Robust Stochastic Analysis with Applications

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

im Fach Mathematik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Dipl.-Math. David Johannes Prömel

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät
Prof. Dr. Elmar Kulke

Gutachter:
1. Prof. Dr. Peter Imkeller
2. Prof. Dr. Peter Friz
3. Prof. Dr. Josef Teichmann

Eingereicht am: 28.07.2015
Tag der Verteidigung: 25.09.2015

Abstract

In this thesis new robust integration techniques, which are suitable for various problems from stochastic analysis and mathematical finance, as well as some applications are presented.

We begin with two different approaches to stochastic integration in robust financial mathematics. The first one is inspired by Itô's integration and based on a certain topology induced by an outer measure corresponding to a minimal superhedging price. The second approach relies on the controlled rough path integral. We prove that this integral is the limit of non-anticipating Riemann sums and that every "typical price path" has an associated Itô rough path. For one-dimensional "typical price paths" it is further shown that they possess Hölder continuous local times. Additionally, we provide various generalizations of Föllmer's pathwise Itô formula.

Recalling that rough path theory can be developed using the concept of controlled paths and with a topology including the information of Lévy's area, sufficient conditions for the pathwise existence of Lévy's area are provided in terms of being controlled. This leads us to study Föllmer's pathwise Itô formulas from the perspective of controlled paths.

A multi-parameter extension to rough path theory is the paracontrolled distribution approach, recently introduced by Gubinelli, Imkeller and Perkowski in [GIP12]. We generalize their approach from Hölder spaces to Besov spaces to solve rough differential equations. As an application we deal with stochastic differential equations driven by random functions.

Finally, considering strongly coupled systems of forward and backward stochastic differential equations (FBSDEs), we extend the existence, uniqueness and regularity theory of so-called decoupling fields to Markovian FBSDEs with locally Lipschitz continuous coefficients. These results allow to solve the Skorokhod embedding problem for a class of Gaussian processes with non-linear drift.

Zusammenfassung

Diese Dissertation präsentiert neue Techniken der Integration für verschiedene Probleme der Finanzmathematik und einige Anwendungen in der Wahrscheinlichkeitstheorie.

Zu Beginn entwickeln wir zwei Zugänge zur robusten stochastischen Integration. Der erste, ähnlich der Itô'schen Integration, basiert auf einer Topologie, welche erzeugt wird von einem äußeren Maß, gegeben durch einen minimalen Superreplikationspreis. Der zweite gründet auf der Integrationtheorie für rauhe Pfade. Wir zeigen, dass das entsprechende Integral als Grenzwert von nicht antizipierenden Riemannsummen existiert und dass sich jedem "typischen Preispfad" ein rauher Pfad im Itô'schen Sinne zuordnen lässt. Für eindimensionale "typische Preispfade" wird sogar gezeigt, dass sie Hölder-stetige Lokalzeiten besitzen. Zudem erhalten wir verschiedene Verallgemeinerungen von Föllmer's pfadweiser Itô-Formel.

Die Integrationstheorie für rauhe Pfade kann mit dem Konzept der kontrollierten Pfade und einer Topologie, welche die Information der Lévy-Fläche enthält, entwickelt werden. Deshalb untersuchen wir hinreichende Bedingungen an die Kontrollstruktur für die Existenz der Lévy-Fläche. Dies führt uns zur Untersuchung von Föllmer's pfadweiser Itô-Formel aus der Sicht kontrollierter Pfade.

Para-kontrollierte Distributionen, kürzlich von Gubinelli, Imkeller und Perkowski [GIP12] eingeführt, erweitern die Theorie rauher Pfade auf den Bereich von mehrdimensionale Parameter. Wir verallgemeinern diesen Ansatz von Hölder'schen auf Besov-Räume, um rauhe Differentialgleichungen zu lösen, und wenden die Ergebnisse auf stochastische Differentialgleichungen an.

Zum Schluß betrachten wir stark gekoppelte Systeme von stochastischen Vorwärts-Rückwärts-Differentialgleichungen (FBSDEs) und erweitern die Theorie der Existenz, Eindeutigkeit und Regularität der sogenannten Entkopplungsfelder auf Markovsche FBSDEs mit lokal Lipschitz-stetigen Koeffizienten. Als Anwendung wird das Skorokhodsche Einbettungsproblem für Gaußsche Prozesse mit nichtlinearem Drift gelöst.

Acknowledgement

Above all, I would like to express my deepest gratitude to Peter Imkeller for his excellent mentoring and the supervision he provided me with while I was writing this thesis. He gave me the freedom, encouragement and guidance to pursue my own research interests. The uncountably many inspiring discussions with him and his stimulating questions have broadened my mathematical horizon and have promoted this thesis enormously.

I would like to thank Peter Friz and Josef Teichmann for agreeing to be co-examiners for this thesis.

I am deeply grateful to Nicolas Perkowski for encouraging me throughout of my Ph.D. studies. His enthusiasm to do mathematics and his patience to explain mathematical ideas gave me a significant push forward. He also pointed my interest towards Vovk's game-theoretic approach to mathematical finance, which led to the first two chapters of this thesis.

I am also greatly indebted to Josef Teichmann for giving me the opportunity to stay at the Department of Mathematics, Eidgenössische Technische Hochschule Zürich. In fruitful and pleasant discussions he gave me a lot of new insights into stochastic analysis and taught me new perspectives in financial mathematics. He encouraged me to grasp the concept of Hairer's theory of regularity structure. Furthermore, I would like to thank all my friends and colleagues in Zürich, who made my time there so enjoyable, especially my colleagues from the "lake-view-office": Leif Döring, Thomas Krabichler, Ariel Neufeld, Oleg Reichmann and Sara Svaluto-Ferro.

The instructive and fruitful collaborations with my co-authors Alexander Fromm and Mathias Trabs have inspired me and I appreciated them immensely. It was really a great pleasure to collaborate with Mathias, who was willing to work in mathematical areas completely novel for both of us. I have profited a lot from the excellent research environment at Humboldt-Universität zu Berlin. Mathematical discussions as well as coffee breaks would not have been so much fun without Randolf Altmeyer, Oliver Janke and Mathias Trabs. Thanks to all of you.

The financial support by a Ph.D. scholarship of the Research Training Group 1845 "Stochastic Analysis with Application in Biology, Finance and Physics" as well as the great infrastructure for Ph.D. students provided by the Berlin Mathematical School is gratefully acknowledged.

