}_n| := \sup_i (t_{i+1}^n - t_i^n) \rightarrow 0$  (see the Proposition 3.3.6). This quantity can be related with a different family of stochastic integrals defined from  $Z$ .

**Definition 4.4.2.** Let  $Y$  be a continuous process on  $[0, T]$  and  $m \geq 1$ . If there exists a sequence of partitions of  $[0, t]$   $\mathcal{Q}_n$  satisfying  $|\mathcal{Q}_n| \rightarrow 0$  such that the limit in probability

$$\lim_{N \rightarrow +\infty} \sum_{t_i \in \mathcal{Q}_n} \frac{Y_{t_i} + Y_{t_{i+1}}}{2} (Z_{t_{i+1}} - Z_{t_i})^m, \quad (4.4.4)$$

exists uniformly on  $t \in [0, T]$ , we call the limiting process the  $m$ -th order symmetric integral of  $Y$  with respect to  $Z$  and we will denote it by the symbol  $\int_0^t Y_r d^{(m)}Z_r$ .

*Remark 4.4.3.* The original notion of  $m$ -th order symmetric integral was developed in [GNRV05, Def. 2.7], [GRV03, Pag. 9] to study a generalisation of the Itô formula for the fractional Brownian motion. This idea was conceived by extending the theory of stochastic calculus by regularization developed by Russo, Vallois and coauthors (see e.g. [RV93], [RV95]) in order to generalise the Stratonovich integration at higher orders. In that literature the  $m$ -th order symmetric integral is defined as

$$\int_0^t Y_r d^{(m)}Z_r := \lim_{\varepsilon \rightarrow 0^+} \int_0^t \frac{Y_{u+\varepsilon} + Y_u}{2\varepsilon} (Z_{u+\varepsilon} - Z_u)^m du, \quad (4.4.5)$$

where the convergence is again a convergence in probability and uniformly on  $[0, T]$ . The two notions are deeply related. Indeed by taking  $\mathcal{U}_n$ , the uniform partition of  $[0, t]$

$$\mathcal{U}_n = \left\{ \frac{kt}{n} : k = 0, \dots, n \right\},$$

we can introduce for any  $\varepsilon > 0$  the regularised symmetric sum process

$$S_t^{n,\varepsilon} := \sum_{t_i \in \mathcal{U}_n} \frac{Y_{t_i} + Y_{t_i+\varepsilon}}{2\varepsilon} (Z_{t_i+\varepsilon} - Z_{t_i})^m \frac{t}{N}.$$

Observing that

$$S_t^{n,t/n} = \sum_{t_i \in \mathcal{U}_n} \frac{Y_{t_i} + Y_{t_{i+1}}}{2} (Z_{t_{i+1}} - Z_{t_i})^m,$$

for any  $t \in [0, T]$  and  $\varepsilon > 0$  we obtain

$$\lim_{n \rightarrow +\infty} S_t^{n,\varepsilon} = \int_0^t \frac{Y_{u+\varepsilon} + Y_u}{2\varepsilon} (Z_{u+\varepsilon} - Z_u)^m du \quad \text{a.s.} \quad (4.4.6)$$

because the process  $S_t^{N,\varepsilon}$  is the Riemann sum associated to the integral in the right hand side of (4.4.6) and the processes  $Y$  and  $Z$  are continuous. Thus the two notions coincide as long as we have the identity

$$\lim_{n \rightarrow +\infty} S_t^{n,t/n} = \lim_{\varepsilon \rightarrow 0^+} \lim_{n \rightarrow +\infty} S_t^{n,\varepsilon}$$

and the sum (4.4.4) converges when we take the partition  $\mathcal{U}_n$ .



In general the definition of a symmetric integral following the Definition 4.4.2 might depend on the approximating sequence of partitions  $\mathcal{Q}_n$ . Then we will always specify which sequence we are choosing. The explicit definition of  $\mathbf{Z}$  as the canonical geometric extension of  $Z$  in (4.2.6) and the definition the  $m$ -th order symmetric integral suggest there should be a link between the rough integration with respect to  $Z$  and the integrals defined in (4.4.4). An explicit formula can be given.

**Theorem 4.4.4.** *Let  $\gamma \in (0, 1)$  and  $N = \lfloor \gamma^{-1} \rfloor > 1$ . Denoting for any random  $\mathbf{Y} \in \mathcal{D}^{N\gamma}(\mathbf{Z})$  the processes  $Y_t^{(i)} := \langle 1^{\circ i}, \mathbf{Y}_t \rangle$ , one has the identity*

$$\int_0^t \mathbf{Y}_r d\mathbf{Z}_r = \sum_{j=0}^{\lfloor N/2 \rfloor - 1} \alpha_j \int_0^t Y_r^{(2j)} d^{\circ(2j+1)} Z_r, \quad a.s. \tag{4.4.7}$$

as long as there exists  $\mathcal{Q}_n$  a deterministic sequence of partition of  $[0, t]$  satisfying  $|\mathcal{Q}_n| \rightarrow 0$  such that only  $\lfloor N/2 \rfloor - 1$  integrals in the right hand side of (4.4.7) are well defined. The sequence  $\{\alpha_l\}_{l \geq 0}$  is given recursively by the conditions:

$$\begin{cases} \alpha_0 = 1 \\ \alpha_{l+1} = \frac{1}{(2l+3)!} - \sum_{k=0}^l \frac{\alpha_k}{2(2l+2-2k)!} \end{cases} \tag{4.4.8}$$