Last but not least, I would like to thank my friends and even more my whole family and Larissa. Without their unconditional love and personal support this thesis would not have come into being.

Contents

1. Introduction

One of the central topics in probability theory is stochastic integration with its numerous applications in stochastic analysis and mathematical finance. Let us begin by illustrating the importance of stochastic integration by two fundamental problems.

In financial mathematics a basic problem is to find “reasonable” prices and hedging strategies for financial derivatives. The first approach to this problem goes back to Bachelier [Bac00]. Seventy years later the works by Merton [Mer73] and Black and Scholes [BS73] revolutionized the mathematical finance. The idea can be described as follows: Assuming the price evolution S of a stock is given by a Brownian motion or a geometric Brownian motion, and we want to price, for instance, a European call option $(S_T - K)^+$. In words this option allows, but not obliges, the holder to buy one stock corresponding to S at time T for the strike price K. By the predictable representation property of the price process S, there exist a constant p and a predictable process H such that

$$(S_T - K)^+ = p + \int_0^T H_s \,\mathrm{d}S_s.$$

Therefore, an investor endowed with an initial capital p can choose the trading strategy H to obtain the same payoff as the option $(S_T - K)^+$. Consequently, in an arbitrage free and frictionless market the “reasonable” price of the European call option $(S_T - K)^+$ should be p.

More generally, in classical financial mathematics, where the price process S is often presumed to be a semimartingale, one can rely on Itô’s stochastic integration to give the appearing capital process $\int H \,\mathrm{d}S$ a rigorous meaning. This requires to postulate a fixed underling probability space together with a probabilistic description of the market dynamics. Unfortunately, this approach fails to include the model risk, also called Knightian uncertainty (cf. Knight [Kni21]). The first works including model risk and dealing with pricing and hedging under volatility uncertainty were authored by Lyons [Lyo95b] and Avellaneda et al. [ALP95]. Instead of assuming a price dynamic under one probability measure $\mathbb{P}$, they consider the price process S simultaneously under a family $\{\mathbb{P}_\alpha\}_{\alpha \in I}$ of probability measures. In this case a suitable stochastic integration theory can be developed using quasi-sure analysis based on capacity theory as done by Denis and Martini [DM06], or using aggregation methods as by Soner et al. [STZ11]. More recently, starting with the pioneering work of Hobson [Hob98], it becomes more and more popular to price and hedge options completely without an underlying model or without presuming any reference measure. In the model-independent context, to develop an appropriate integration theory is an even more challenging and widely open problem.

Our second application of stochastic integration is linked with the area of controlled differential equations, which are omnipresent in stochastic analysis. They form a very important subclass of classical ordinary differential equations gaining extra interest from their various fields of application. The dynamic of such a controlled differential equation is described by

$$\dot{u}(t) = F(u(t))\dot{\vartheta}(t), \quad u(0) = u_0, \quad t \in [0,T], \tag{1.1}$$

where $u_0 \in \mathbb{R}^m$ is the initial condition, F is a suitable vector field and $\vartheta\colon [0,T] \to \mathbb{R}^n$ is a deterministic smooth function, for $m,n \in \mathbb{N}$. In probability theory, $\dot{\vartheta}$ is often replaced by a stochastic driving signal, for example white noise, given formally as derivative of Brownian motion B. It is well-known that Brownian motion is α-Hölder continuous for any $\alpha < 1/2$ and nowhere differentiable almost surely. This makes it impossible to give directly a rigorous meaning to the product $F(u)\dot{\vartheta}$. One approach to overcome this problem is to formally integrate equation (1.1), which gives

$$u(t) = u_0 + \int_0^t F(u(s))\,\mathrm{d}\vartheta(s), \quad t \in [0,T].$$

In this way the problem of defining the product $F(u)\dot{\vartheta}$ translates into understanding the integral $\int_0^t F(u(s))\,\mathrm{d}\vartheta(s)$. Roughly speaking, there are two different approaches to construct this integral. The first strategy relies on the probabilistic nature of the involved process ϑ. This leads to stochastic integration theory as, for instance, Itô or Stratonovich integration. The second one ignores the stochastic structure, but assumes additional information about ϑ in order to build up a deterministic integration theory rich enough to handle paths with the regularity of Brownian motion trajectories. To the later strategy we refer as *pathwise approach* or *pathwise stochastic integration* since it will be used for problems coming from probability theory in the present thesis.

Let us briefly display the most common concepts of integration suitable for stochastic analysis and financial mathematics.

Riemann and Young integration

The most classical attempt to define an integral of two functions $X\colon [0,T] \to \mathbb{R}^n$ and $Y\colon [0,T] \to \mathbb{R}^n$ is to start with left-point Riemann sums and to set

$$\int_0^T Y_s\,\mathrm{d}X_s := \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} Y_s(X_t - X_s), \tag{1.2}$$

where π belongs to the collection of all partitions of $[0,T]$ and $|\pi|$ denotes the mesh size of π. Especially, in view of the above mentioned application to finance, where the integral is meant to be the capital process of a hedging strategy Y investing on a market with price dynamics X, the left-point Riemann sums are the canonical choice. Assuming that the path Y is continuous and X is of finite variation, i.e.

$$\lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} |X_t - X_s| < \infty,$$

it is commonly known that the limit (1.2) exists along every sequence of partitions with mesh size going to zero. Already this basic construction has its applications in recent model-independent financial mathematics. We refer, for instance, to Dolinsky and Soner [DS14].

A more elaborated approach was developed by Young [You36] relying on the concept of p-variation. The p-variation of a continuous path X is given by

$$\sup_{\pi}\bigg(\sum_{[s,t]\in\pi}|X_t-X_s|^p\bigg)^{1/p},\quad p\geq 1,$$

where the supremum is taken over the collection of all partitions of $[0,T]$. Provided X and Y are of finite p- and q-variation, respectively, with $1/p+1/q>1$, Young proved that the limit of Riemann sums in (1.2) exists independently of the chosen sequence of partitions, and that the integral operator $(Y,X)\mapsto\int Y\,\mathrm{d}X$ is continuous with respect to the q- and p-variation norm. To point out the necessity of the assumption $1/p+1/q>1$, Young also constructed an example of two paths of finite 2-variation for which the Riemann sums in (1.2) diverge. Young integration already allows for treating controlled differential equations (1.1) under the assumption that ϑ has finite p-variation for some $p<2$. This was proven for the first time by Lyons [Lyo94] using a Picard iteration. Although this result covers interesting examples from stochastic analysis such as fractional Brownian motion with Hurst index $H>1/2$, it excludes frequently appearing stochastic processes like Brownian motion and continuous martingales.