We will prove the identity (4.4.7) by means of three preliminary lemmas. The the first couple of results involves the sequence  $\{\alpha_l\}_{l \geq 0}$  given in (4.4.8). Here are its first six terms

$$1, -\frac{1}{12}, \frac{1}{120}, -\frac{17}{20160}, \frac{5}{58529}, -\frac{2}{231067},$$

No explicit formula is known for  $\alpha_l$ . In order to prove a general identity between its terms, we consider its generating function

$$F(x) := \sum_{l=0}^{+\infty} \alpha_l x^l \tag{4.4.9}$$

which exists at least for any  $x \in (-1, 1)$ , since it is straightforward to prove from the definition in (4.4.8) that  $|\alpha_l| \leq 1$  for every  $l \geq 0$ .

**Lemma 4.4.5.** *for any  $x \in [0, 1)$  one has*

$$F(x) = \frac{2 \sinh(x^{1/2})}{x^{1/2}(1 + \cosh(x^{1/2}))} \tag{4.4.10}$$

where the right hand side can be extended smoothly on 0.

*Proof.* For  $x = 0$  this identity holds trivially. Let  $x \in (0, 1)$ ; we fix  $N > 0$  and and we sum the recursive identity of (4.4.8) until  $N$  obtaining

$$\sum_{l=0}^N \alpha_{l+1} x^l = \sum_{l=0}^N \frac{1}{(2l+3)!} x^l - \sum_{l=0}^N \left( \sum_{k=0}^l \frac{\alpha_k}{2(2l+2-2k)!} \right) x^l. \tag{4.4.11}$$

Using Cauchy's product formula, the right hand side of (4.4.11) converges to

$$\sum_{l=0}^{+\infty} \frac{1}{(2l+3)!} x^l - \frac{1}{2} \left( \sum_{l=0}^{+\infty} \frac{1}{(2l+2)!} x^l \right) F(x) =: H(x) - \frac{1}{2} G(x) F(x).$$

Thanks to some elementary Taylor's series formulae we can write

$$H(x) = \frac{\sinh(x^{1/2}) - x^{1/2}}{x^{3/2}}, \quad G(x) = \frac{\cosh(x^{1/2}) - 1}{x}. \quad (4.4.12)$$

On the other hand, the right hand side of (4.4.11) becomes

$$\sum_{l=0}^{+\infty} \alpha_{l+1} x^l = \frac{F(x) - 1}{x}.$$

By uniqueness of the limit one has

$$\frac{F(x) - 1}{x} = H(x) - \frac{1}{2} G(x) F(x) \iff F(x) = \frac{xH(x) + 1}{1 + \frac{x}{2}G(x)} = \frac{2xH(x) + 2}{2 + xG(x)}.$$

Thus we obtain the thesis.  $\square$

**Lemma 4.4.6.** *For every integer  $l \geq 0$  one has*

$$\sum_{k=0}^l \frac{\alpha_k}{2(2l+1-2k)!} = \frac{1}{(2l+2)!} \quad (4.4.13)$$

*Proof.* We will prove this identity by calculating explicitly the generating function of both sides of the equation (4.4.13), obtaining an identity if both functions coincide on a right neighbourhood of 0. The right hand side of (4.4.13) was already calculated and it is given by the function  $G$  defined in (4.4.12). On the other hand, using again Cauchy's product formula one has for every  $x \in [0, 1)$

$$\sum_{l=0}^{+\infty} \sum_{k=0}^l \frac{\alpha_k}{2(2l+1-2k)!} x^l = \frac{1}{2} G_1(x) F(x),$$

where the function  $G_1$  is given by for any  $x \in [0, 1)$  by

$$G_1(x) := \sum_{p=0}^{+\infty} \frac{1}{(2p+1)!} x^p = \frac{\sinh(x^{1/2})}{x^{1/2}}.$$

From an elementary calculation we obtain for any  $x \in [0, 1)$

$$\frac{1}{2} G_1(x) F(x) = \frac{\sinh^2(x^{1/2})}{x(1 + \cosh(x^{1/2}))} = \frac{\cosh^2(x^{1/2}) - 1}{x(1 + \cosh(x^{1/2}))} = G(x).$$

Thereby obtaining the thesis.  $\square$

**Lemma 4.4.7.** *Let  $\gamma \in (0, 1)$ ,  $N = \lfloor \gamma^{-1} \rfloor > 1$  and  $\mathbf{Y} \in \mathcal{D}^{N\gamma}(\mathbb{X})$  a random controlled rough path. Then there exists  $\varepsilon > 0$  and a positive random variable  $K$  such that*

$$\left| \sum_{j=0}^{\lfloor N/2 \rfloor - 1} \alpha_j \left( \frac{Y_t^{(2j)} + Y_s^{(2j)}}{2} \right) (Z_t - Z_s)^{2j+1} - \sum_{i=0}^{N-1} Y_s^i \frac{(Z_t - Z_s)^{i+1}}{(i+1)!} \right| \leq K(\omega) |t - s|^{1+\varepsilon}.$$

*Proof.* This result is a direct consequence of the controlled rough path hypothesis on  $\mathbf{Y}$  given in the Definition 3.3.2. Indeed by definition of  $\mathbf{Y}$  there exists a positive random variable  $C'$  not depending on  $t, s$  such that for any  $k \geq 0$  one has

$$|Y_t^{(2k)} - \langle \mathbf{Z}_{st} \otimes 1^{\otimes 2k}, \nabla \mathbf{Y}_s \rangle| \leq C'(\omega) |t - s|^{(N-2k)\gamma}. \quad (4.4.14)$$

Writing  $\mathbf{Y}_t = \sum_{l=0}^{N-1} Y_t^{(l)} 1^{\otimes l}$  and recalling the definition of  $\mathbf{Z}$ , the bound (4.4.14) becomes

$$\left| Y_t^{(2k)} - \sum_{l=2k}^{N-1} Y_s^{(l)} \frac{(Z_t - Z_s)^{l-2k}}{(l-2k)!} \right| \leq C'(\omega) |t - s|^{(N-2k)\gamma}.$$

By multiplying both sides by  $|\alpha_k (Z_t - Z_s)^{2k+1}|/2$  there exists a positive random variable  $C''$  satisfying

$$\left| \alpha_k \frac{Y_t^{(2k)}}{2} (Z_t - Z_s)^{2k+1} - \alpha_k \sum_{l=2k}^{N-1} \frac{Y_s^{(l)}}{2} \frac{(Z_t - Z_s)^{l+1}}{(l-2k)!} \right| \leq C''(\omega) |t - s|^{1+\varepsilon}, \quad (4.4.15)$$

where  $1 + \varepsilon = (N + 1)\gamma$ . Summing the identity (4.4.15) from  $k = 0$  to  $k = \lfloor N/2 \rfloor - 1$  all the terms inside the absolute value there exists a random variable  $K > 0$  such that

$$\left| \sum_{k=0}^{\lfloor N/2 \rfloor - 1} \left( \alpha_k \frac{Y_t^{(2k)}}{2} (Z_t - Z_s)^{2k+1} - \alpha_k \sum_{l=2k}^{N-1} \frac{Y_s^{(l)}}{2} \frac{(Z_t - Z_s)^{l+1}}{(l-2k)!} \right) \right| \leq K(\omega) |t - s|^{1+\varepsilon}. \quad (4.4.16)$$

By adding and subtracting the quantity  $\alpha_k Y_s^{(2k)}/2 (Z_t - Z_s)^{2k+1}$  inside the sum of the left hand side of (4.4.16) we rewrite this term as

$$\begin{aligned} & \sum_{k=0}^{\lfloor N/2 \rfloor - 1} \alpha_k \frac{Y_t^{(2k)} + Y_s^{(2k)}}{2} (Z_t - Z_s)^{2k+1} - \sum_{k=0}^{\lfloor N/2 \rfloor - 1} \alpha_k Y_s^{(2k)} (Z_t - Z_s)^{2k+1} \\ & - \sum_{k=0}^{\lfloor N/2 \rfloor - 1} \sum_{l=2k+1}^{N-1} \alpha_k Y_s^{(l)} \frac{(Z_t - Z_s)^{l+1}}{2(l-2k)!} \\ = & \sum_{k=0}^{\lfloor N/2 \rfloor - 1} \alpha_k \frac{Y_t^{(2k)} + Y_s^{(2k)}}{2} (Z_t - Z_s)^{2k+1} - \sum_{k=0}^{\lfloor N/2 \rfloor - 1} \alpha_k Y_s^{(2k)} (Z_t - Z_s)^{2k+1} \\ & - \sum_{j=1}^{N-1} Y_s^{(j)} (Z_t - Z_s)^{j+1} \sum_{k=0}^{\lfloor (j-1)/2 \rfloor} \frac{\alpha_k}{2(j-2k)!} \\ = & \sum_{k=0}^{\lfloor N/2 \rfloor - 1} \alpha_k \frac{Y_t^{(2k)} + Y_s^{(2k)}}{2} (Z_t - Z_s)^{2k+1} - \sum_{j=0}^{N-1} \beta_j Y_s^{(j)} (Z_t - Z_s)^{j+1}, \end{aligned}$$

where the sequence  $\{\beta_j\}_{j \geq 0}$  is defined by the conditions

$$\beta_j := \begin{cases} \alpha_0 & j = 0 \\ \sum_{k=0}^l \frac{\alpha_k}{2^{(j-2k)!}} & j = 2l + 1 \\ \alpha_{l+1} + \sum_{k=0}^l \frac{\alpha_k}{2^{(j-2k)!}} & j = 2l + 2 \end{cases}$$

Therefore the thesis is verified as long as  $\beta_i = 1/(i+1)!$  for every  $i \geq 0$ . By definition of  $\{\alpha_l\}_{l \geq 0}$ , one has  $\beta_{2l+2} = 1/(2l+3)!$  and the equality  $\beta_{2l+1} = 1/(2l+2)!$  is the content of the Proposition 4.4.6.  $\square$