Itô integration

The most frequently used notion of integration in probability theory is the so called Itô integration initiated by Itô [Itô44]. Like ordinary integrals, stochastic Itô integrals are constructed by a limiting procedure. To briefly sketch this construction, we suppose that X in (1.2) is replaced by a Brownian motion B, which generates a filtration $(\mathcal{F}_t)$. Let us denote the space of simple integrands by $\mathcal{E}$ consisting of all stochastic processes which are piecewise constant, left-continuous and adapted to the Brownian filtration $(\mathcal{F}_t)$. Adaptedness is a probabilistic concept, which does not appear in the other pathwise approaches to integration, but is crucial for constructing the Itô integral. Heuristically, it says that the integrand Y at time t does not have more "information" about the Brownian motion B than is available at time t. If Y is a simple integrand, the integral $\int Y_s\,\mathrm{d}B_s$ is well-defined as Riemann sum and the integral process is a martingale. Hence, a fairly elementary calculation reveals the Itô isometry:

$$\mathbb{E}\Big[\int_0^T Y_t^2\,\mathrm{d}t\Big]=\mathbb{E}\Big[\Big(\int_0^T Y_t\,\mathrm{d}B_t\Big)^2\Big].$$

Therefore, one sees that the integral map $I\colon\mathcal{E}\to L^2(\mathbb{P})$ defined by $Y\mapsto\int Y\,\mathrm{d}B$ is a linear isometry, which can be uniquely extended from $\mathcal{E}$ to the space $L^2(\mathrm{d}\mathbb{P}\otimes\mathrm{d}t)$. This extension is then called *Itô integral*.

With this notion of stochastic integration Itô [Itô51] was able to define and solve the differential equation (1.1) driven by Brownian motion. Subsequently, the Itô integral was extended to stochastic integration with respect to martingales (Kunita and Watanabe [KW67]), local martingales (Meyer [Mey67], Doleans-Dade and Meyer [DDM69]) and semimartingales (Jacod [Jac79], Dellacherie and Meyer [DM82]).

Föllmer integration

In his seminal paper Föllmer [Fö81] developed the first deterministic approach which allows for defining pathwise integrals with respect to Brownian motion or continuous martingales. His starting point was the hypothesis that the quadratic variation of the continuous path X exists along a sequence of partitions (π_n) whose mesh size tends to zero. This is almost surely the case, for instance, for the just mentioned stochastic processes. According to his concept, a continuous function $X\colon [0,T] \to \mathbb{R}$ has *quadratic variation* if the sequence of discrete measures on $([0,T], \mathcal{B}([0,T]))$ given by

$$\mu_n := \sum_{[s,t]\in\pi_n} |X_t - X_s|^2 \delta_s$$

converges weakly to a measure μ along an increasing sequence of partitions (π_n) such that $\lim_{n\to\infty} |\pi_n| = 0$, where δ_s denotes the Dirac measure at $s \in [0,T]$. It turns out that this concept is actually sufficient to construct certain stochastic integrals in a pathwise manner. More precisely, provided F is a twice continuously differentiable function and X has quadratic variation along (π_n), Föllmer presents a pathwise version of Itô's formula

$$F(X_T) = F(X_0) + \int_0^T \mathrm{D}F(X_t)\,\mathrm{d}^{\pi_n} X_t + \int_0^T \mathrm{D}^2 F(X_t)\,\mathrm{d}\langle X\rangle_t.$$

The appearing "stochastic" integral is given by the limit of Riemann sums

$$\int_0^T \mathrm{D}F(X_t)\,\mathrm{d}^{\pi_n} X_t := \lim_{n\to\infty} \sum_{[s,t]\in\pi_n} \mathrm{D}F(X_s)(X_t - X_s), \tag{1.3}$$

which has to converge by the assumption on quadratic variation. This special kind of integration is today named *Föllmer integration*.

Let us stress that this construction of integrals comes with a clear financial interpretation thanks to its approximation by left-point Riemann sums. Unsurprisingly, Föllmer integration has applications in finance. In particular, it is recently used in model-independent financial mathematics to derive price bounds of certain financial options, see for example Lyons [Lyo95b] and Davis et al. [DOR14].

Rough path integration

The theory of rough paths has established an analytical frame which allows for treating stochastic differential and integral calculus beyond Young's classical notions. It simultaneously extends the Riemann-Stieltjes, the Young and Föllmer integrals. Lyons [Lyo98] provided a systematic approach to handle pathwise integrals of the form $\int f(X)\,\mathrm{d}X$ if X is of finite p-variation for some $p \geq 1$. The main purpose of his seminal work [Lyo98] was to analyze controlled differential equations (1.1) driven by such irregular paths. His significant insight was to enhance the path X with its iterated integrals in an abstract setting in order to define $\int f(X)\,\mathrm{d}X$ as a linear and continuous map. As shown by Young, the iterated integrals cannot be defined in general as limit of Riemann sums, but they are supposed to be objects which mimic the iterated integrals in an algebraic and analytic way. In his well-known Extension Theorem [Lyo98, Theorem 2.2.1] Lyons proves that the number of iterated integrals required to define $\int f(X)\,\mathrm{d}X$ depends on the regularity of the path X. More precisely, if X is a path of finite p-variation with existing iterated integrals up to order p, then the iterated integrals of higher order are uniquely determined. In parallel to the p-variation setting, rough path theory can be developed in the Hölder topology with similar tools, cf. [FH14].

In this thesis we shall focus on paths of finite p-variation for $p \in (2,3)$, which require only the existence of the first iterated integral. A p-rough path $(X, \mathbb{X})$ for $p \in (2,3)$ is a pair of $X\colon [0,T] \to \mathbb{R}^n$ and $\mathbb{X}\colon [0,T]^2 \to \mathbb{R}^{n\times n}$ such that X has finite p-variation, $\mathbb{X}$ has finite $p/2$-variation and $(X, \mathbb{X})$ satisfies Chen's relation, i.e.

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = X_{s,u} \otimes X_{u,t}, \quad \text{for } 0 \le s \le u \le t \le T.$$

In that case $\mathbb{X}$ is sometimes referred to as the *area* of X. For such a rough path $(X, \mathbb{X})$ and a twice continuously differentiable function F the rough path integral is defined by

$$\int_0^T F(X_s)\,\mathrm{d}X_s := \lim_{|\mathcal{P}|\to 0} \sum_{[s,t]\in\mathcal{P}} \big(F(X_s)(X_t - X_s) + F'(X_s)\mathbb{X}_{s,t}\big). \tag{1.4}$$

A significant refinement of the rough path integral was due to Gubinelli [Gub04]: Ensuring still the existence of the rough path integral (1.4), the integrand $F(X_s)$ can be replaced by any path Y which is *controlled* by X. To be more exact, $Y\colon [0,T] \to \mathbb{R}^n$ is controlled by X if there is a process Y' of finite q-variation such that the remainder $R\colon [0,T]^2 \to \mathbb{R}^n$ implicitly given by

$$Y_t - Y_s = Y'_s(X_t - X_s) + R_{s,t}$$

is of finite r-variation with $1/r = 1/q + 1/q$ and $2/p + 1/q > 1$.

As we have seen in the various approaches to integration, the information structure, usually modeled by a filtration, plays a crucial role in Itô concept while it was completely ignored in the pathwise approaches. For example changing the initial condition to a terminal one in (1.1) makes a massive difference in the dynamics of the stochastic differential equation, dependently of whether one wants to be consistent with the information flow of the driving signal or not. Stochastic differential equations with a terminal condition are called *backward stochastic differential equations* (BSDEs) and have been introduced in the linear case by Bismut [Bis73, Bis78] as adjoint process in the stochastic version of the Pontryagin maximum principle. Almost 20 years later Pardoux and Peng [PP90] were the first to consider general BSDEs and to solve the question of existence and uniqueness. Many phenomena in stochastic analysis and financial mathematics can be described by or in fact require to solve more general systems of forward and backward stochastic differential equations (FBSDEs), whose dynamics can be stated by

$$\begin{aligned} X_s &= X_0 + \int_0^s \mu(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_0^s \sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r, \\ Y_t &= \xi(X_T) - \int_t^T f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \quad s,t \in [0,T], \end{aligned}$$

where μ, σ, f and ξ are supposed to be suitable coefficient functions. The theory of FBSDEs is closely connected to the theory of quasi-linear partial differential equations. It received strong interest from its numerous areas of applications among which stochastic control and mathematical finance are the most vivid ones in recent decades. See [EPQ97] or [PW99] for surveys.

Such a system of equations can be even linked to an old and central problem in probability theory: the Skorokhod embedding problem. This problem was originally formulated by Skorokhod [Sko61, Sko65]. Given a Brownian motion and a probability measure, it consists in finding a stopping time such that the stopped Brownian motion has the law given by the prescribed probability measure. Over 50 years this problem has received a lot of attention in probability theory, see [Obł04] for a survey. It recently gained additional interest because of its applications in robust finance starting with the seminal work of [Hob98].

To summarize the previous discussion about stochastic integration and its applications, this thesis contributes mainly to two general tasks from financial mathematics and stochastic analysis:

(i) Develop a stochastic (pathwise) integration concepts suitable for the requirements of model-independent financial mathematics.

(ii) Use a (robust) theory of stochastic integration to solve (pathwise forward) stochastic differential equations and system of forward-backward stochastic differential equations.

Roughly speaking, the first part (Chapters 2, 3 and 4) of this thesis is mainly related and motivated by the first task, while the second part (Chapters 4, 5 and 6) is concerned with the second one. Every Chapter is relatively self-contained and can be read independently. In the following we give a brief outline of each Chapter and sketch its main contributions.

Chapter 2: Pathwise stochastic integrals for model free finance

This chapter is based on [PP13] by Perkowski and Prömel. It uses Vovk's [Vov12] game-theoretic approach to develop two different techniques of stochastic integration in frictionless and model free financial mathematics. As discussed above in the application of stochastic integration in financial mathematics, integration is highly non-trivial in the model free context since we do not want to assume any probabilistic or semimartingale structure. Therefore, we do not have access to Itô integration and most known techniques completely break down.

In a recent series of papers [Vov11a, Vov11b, Vov12], Vovk introduced an outer measure given by the cheapest pathwise superhedging price of the indicator function of a set. His aim was to characterize "typical price paths". The basic idea of Vovk, which we shall slightly modify in the following, is that "non-typical price paths" can be excluded since they are in a certain sense "too good to be true": they would allow investors to realize infinite profit while at the same time taking essentially no risk.

To be more precise, let Ω be the space of d-dimensional continuous paths (which represent all possible asset price trajectories), with coordinate process S. A process

$H \colon \Omega \times [0,T] \to \mathbb{R}^d$ is called a *simple strategy* if there exist a sequence of stopping times (τ_n) and $\mathcal{F}_{\tau_n}$-measurable bounded functions $F_n \colon \Omega \to \mathbb{R}^d$, such that for every $\omega \in \Omega$ we have $\tau_n(\omega) = \infty$ for all but finitely many n, and such that

$$H_t(\omega) = \sum_{n=0}^{\infty} F_n(\omega) \mathbf{1}_{(\tau_n(\omega), \tau_{n+1}(\omega)]}(t).$$

The *outer measure* of $A \subseteq \Omega$ is defined as the cheapest superhedging price for $\mathbf{1}_A$, that is

$$\overline{P}(A) := \inf \Big\{ \lambda > 0 : \exists\, (H^n) \subseteq \mathcal{H}_\lambda \text{ s.t. } \liminf_{n \to \infty} (\lambda + (H^n \cdot S)_T(\omega)) \geq \mathbf{1}_A(\omega) \,\forall \omega \in \Omega \Big\},$$

where $(H^n \cdot S)$ is the integral process of H with respect to S and for $\lambda > 0$ the set $\mathcal{H}_\lambda$ consists of all λ-admissible simple strategies, i.e. $H \in \mathcal{H}_\lambda$ if H is a simple strategy such that $(H \cdot S)_t(\omega) \geq -\lambda$ for all $\omega \in \Omega$, $t \in [0,T]$.

We start by observing that $\overline{P}$ is indeed an outer measure, which simultaneously dominates all local martingale measures on Ω. It comes with a natural arbitrage interpretation in terms of "no arbitrage of the first kind" (NA1): A set $A \subseteq \Omega$ is a null set under $\overline{P}$ if and only if there exists a sequence of 1-admissible simple strategies $(H^n) \subset \mathcal{H}_1$ such that

$$\liminf_{n \to \infty} (1 + (H^n \cdot S)_T(\omega)) \geq \infty \cdot \mathbf{1}_A(\omega), \quad \text{for all } \omega \in \Omega.$$

In other words, if a set A has outer measure 0, then we can make infinite profit by investing in the paths of A, without ever risking to lose more than the initial capital 1. Hence, we say that a property (P) holds for *typical price paths* if the set A where (P) is violated is a null set under $\overline{P}$.