*Proof of Theorem 4.4.4.* We suppose without loss of generality that all the terms

$$\sum_{t_i \in \mathcal{Q}_n} \frac{Y_{t_i}^{(2j)} + Y_{t_{i+1}}^{(2j)}}{2} (Z_{t_{i+1}} - Z_{t_i})^{2j+1}$$

converge to the corresponding symmetric integral for any  $j \geq 1$ . Thus the theorem will follow by showing that the sum

$$S_{\mathcal{Q}_n} := \alpha_0 \sum_{t_i \in \mathcal{Q}_n} \frac{Y_{t_{i+1}}^{(0)} + Y_{t_i}^{(0)}}{2} (Z_{t_{i+1}} - Z_{t_i}) = \sum_{t_i \in \mathcal{Q}_n} \frac{Y_{t_{i+1}}^{(0)} + Y_{t_i}^{(0)}}{2} (Z_{t_{i+1}} - Z_{t_i})$$

converges in probability to

$$\int_0^t \mathbf{Y}_r \mathbf{Z}_r - \sum_{j=1}^{\lfloor N/2 \rfloor - 1} \alpha_j \int_0^t Y_r^{(2j)} d^{(2j+1)} Z_r,$$

uniformly on  $t \in [0, T]$ . Thanks to the Lemma 4.4.7, there exists a positive random variable  $K$  and  $\varepsilon > 0$  not depending on  $t$  and a sequence of processes  $R_t^n$  such that

$$\begin{aligned} S_{\mathcal{Q}_n} &= \sum_{t_i \in \mathcal{Q}_n} \sum_{i=0}^{N-1} Y_s^{(i)} \frac{(Z_{t_{i+1}} - Z_{t_i})^{i+1}}{(i+1)!} \\ &\quad - \sum_{j=1}^{\lfloor N/2 \rfloor - 1} \sum_{t_i \in \mathcal{Q}_n} \alpha_j \left( \frac{Y_{t_i} + Y_{t_{i+1}}}{2} (Z_{t_{i+1}} - Z_{t_i})^{2j+1} \right) + R_t^n, \\ |R_t^n| &\leq K(\omega) \sum_{t_i \in \mathcal{Q}_n} |t_{i+1} - t_i|^{1+\varepsilon} \leq 2K(\omega) T |\mathcal{Q}_n|^\varepsilon. \end{aligned}$$

Thus  $R_t^n$  converges a.s. to 0 uniformly on  $t \in [0, T]$ . Since the approximations of the rough integral in (4.4.3) converge also a.s. and uniformly on  $t \in [0, T]$  and the remaining terms converge by hypothesis we obtain the desired result.  $\square$

By choosing  $\mathbf{Y}_t = \Phi'(Z_t)$  where  $\langle 1^{\otimes i}, \mathbf{Y}_t \rangle = \varphi^{(i+1)}(Z_t)$ , we combine the Theorem 4.4.4 with the formula (4.2.4) to obtain a change of variable formula with the symmetric integrals.

**Corollary 4.4.8.** *Let  $N = \lfloor \gamma^{-1} \rfloor$ . For any  $\varphi \in C_b^{N+1}(\mathbb{R}, \mathbb{R})$  one has the identity*

$$\varphi(Z_t) = \varphi(Z_s) + \sum_{i=0}^{\lfloor N/2 \rfloor - 1} \alpha_i \int_s^t \varphi^{(2i+1)}(Z_r) d^{\circ(2i+1)} Z_r, \quad (4.4.17)$$

as long as there exists  $\mathcal{Q}_n$  a deterministic sequence of partition of  $[0, t]$  satisfying  $|\mathcal{Q}_n| \rightarrow 0$  such that only  $\lfloor N/2 \rfloor - 1$  integrals in the right hand side of (4.4.17) are well defined.

*Remark 4.4.9.* A similar type of formula was obtained in the proof of [GRV03, Thm 4.1] in the context of the fractional Brownian motion  $B$  with Hurst index  $H = 1/4$ . Indeed using the definition of the symmetric integral defined as in (4.4.5) the authors were able to provide for any  $\varphi \in C^4(\mathbb{R})$  the identity

$$\varphi(B_t) = \varphi(B_s) + \int_s^t \varphi'(B_r) d^{\circ(1)} B_r - \frac{1}{12} \int_s^t \varphi^{(3)}(B_r) d^{\circ(3)} B_r. \quad (4.4.18)$$

Even without entering in the technical details of the statement, since the process  $B$  admits a continuous modification which is almost  $1/4$ -Hölder (see [NZ98, Chap. 5]), the formula (4.4.17) for  $N = 4$  is formally identical to (4.4.18) and the Corollary 4.4.8 generalises such identity for any  $\gamma$ -Hölder process  $X$ . However we should not forget that this result is based upon the *trajectorial* properties of  $X$  and it does not include the specific *probabilistic law* of the process  $X$ . Indeed it turns out that in the specific case of  $B$ , the Gaussian nature of its process allows to prove in [GRV03, Thm 3.4] the additional property

$$\int_s^t \varphi^{(3)}(B_r) d^{\circ(3)} B_r = 0,$$

where the integral is again interpreted in the sense of (4.4.5). Thus the formula (4.4.18) simplifies and we obtain a Stratonovich type formula. By analogy we deduce that even the the generic formula (4.4.17) might simplify by choosing a specific process  $X$ .