In our first approach we do not restrict the set of paths and work on the whole space Ω. Vovk's outer measure allows us to define a topology on processes on Ω, and we observe that the "natural Itô integral" on simple functions is in a certain sense continuous in that topology. This is made precise in our "Model free version of Itô's isometry" (Lemma 2.2.4): We denoted by d_{loc} and d_∞ pseudometrics induced for $\overline{P}$, for details we refer to Section 2.1.4. For all $a, b, c > 0$ and a simple process F we have

$$\overline{P}\left(\{ \|(F \cdot S)\|_\infty \geq ab\sqrt{c} \} \cap \{ \|F\|_\infty \leq a \} \cap \{ \langle S \rangle_T \leq c \} \right) \leq 2 \exp(-b^2/(2d)),$$

where $\{\langle S \rangle_T \leq c\}$ denotes the set of all paths for which the quadratic variation $\langle S \rangle_T$ exist and is smaller than c.

This allows us to extend the integral from simple integrands to adapted càdlàg processes. The resulting integral is called "model free Itô integral" (Theorem 2.2.5): For any adapted càdlàg process F there exists an adapted process $\int_0^\cdot F \,\mathrm{d}S$ such that for every sequence of step functions (F^n) satisfying $\lim_n d_\infty(F^n, F) = 0$, we have $\lim_n d_{\mathrm{loc}}((F^n \cdot S), \int F \,\mathrm{d}S) = 0$. Furthermore, the map $F \mapsto \int F \,\mathrm{d}S$ is linear, satisfies

$$d_{\mathrm{loc}}\Big(\int F \,\mathrm{d}S, \int G \,\mathrm{d}S \Big) \lesssim d_\infty(F,G)^{1/2-\varepsilon}$$

for all $\varepsilon > 0$, and also the model free version of Itô's isometry extends to this setting. We stress once more that the entire construction is based only on financial arguments. Therefore, it has a purely financial interpretation and does not come from an artificially imposed probabilistic structure.

The second approach relies on the controlled rough path integral, which is more in the spirit of [Lyo95b, DOR14, DS14]. The controlled rough path integration has the advantage of being an entirely linear Banach space theory.

For a p-rough path (S, A) with $p \in (2, 3)$ and a function F controlled by S in the sense of Gubinelli, we recall that the rough path integral is defined as limit of compensated Riemann sums

$$\int_0^T F_s \, \mathrm{d}S_s := \lim_{|\pi| \to 0} \sum_{[s,t] \in \pi} F_s(S_t - S_s) + F'_s A_{s,t}.$$

We show that every typical price path has a natural Itô rough path associated to it. While its existence is directly ensured by our model free Itô integral, we need additional fine estimates to obtain the required regularity of the area process. This seems to be the first time that the area of a path is constructed in a nontrivial setting without using probability theory. Since in financial applications we can always restrict to typical price paths, this observation opens the door for applications of the controlled rough path integral in model free finance.

There is only one pitfall: the rough path integral is usually defined as a limit of compensated Riemann sums, which have no obvious financial interpretation. This sabotages our entire philosophy of only using arguments related to portfolio processes. This is why we show that under some weak regularity condition every rough path integral $\int F \, \mathrm{d}S$ is given as limit of non-anticipating Riemann sums that do not need to be compensated. Of course, this will not change anything in particular application, but it is of utmost importance from a philosophical point of view. Indeed, the justification for using the Itô integral in classical financial mathematics is crucially based on the fact that it is the limit of non-anticipating Riemann sums, even if in "every day applications" one never makes reference to that; see for example the discussion in [Lyo95b].

We use a certain "coarse-grained" regularity condition to obtain the rough path integral as limit of Riemann sums, which roughly says: Let (π_n) be an increasing sequence of partitions. Suppose S and A have finite p- respectively $p/2$-variation along the grid induced by the partition (π_n) and A is given as limit of Riemann sums along (π_n). This assumption (cf. Assumption (RIE) in Section 2.3.3) is weaker than the one required by Davie [Dav07]. Given our regularity condition the rough path integral can be obtained as limit of non-anticipating Riemann sums (see Theorem 2.3.19), i.e.

$$\int_0^T F_s \, \mathrm{d}S_s = \lim_{n \to \infty} \sum_{[t_k^n, t_{k+1}^n] \in \pi_n}^{N_n - 1} F_{t_k^n} S_{t_k^n, t_{k+1}^n}.$$

More importantly, every typical price path satisfies our "coarse-grained" assumption if we choose (t_k^n) to be a partition by suitable stopping times (τ_k^n).

Chapter 3: Local times for typical price paths and pathwise Tanaka formulas

This chapter is based on [PP15] by Perkowski and Prömel. It uses Vovk's [Vov12] game-theoretic approach to mathematical finance to construct local times for one-dimensional typical price paths.

While techniques of Chapter 2 were capable of treating integrands that are not necessarily functions of the integrator, the construction of $\int f(S)\,\mathrm{d}S$ required $f \in C^{1+\varepsilon}$ in the last Chapter. In Chapter 3 we show that for one-dimensional price processes this assumption can be greatly relaxed. Using a pathwise concept of local times, we derive various pathwise change of variable formulas. They generalize Föllmer's pathwise Itô formula in the same way as the classical Tanaka formula generalizes the classical Itô formula for continuous semimartingales.

In order move along the some lines in a purely analytic way, we define a discrete pathwise local time by setting

$$L_t^{\pi_n}(S,u) := \sum_{t_j^n \in \pi_n} \mathbf{1}_{(\!(S(t_j^n \wedge t), S(t_{j+1}^n \wedge t)]\!]}(u) |S(t_{j+1}^n \wedge t) - u|, \quad u \in \mathbb{R},$$

where (π_n) is a sequence of partitions with mesh size going to 0. This is our starting point for a pathwise version of a generalized Itô formula and Tanaka's formula, respectively.

Let us suppose that $S\colon [0,T] \to \mathbb{R}$ and $(L_t^{\pi_n}(S,\cdot))$ converge in L^2 along a sequence of partitions (π^n). Then Würmli [Wue80] proved for f in the Sobolev space H^2 that the generalized pathwise Itô formula

$$f(S(t)) = f(S(0)) + \int_0^t f'(S(s))\,\mathrm{d}S(s) + \int_{-\infty}^{\infty} f''(u) L_t(S,u)\,\mathrm{d}u$$

holds with

$$\int_0^t f'(S(s))\,\mathrm{d}S(s) := \lim_{n\to\infty} \sum_{t_j \in \pi^n} f'(S(t_j))(S(t_{j+1} \wedge t) - S(t_j \wedge t)), \quad t \in [0,\infty).$$

In Section 3.1 a pattern emerges: the more regular the local time is, the larger is the space of functions to which we can extend our pathwise stochastic integral. Indeed, Würlmi's result is based on the duality between the derivative of the integrand and the occupation measure. In his setting, the occupation measure has a density in L^2 and therefore defines a bounded functional on L^2. If the local time is continuous, then we can even integrate Radon measures against it.

Using the Young integral, it is possible to extend the pathwise Tanaka formula to a larger class of integrands, allowing us to integrate $\int g(S)\,\mathrm{d}S$ provided that g has finite q-variation for some $q > 1$, see Theorem 3.1.8. Therefore, we obtain a pathwise integral, which is given very naturally as a limit of Riemann sums.

To make our pathwise Tanaka formulas applicable in a framework of model free finance, we verify that every typical price path has a local time which satisfies all the requirements needed to apply our most general version of Itô-Tanaka formula, Theorem 3.1.8. Indeed, for typical price paths, the discrete local times $L^{\pi^n}(S,\cdot)$ converge uniformly in $(t,u) \in [0,T] \times \mathbb{R}$ to a continuous limit $L(S,\cdot)$, and for all $p > 2$ we have the discrete local times $(L_t^{\pi_n})$ have uniformly bounded p-variation for typical price paths, see Theorem 3.2.5. In particular, we can integrate $f(S)$ against a typical price path S whenever f has finite q-variation for some $q < 2$.

While we worked in Chapter 2 on a finite time horizon and with multidimensional price paths, the price paths are now assumed to be one-dimensional but may live on an infinite time horizon.

A remarkable result of [Vov12] roughly states: if $A \subseteq \Omega$ is "invariant under time changes" and such that $S_0(\omega) = 0$ for all $\omega \in A$, then $A \in \mathcal{F}$ and $\overline{P}(A) = \mathbb{W}(A)$, where $\mathbb{W}$ denotes the Wiener measure. This can be interpreted as a pathwise Dambis Dubins-Schwarz theorem. For instance, based on this theorem a model-independent super-replication theorem for time-invariant options in continuous time, given information on finitely many marginals, can be derived as done in Beiglböck et al. [BCH$^+$15]. Time-invariant options cover a broad range of exotic derivatives, including lookback options, discretely monitored Asian options, and options on realized variance.

In Appendix A.3, we sketch an alternative proof for the existence of local times based on Vovk's pathwise Dambis Dubins-Schwarz theorem, which allows for translating standard results for Brownian motion to typical price paths. For the Brownian local time all statements of Theorem 3.2.5 are well-known except one: the uniform boundedness in p-variation of the discrete local times.

Chapter 4: Existence of Lévy's area and pathwise integration

The theory of rough paths has established an analytical frame in which stochastic differential and integral calculus is traced back to properties of the trajectories of stochastic processes without reference to a particular probability measure, as we discuss in Section 2.3.

More recently, an alternative calculus with a more Fourier analytic touch has been designed (see [GIP14, Per14]) in which an older idea by Gubinelli [Gub04] is further developed. The existence of the rough path integral in this approach is seen to be linked to the existence of the corresponding Lévy area and the concept of "controlled paths". This raises the question about the relative power of the hypotheses leading to the existence of the integral. In Chapter 4, which was published in [IP15] by Imkeller and Prömel, we deal with this natural question.

In probability theory Lévy's area was first introduced by Lévy [Lé40]. For a d-dimensional path $X := (X^1, \dots, X^d)$ it can be defined via Riemann sums by

$$\mathbb{L}(X)^{i,j} := \int_0^T X_t^i \,\mathrm{d}X_t^j - \int_0^T X_t^j \,\mathrm{d}X_t^i := \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} (X_{s'}^i X_{s,t}^j - X_{s'}^j X_{s,t}^i),$$

for $1 \leq i, j \leq d$, where $s' \in [s,t] \in \pi$. Here the limit means "along every sequence of partitions with mesh going to zero".

Keeping in mind the concept of controlled paths in the sense of Gubinelli, to ensure the pathwise existence of Lévy's area, a suitable structure of control turns out to be sufficient. This new modification is called "self-control". We call a d-dimensional path X *self-controlled* if X^i is controlled by X^j or vice versa, for $i \neq j$. In Theorem 4.1.5 and Lemma 4.1.5 it is proven that this specific type of control always implies the existence of Lévy's area independently of the choice of the $s' \in [s,t] \in \pi$ and without any reference probability measure. If a path is not self-controlled, Lévy's area may not exit as we demonstrate in Example 4.1.8.

As already mentioned, these two concepts of control and Lévy's area play an important role in integration theory suitable for applications in stochastic analysis. Therefore, we link these concepts to Föllmer's pathwise integrals. Our analysis relies on the decomposition of the integral into its symmetric and its antisymmetric part, which is closely related to Lévy's area, i.e.

$$\gamma\text{-}\int_0^T Y_t \, \mathrm{d}X_t := \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} (Y_s + \gamma(Y_t - Y_s))(X_t - X_s) = \frac{1}{2}\mathbb{S}_\gamma\langle X, Y\rangle + \frac{1}{2}\mathbb{A}_\gamma\langle X, Y\rangle,$$

for $\gamma \in [0,1]$. We identify two different additional assumptions on the control relationship between Y and X which both lead to the existence of the antisymmetric part $\mathbb{A}_\gamma$ along every sequence of partitions and independently of γ, cf. Theorem 4.2.4. The first assumption imposes that Y' is symmetric and the second one is that X and Y are controlled mutually by each other. Each of these requirements directly gives the existence of the Stratonovich integral corresponding to $\gamma = \frac{1}{2}$, as seen in Corollary 4.2.6, since the symmetric part simplifies considerably in that case. For arbitrary γ the symmetric part can only be obtained along sequences of partitions for which the quadratic variation in the sense of Föllmer is guaranteed. Under the latter assumption we provide in Theorem 4.2.11 an Itô formula for controlled path Y with symmetric Y':

$$\gamma\text{-}\int_0^T Y_t \, \mathrm{d}^{\pi_n} X_t = \frac{1}{2}\text{-}\int_0^T Y_t \, \mathrm{d}X_t + \frac{1}{2}(2\gamma - 1) \sum_{1\le i,j\le d} \int_0^T Y_t'(i,j) \, \mathrm{d}^{\pi_n}[X^i, X^j]_t,$$

where $Y_t' = (Y_t'(i,j))_{1\le i,j\le d}$. As a consequence, this yields Föllmer's pathwise Itô formula, see Corollary 4.2.13