## 4.4.2 A simplified proof of the Stratonovich formula

We apply the Corollary 4.4.8 to give a simplified proof of the Theorem 4.4.1 over a class of test functions with lower regularity. Indeed the sequence  $T_t^n(g')$  given by (4.4.1) is almost the approximating sequence of a symmetric integral defined in (4.4.4) when  $m = 1$  and with respect to a specific partition of  $[0, t]$  (the last point  $t$  is not included in the sum defining  $T_t^n(g')$ ). To show the result it is sufficient to recall only one technical lemma to obtain the Theorem 4.4.1. For its proof see [BS10, Thm 4.3].

**Lemma 4.4.10.** *Let  $g \in C^4(\mathbb{R}, \mathbb{R})$ . Choosing the sequence  $t_j = j/n$  we then have the following convergences*

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{j=1, j \text{ odd}}^{\lfloor nt \rfloor} g(X_{t_{j-1}}) (X_{t_j} - X_{t_{j-1}})^3 &= \lim_{n \rightarrow \infty} \sum_{j=1, j \text{ even}}^{\lfloor nt \rfloor} g(X_{t_{j-1}}) (X_{t_j} - X_{t_{j-1}})^3 \\ &= -\frac{3}{2\pi} \int_0^t g(X_s) ds, \end{aligned}$$

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{j=1, j \text{ odd}}^{\lfloor nt \rfloor} g(X_{t_j})(X_{t_j} - X_{t_{j-1}})^3 &= \lim_{n \rightarrow \infty} \sum_{j=1, j \text{ even}}^{\lfloor nt \rfloor} g(X_{t_j})(X_{t_j} - X_{t_{j-1}})^3 \\ &= \frac{3}{2\pi} \int_0^t g(X_s) ds, \end{aligned}$$

where the limit is in probability, uniformly on  $t \in [0, T]$ .

**Theorem 4.4.11.** For any  $\varphi \in C_b^5(\mathbb{R}, \mathbb{R})$  the trapezoid-type approximations sums  $T_t^n(\varphi)$  given in (4.4.1) converges in probability uniformly on  $t \in [0, T]$  to some continuous process  $\int_0^t \varphi'(X_r) d^\circ X_r$  such that we have the Stratonovich-type formula (4.4.2) and the a.s. identity

$$\int_0^t \Phi'(X_r) d\mathbf{X}_r^1 = \int_0^t \varphi'(X_r) d^\circ X_r. \quad (4.4.19)$$

*Proof.* Applying the Corollary 4.4.8 when  $N = 4$  to the process  $X$  if there exists  $\mathcal{Q}_n$  a deterministic sequence of partition of  $[0, t]$  satisfying  $|\mathcal{Q}_n| \rightarrow 0$  such that the symmetric integral

$$\int_s^t \varphi^{(3)}(X_r) d^{\circ(3)} X_r$$

is well defined, then the first order symmetric integral

$$\int_s^t \varphi^{(1)}(X_r) d^{\circ(1)} X_r$$

is well defined along the same sequence and we have the identities

$$\begin{aligned} \varphi(X_t) &= \varphi(X_s) + \int_s^t \varphi^{(1)}(X_r) d^{\circ(1)} X_r - \frac{1}{12} \int_s^t \varphi^{(3)}(X_r) d^{\circ(3)} X_r, \\ \int_0^t \Phi'(X_r) d\mathbf{X}_r^1 &= \int_s^t \varphi^{(1)}(X_r) d^{\circ(1)} X_r - \frac{1}{12} \int_s^t \varphi^{(3)}(X_r) d^{\circ(3)} X_r. \end{aligned}$$

Let us show that the sequence of processes

$$S_t^n := \sum_{t_i \in \mathcal{Q}_n} \frac{\varphi^{(3)}(X_{t_i}) + \varphi^{(3)}(X_{t_{i+1}})}{2} (X_{t_{i+1}} - X_{t_i})^3 \xrightarrow{\mathbb{P}} 0 \quad (4.4.20)$$

uniformly on  $t$  along the specific sequence of partitions

$$\mathcal{Q}_n := \left\{ \frac{k}{n} : k = 0, \dots, \lfloor nt \rfloor \right\} \cup \{t\}.$$

Using the notation

$$\begin{aligned} A_t^n &:= \frac{1}{2} \sum_{j=1}^{\lfloor nt \rfloor} \varphi^{(3)}(X_{t_{j-1}})(X_{t_j} - X_{t_{j-1}})^3, & B_t^n &:= \frac{1}{2} \sum_{j=1}^{\lfloor nt \rfloor} \varphi^{(3)}(X_{t_j})(X_{t_j} - X_{t_{j-1}})^3, \\ C_t^n &:= \frac{\varphi^{(3)}(X_{t_{\lfloor nt \rfloor}}) + \varphi^{(3)}(X_t)}{2} (X_t - X_{t_{\lfloor nt \rfloor}})^3, \end{aligned}$$

We can easily check that we have the a.s. identity  $S_t^n = A_t^n + B_t^n + C_t^n$ . Moreover by applying the Lemma 4.4.10 and the general estimate (4.3.4), we obtain also the following convergences uniformly on  $t \in [0, T]$

$$A_t^n \xrightarrow{\mathbb{P}} -\frac{3}{2\pi} \int_0^t g(X_s) ds, \quad \lim_{n \rightarrow \infty} B_t^n \xrightarrow{\mathbb{P}} \frac{3}{2\pi} \int_0^t g(X_s) ds, \quad C_t^n \xrightarrow{\mathbb{P}} 0,$$

thereby obtaining the that the first order symmetric integral satisfies

$$\varphi(X_t) = \varphi(X_s) + \int_s^t \varphi^{(1)}(X_r) d^{\circ(1)} X_r, \quad \int_0^t \Phi'(X_r) d\mathbf{X}_r^1 = \int_s^t \varphi^{(1)}(X_r) d^{\circ(1)} X_r. \quad (4.4.21)$$