In recent years, functional Itô calculus, which extends classical calculus to functionals depending on the whole path of a stochastic process and not only on its current value, has received much attention. Based on the notion of derivatives due to Dupire [Dup09], in a series of papers Cont and Fournié [CF10a, CF10b, CF13] developed a functional Itô formula. One drawback of their approach is that the involved functional has to be defined on the space of càdlàg functions, or at least on a subspace strictly larger than the space of continuous functions $C([0,T],\mathbb{R}^d)$ (see [CR14]). In the spirit of Föllmer the paper [CF10b] provides a non-probabilistic version of a probabilistic Itô formula shown in [CF10a, CF13]. Referring to this approach we generalize in Theorem 4.2.14 Föllmer's pathwise Itô formula to twice Fréchet differentiable functionals defined on the space of continuous functions. Our functional Itô formula might be seen as the pathwise analogue to formulas stated in [Ahn97]. Let us stress that twice Fréchet differentiable functionals are generally beyond the scope of the concept of controlled paths as illustrated in Example 4.2.15.

Chapter 5: Rough differential equations on Besov spaces

The paracontrolled distribution approach recently introduced by Gubinelli, Imkeller and Perkowski [GIP12] is an extension of rough path theory to a multiparameter setting. It contains a concept of rough integration respectively multiplication of distributions, which very well fits with Hairer's theory of regularity structures [Hai14] to certain singular stochastic partial differential equations. The paracontrolled distribution approach works with tools from analysis like Bony's paraproduct and Littlewood-Paley theory.

In Chapter 5, which also appeared in [PT15] by Prömel and Trabs, we deal with rough differential equations (RDEs) on the very large and flexible class of Besov spaces $B^\alpha_{p,q}$ based on paracontrolled distributions. Intuitively, in $B^\alpha_{p,q}(\mathbb{R}^d)$ the regularity α is measured in the L^p-norm while q can be seen as a fine tuning parameter. The RDE considered is given by

$$\mathrm{d}u(t) = F(u(t))\xi(t), \quad u(0) = u_0, \quad t \in \mathbb{R}, \tag{1.5}$$

where $u_0 \in \mathbb{R}^m$ is the initial condition and F a family of vector fields on $\mathbb{R}^m$. The immediate and highly non-trivial problem appearing in equation (1.5) is that the product $F(u)\xi$ is not well-defined for very irregular signals ξ. While classical approaches as rough path theory formally integrate equation (1.5) and then give the appearing integral a meaning, the first step of our analysis is to give a direct meaning to the product in (1.5). For this purpose we generalize paracontrolled distributions from Hölder spaces $B^\alpha_{\infty,\infty}$, as studied in [GIP12], to $B^\alpha_{p,q}$.

It is well-known that the continuity of the Itô map, defined by mapping ξ to the solution trajectory u, can be restored with respect to a p-variation topology, cf. [Lyo98], as well as with respect to a Hölder topology, cf. [Fri05]. One core goal of Chapter 5 is to unify these two approaches in a common framework.

Our analysis relies on Littlewood-Paley theory: Taking a dyadic partition of unity (χ, ρ), every function f in a Besov spaces can be approximated in term of *Littlewood-Paley blocks*, i.e.

$$f = \lim_{j\to\infty} \sum_{i=-1}^{j-1} \Delta_i f \quad \text{with} \quad \Delta_i f := \mathcal{F}^{-1}(\rho_i \mathcal{F} f),$$

with $\rho_{-1} := \chi$, $\rho_i := \rho(2^{-j}\cdot)$ for $i \geq 0$, and $\mathcal{F}$ denotes the Fourier transform. Applying Bony's decomposition to $F(u)\xi$, which gives

$$F(u)\xi = \underbrace{T_{F(u)}\xi}_{\in B^{\alpha-1}_{p,q}} + \underbrace{\pi(F(u),\xi)}_{\in B^{2\alpha-1}_{p/2,q/2} \text{ if } 2\alpha-1>0} + \underbrace{T_\xi(F(u))}_{\in B^{2\alpha-1}_{p/2,q/2}}, \quad \text{for } \xi \in B^\alpha_{p,q},$$

where

$$T_{F(u)}\xi := \sum_{i\geq -1} S_{i-1}F(u)\Delta_i\xi \quad \text{and} \quad \pi(F(u),\xi) := \sum_{|i-j|\leq 1} \Delta_i F(u)\Delta_j\xi.$$

The stated Besov regularity of the different terms is presented in Lemma 5.1.1.

Provided the driving signal ξ is in $B^{\alpha-1}_{p,q}$ for $\alpha > 1/2$, $p \geq 2$, $q \geq 1$, the existence and uniqueness of a solution u to the RDE (1.5) is proven in Theorem 5.2.1 and further it is shown that the corresponding Itô map is locally Lipschitz continuous with respect to the Besov topology, see Theorem 5.2.2. In particular, with these results we recover the classical Young integration on Besov spaces.

In order to handle a more irregular driving signal ξ in $B^{\alpha-1}_{p,q}$ for $\alpha > 1/3$, $p \geq 3$, $q \geq 1$, we assume a control relation given by $u = T_{F(u)}\vartheta + u^{\#}$, where ϑ is the antiderivative of ξ and $u^{\#} \in B^{2\alpha}_{p/2,q}$. This yields that understanding the so called *resonant term* $\pi(u,\xi)$ boils down to the analysis of $\pi(\vartheta,\xi)$ thanks to the commutator lemma (Lemma 5.3.4) and the paralinearization result (Lemma (5.3.2)):

$$\pi(u,\xi) = F(u)\pi(\vartheta,\xi) + \underbrace{\pi(T_{F(u)}\vartheta,\xi) - F(u)\pi(\vartheta,\xi)}_{\in B^{3\alpha-1}_{p/3,q}} + \underbrace{\pi(u^{\#},\xi).}_{\in B^{3\alpha-1}_{p/3,q}}$$

Therefore, the path itself has to be enhanced with the "information" of the resonant term $\pi(\vartheta,\xi)$ instead of the first iterated integral common in rough path theory. In the spirit of the usual notion of geometric rough paths, this leads naturally to the new definition of the space of *geometric Besov rough paths*, cf. Definition 5.4.1, which is the closure of smooth paths ϑ enhanced with they resonant terms $\pi(\vartheta,\xi)$ in the Besov topology. Starting with a smooth path ϑ, it is shown that the Itô map associated to the RDE (1.5) extends continuously to the space of geometric Besov rough paths, cf. Theorem 5.4.8. As a consequence there exists a unique pathwise solution to the RDE (1.5) driven by a geometric Besov rough path.