Recalling the definition of the sequence  $T_t^n(\varphi')$  in (4.4.1), by hypothesis on  $\varphi$  there exists a constant  $M$  such that

$$\left| T_t^n - \sum_{t_i \in \mathcal{Q}_n} \frac{\varphi'(X_{t_i}) + \varphi'(X_{t_{i+1}})}{2} (X_{t_{i+1}} - X_{t_i}) \right| \leq M |X_t - X_{t_{\lfloor nt \rfloor}}|.$$

Using the estimate (4.3.2) and the property  $t_{\lfloor nt \rfloor} \rightarrow t$ , we deduce that  $T_t^n$  converges in probability uniformly on  $t \in [0, T]$  and its limit coincides with  $\int_0^t g'(X_r) d^{\circ(1)} X_r$  and using the equality (4.4.21) we conclude.  $\square$

*Remark 4.4.12.* Looking at the formula (4.4.2), we simply observe that the procedure to prove it is the same as we briefly recalled in the Remark 4.4.9 for the fractional Brownian motion with Hurst parameter equal to  $1/4$ . Considering the identity (4.4.19) we recall that this identity holds only when the converging sums is chosen along some specific converging sequence of partitions of  $[0, t]$ , whereas the rough integral does not depends on that.

# Bibliography

- [Abe04] Eiichi Abe. *Hopf Algebras*. Cambridge Tracts in Mathematics. Cambridge University Press, 2004.
- [BCCH17] Y. Bruned, A. Chandra, I. Chevyrev, and M. Hairer. Renormalising SPDEs in regularity structures. *arXiv e-prints*, 2017.
- [BCD11] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011.
- [BCEF18] Yvain Bruned, Charles Curry, and Kurusch Ebrahimi-Fard. Quasi-shuffle algebras and renormalisation of rough differential equations. *arXiv preprint 1801.02964*, 2018.
- [Bel18] Carlo Bellingeri. An Itô type formula for the additive stochastic heat equation. *arXiv e-prints :1803.01744*, 2018.
- [Bel19a] Carlo Bellingeri. Rough change of variable formulae and quasi-geometric rough paths. *In preparation*, 2019.
- [Bel19b] Carlo Bellingeri. Rough Itô formulae for the stochastic heat equation. *In preparation*, 2019.
- [BFG<sup>+</sup>17] C. Bayer, P. K. Friz, P. Gassiat, J. Martin, and B. Stemper. A regularity structure for rough volatility. *ArXiv e-prints*, 2017.
- [BHZ19] Y. Bruned, M. Hairer, and L. Zambotti. Algebraic renormalisation of regularity structures. *Inventiones mathematicae*, 215(3):1039–1156, 2019.
- [Boe18] Horatio Boedihardjo. Decay rate of iterated integrals of branched rough paths. *Annales de l’Institut Henri Poincaré. Analyse Non Linéaire*, 35(4):945–969, 2018.
- [Bru15] Yvain Bruned. *Singular KPZ Type Equations*. Theses, Université Pierre et Marie Curie - Paris VI, 2015.
- [Bru18] Yvain Bruned. Recursive formulae in regularity structures. *Stochastic Partial Differential Equations. Analysis and Computations*, 6(4):525–564, 2018.
- [BS10] Krzysztof Burdzy and Jason Swanson. A change of variable formula with itô correction term. *Annals of Probability*, 38(5):1817–1869, 2010.



- [But72] John Charles Butcher. An algebraic theory of integration methods. *Math. Comp.*, 26:79–106, 1972.
- [Can84] John R. Cannon. *The one-dimensional heat equation*, volume 23 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1984.
- [CEFMMK18] C. Curry, K. Ebrahimi-Fard, D. Manchon, and H. Munthe-Kaas. Planarly branched rough paths and rough differential equations on homogeneous spaces. *arXiv preprint 1804.08515v3*, 2018.
- [CH16] A. Chandra and M. Hairer. An analytic BPHZ theorem for regularity structures. *ArXiv e-prints*, 2016.
- [CK98] Alain Connes and Dirk Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Communications in Mathematical Physics*, 199(1):203–242, Dec 1998.
- [CN05] Patrick Cheridito and David Nualart. Stochastic integral of divergence type with respect to fractional Brownian motion with Hurst parameter  $H \in (0, 1/2)$ . *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 41(6):1049 – 1081, 2005.
- [Col84] Jean-François Colombeau. *New generalized functions and multiplication of distributions*, volume 84 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1984. Notas de Matemática [Mathematical Notes], 90.
- [CP19] Rama Cont and Nicolas Perkowski. Pathwise integration and change of variable formulas for continuous paths with arbitrary regularity. *Transactions of the American Mathematical Society. Series B*, 6:161–186, 2019.
- [CR99] Rene A. Carmona and Boris Rozovskii, editors. *Stochastic partial differential equations: six perspectives*, volume 64 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1999.
- [CW75] R. Cairoli and John B. Walsh. Stochastic integrals in the plane. *Acta Mathematica*, 134:111–183, 1975.
- [DP99] Giuseppe Da Prato. On perturbations of symmetric gaussian diffusions. *Stochastic Analysis and Applications*, 17(3):369–381, 1999.
- [DPZ14] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 2014.
- [EFMPW15] Kurusch Ebrahimi-Fard, Simon J.A. Malham, Frédéric Patras, and Anke Wiese. Flows and stochastic Taylor series in Itô calculus. *Journal of Physics A: Mathematical and Theoretical*, 48(49):495202, 2015.
- [EL15] Alison M. Etheridge and Cyril Labbé. Scaling limits of weakly asymmetric interfaces. *Communications in Mathematical Physics*, 336(1):287–336, 2015.

- [ERV02] Mohammed Errami, Francesco Russo, and Pierre Vallois. Itô's formula for  $C^1$   $\lambda$ -functions of a càdlàg process and related calculus. *Probability Theory and Related Fields*, 122(2):191, February 2002.
- [EZ18] Henri Elad Altman and Lorenzo Zambotti. Bessel SPDEs and renormalized local times. *arXiv e-prints :1811.00518*, 2018.
- [FGGR16] P. K. Friz, B. Gess, A. Gulisashvili, and S. Riedel. The Jain–Monrad criterion for rough paths and applications to random Fourier series and non-Markovian Hörmander theory. *The Annals of Probability*, 44(1):684–738, 2016.
- [FH14] Peter Friz and Martin Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. Universitext. Springer International Publishing, 2014.
- [FM17] Frédéric Fauvet and Frederic Menous. Ecalle's arborification coarborification transforms and Connes-Kreimer Hopf algebra. *Annales scientifiques de l'ENS*, 50(1):39 – 83, 2017.
- [Föll81] Hans Föllmer. Calcul d'Ito sans probabilités. *Séminaire de probabilités de Strasbourg*, 15:143–150, 1981.
- [FOT11] Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011.
- [FR02] Franco Flandoli and Francesco Russo. Generalized Integration and Stochastic ODEs. *Annals of Probability*, 30(1):270–292, 2002.
- [Fun83] Tadahisa Funaki. Random motion of strings and related stochastic evolution equations. *Nagoya Mathematical Journal*, 89:129–193, 1983.
- [FV10a] Peter Friz and Nicolas Victoir. Differential equations driven by Gaussian signals. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(2):369–413, 2010.
- [FV10b] Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Theory and applications.
- [GH80] Donald Geman and Joseph Horowitz. Occupation Densities. *Annals of Probability*, 8(1):1–67, 02 1980.
- [GH17] M. Gerencsér and M. Hairer. Singular SPDEs in domains with boundaries. *ArXiv e-prints*, 2017.
- [GIP15] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi*, 3, 2015.
- [GNRV05] Mihai Gradinaru, Ivan Nourdin, Francesco Russo, and Pierre Vallois. m-order integrals and generalized Itô's formula; the case of a fractional

- Brownian motion with any Hurst index. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 41(4):781–806, 2005.
- [GNT05] Mihai Gradinaru, Ivan Nourdin, and Samy Tindel. Ito's- and Tanaka's-type formulae for the stochastic heat equation: The linear case. *Journal of Functional Analysis*, 228:114–143, 2005.
- [GRV03] Mihai Gradinaru, Francesco Russo, and Pierre Vallois. Generalized covariations, local time and Stratonovich Itô's formula for fractional Brownian motion with Hurst index  $H \geq \frac{1}{4}$ . *The Annals of Probability*, 31(4):1772–1820, 2003.
- [Gub04] M. Gubinelli. Controlling rough paths. *Journal of Functional Analysis*, 216(1):86–140, 2004.
- [Gub10] Massimiliano Gubinelli. Ramification of rough paths. *Journal of Differential Equations*, 248(4):693 – 721, 2010.
- [Gyu16] Lajos Gergely Gyurkó. Differential equations driven by  $\pi$ -rough paths. *Proceedings of the Edinburgh Mathematical Society*, 59(3):741–758, 2016.
- [Hai11] M. Hairer. Rough stochastic PDEs. *Communications on Pure and Applied Mathematics.*, 64(11):1547–1585, 2011.
- [Hai13] Martin Hairer. Solving the KPZ equation. *Annals of Mathematics*, 178(2):559–664, Jul 2013.
- [Hai14] M. Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2):269–504, 2014.
- [Hai15] Martin Hairer. Introduction to regularity structures. *Brazilian Journal of Probability and Statistics*, 29(2):175–210, 2015.
- [Hai16] Martin Hairer. The motion of a random string. *ArXiv e-prints*, 2016.
- [HI17] Michael E. Hoffman and Kentaro Ihara. Quasi-shuffle products revisited. *Journal of Algebra*, 481:293 – 326, 2017.
- [HK15] Martin Hairer and David Kelly. Geometric versus non-geometric rough paths. *Ann. Inst. H. Poincaré Probabilité et Statistique*, 51(1):207–251, 02 2015.
- [Hof00] Michael E. Hoffman. Quasi-Shuffle Products. *Journal of Algebraic Combinatorics*, 11(1):49–68, Jan 2000.
- [HP15] Martin Hairer and Étienne Pardoux. A Wong-Zakai theorem for stochastic PDEs. *Journal of the Mathematical Society of Japan*, 67(4):1551–1604, 2015.
- [HW13] Martin Hairer and Hendrik Weber. Rough Burgers-like equations with multiplicative noise. *Probability Theory and Related Fields*, 155(1-2):71–126, 2013.
- [Itô42] Kiyosi Itô. Differential equations determining Markov processes. *Zenkoku Shijo Sugaku Danwakai*, 1077:1352–1400, 1942.

- [Itô44] Kiyosi Itô. Stochastic integral. *Proceedings of the Imperial Academy of Tokyo*, 20:519–524, 1944.
- [Kel12] David Kelly. *Itô corrections in stochastic equations*. PhD thesis, University of Warwick, 2012.
- [KPZ86] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. Dynamic Scaling of Growing Interfaces. *Physical Review Letters*, 56:889–892, Mar 1986.
- [Kry96] N.V. Krylov. *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*. Graduate studies in mathematics. American Mathematical Society, 1996.
- [Lan07] Alberto Lanconelli. On a new version of the Itô’s formula for the stochastic heat equation. *Communications on Stochastic Analysis*, 1(2):311–320, 2007.
- [Lot97] M. Lothaire. *Combinatorics on words*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997.
- [LV07] Terry Lyons and Nicolas Victoir. An extension theorem to rough paths. *Annales de l’Institut Henri Poincaré. Analyse Non Linéaire*, 24(5):835–847, 2007.
- [Lyn54] R. C. Lyndon. On burnside’s problem. *Transactions of the American Mathematical Society*, 77(2):202–215, 1954.
- [Lyo98] Terry J. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310, 1998.
- [Man04] Dominique Manchon. Hopf algebras, from basics to applications to renormalization. *arXiv: 0408405*, 2004.
- [MR89] Guy Melançon and Christophe Reutenauer. Lyndon words, free algebras and shuffles. *Canadian Journal of Mathematics. Journal Canadien de Mathématiques*, 41(4):577–591, 1989.
- [MT02] C. Mueller and R. Tribe. Hitting properties of a random string. *Electronic Journal of Probability*, 7:no. 10, 29, 2002.
- [NN10] Ivan Nourdin and David Nualart. Central Limit Theorems for Multiple Skorokhod integrals. *Journal of Theoretical Probability*, 23(1):39–64, Mar 2010.
- [Nua95] David Nualart. *The Malliavin Calculus and Related Topics*. Probability and its applications. Springer-Verlag, 1995.
- [NZ98] David Nualart and Moshe Zakai. Generalized multiple stochastic integrals and the representation of Wiener functionals. *Stochastics*, 23(3):311–330, 1998.
- [PT11] Giovanni Peccati and Murad Taqqu. *Wiener Chaos: Moments, Cumulants and Diagrams*. Springer, 2011.
- [Reu93] Christophe Reutenauer. *Free Lie Algebras*. LMS monographs. Clarendon Press, 1993.

- [RT06] Francesco Russo and Ciprian A. Tudor. On bifractional Brownian motion. *Stochastic Processes and their Applications*, 116(5):830–856, 2006.
- [RV93] Francesco Russo and Pierre Vallois. Forward, Backward and Symmetric Stochastic integration. *Probability Theory and Related Fields*, 97(3):403–421, 1993.
- [RV95] Francesco Russo and Pierre Vallois. The generalized covariation process and Itô formula. *Stochastic Processes and their applications*, 59(1):81–104, 1995.
- [RY04] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2004.
- [Sch54] Laurent Schwartz. Sur l'impossibilité de la multiplication des distributions. *Comptes Rendus de l'Académie des sciences de Paris*, 239:847–848, 1954.
- [Sim97] Leon Simon. Schauder estimates by scaling. *Calculus of Variations and Partial Differential Equations*, 5(5):391–407, Jul 1997.
- [Str66] Ruslan Stratonovich. A new representation for stochastic integrals and equations. *SIAM Journal on Control*, 4(2):362–371, 1966.
- [Swa07] Jason Swanson. Variations of the solution to a stochastic heat equation. *Annals of Probability*, 35(6):2122–2159, 2007.
- [TZ18] Nikolas Tapia and Lorenzo Zambotti. The geometry of the space of branched Rough Paths. *arXiv preprint 1810.12179v3*, 2018.
- [Wal84] John B. Walsh. An introduction to stochastic partial differential equations. *École d'été de probabilités de Saint-Flour, XIV - 1984*, 1984.
- [You36] William H. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Mathematica*, 67(1):251–282, 1936.
- [Zam01] Lorenzo Zambotti. A reflected stochastic heat equation as symmetric dynamics with respect to the 3-d Bessel bridge. *Journal of Functional Analysis*, 180(1):195–209, 2001.
- [Zam06] Lorenzo Zambotti. Itô-tanaka's formula for stochastic partial differential equations driven by additive space-time white noise. *Stochastic partial differential equations and applications - VII*, 245:337–347, 2006.
- [Zam17] Lorenzo Zambotti. Random Obstacle Problems. *École d'été de probabilités de Saint-Flour, XLV - 2015*, 2017.