Generalizing the approach from [GIP12] to Besov spaces poses severe additional problems, which are solved by using the Besov space characterizations via Littlewood-Paley blocks as well as the one via the modulus of continuity, cf (5.3). Besov spaces are a Banach algebra if and only if $p = q = \infty$. Hence, in general our results can only rely on pointwise multiplier theorems, Bony's decomposition and Besov embeddings. We thus need to generalize certain results in [BCD11] and [GIP12], including the already mentioned commutator lemma, see Lemma 5.3.4. A second difficulty is that the condition $u \in B^{\alpha}_{p,q}$ leads to an L^p-integrability requirement for u. To overcome this problem, we localize the signal and actually consider a weighted Itô(-Lyons) map, both done in a way that does not change the dynamics of the RDE (1.5) on a compact interval centered at the origin.

As an application we consider stochastic differential equations in Section 5.5, where the driving signal ξ is replaced by typical trajectories of stochastic processes including for example continuous martingales and Gaussian processes. But the prototypical example for our approach is a driving signal given by random functions via wavelet expansions, see Proposition 5.5.6.

Chapter 6: An FBSDE approach to the Skorokhod embedding problem for Gaussian processes with non-linear drift

In the last chapter we deal with fully coupled forward backward stochastic differential equations (FBSDEs) with the purpose to solve the Skorokhod embedding problem (SEP) for Gaussian processes with non-linear drift. This chapter is based on the joint work [FIP14] by Fromm, Imkeller and Prömel.

The classical goal of the SEP consists in finding, for a given Brownian motion B and a probability measure ν, a stopping time τ such that B_τ possesses the law ν. Since the first solution by Skorokhod [Sko61], there appeared many different constructions for the stopping time τ and generalizations of the original problem in the literature. Just to name some of the most famous solutions to the SEP we refer to Root [Roo69], Rost [Ros71] and Azéma-Yor [AY79]. A comprehensive survey can be found in [Obł04].

Recently, Skorokhod embedding raised additional interest because of its applications in financial mathematics, as for instance to obtain model-independent price bounds for lookback options [Hob98] or options on variance [CL10, CW13, OdR13].

In general terms, let us recall that a fully coupled system of FBSDEs is given by

$$\begin{aligned} X_s &= X_0 + \int_0^s \mu(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_0^s \sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r, \\ Y_t &= \xi(X_T) - \int_t^T f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \quad s, t \in [0, T]. \end{aligned} \tag{1.6}$$

Here the coefficient functions μ, σ of the forward part, the terminal condition ξ and the driver f of the backward component are supposed to be suitable functions.

There are mainly three methods to show the existence of solutions for a system of FBSDEs: the *contraction method* [Ant93, PT99], the *four step scheme* [MPY94] and the *method of continuation* [HP95, Yon97, PW99]. As an unified approach, [MWZZ15] and also [Del02] designed the theory of decoupling fields for FBSDEs, which was significantly refined in [FI13] to a multidimensional setting. We call a function u *decoupling field* if $Y = u(\cdot, X)$ holds, i.e. if the backward part Y of the FBSDE can be written as a functional of the forward part X, cf. Definition 6.2.1. The method of decoupling fields can primarily be seen as an extension of the contraction method.

In Chapter 6 we deal with Markovian systems of FBSDEs, that means all the involved coefficient functions $(\xi, (\mu, \sigma, f))$ are deterministic. This comes with the crucial advantage that heuristically we have

$$Z_s = u_x(s, X_s) \cdot \sigma(s, X_s, Y_s, Z_s), \quad s \in [0, T].$$

Allowing the coefficients (μ, σ, f) to be locally Lipschitz continuous in the control variable z, we develop an existence, uniqueness and regularity theory for FBSDEs in Section 6.2.2.

Equipped with these tools we are able to solve the Skorokhod embedding problem for Gaussian processes G of the form

$$G_t := G_0 + \int_0^t \alpha_s\,\mathrm{d}s + \int_0^t \beta_s\,\mathrm{d}W_s,$$

where $G_0 \in \mathbb{R}$ is a constant and $\alpha, \beta \colon [0, \infty) \to \mathbb{R}$ are suitable deterministic functions.

The spirit of our approach is related to the one by Bass [Bas83], who employed martingale representation to find a solution of the SEP for the Brownian motion. His approach was further developed for Brownian motion with linear drift in [AHI08] and time-homogeneous diffusion in [AHS15]. Bass' approach rests upon the observation that the SEP may be viewed as the weak version of a stochastic control problem: the goal is to steer G in such a way that it takes the distribution of a prescribed law. In the case of a Gaussian process G we reformulate the SEP in terms of FBSDEs in the form

$$X_s^{(1)} = \int_t^s 1\,\mathrm{d}W_r, \quad X_s^{(2)} = \int_t^s Z_r^2\,\mathrm{d}r, \quad Y_s = g(X_T^{(1)}) - \delta(X_T^{(2)}) - \int_s^T Z_r\,\mathrm{d}W_r,$$

where the function g is taken such that $g(X_T^{(1)})$ has the prescribed law ν and δ in such a way hat it encodes the information of the drift of G.

As a first step we construct a weak solution, i.e. we obtain a Gaussian process of the above form and an integrable random time such that, stopped at this time, the process possesses the given distribution ν. More precisely, if g and δ are both Lipschitz continuous, then there exist a Brownian motion $\tilde{B}$, a bounded stopping time $\tilde{\tau}$ and a constant $c \in \mathbb{R}$ such that $c + \int_0^{\tilde{\tau}} \alpha_s \, \mathrm{d}s + \int_0^{\tilde{\tau}} \beta_s \, \mathrm{d}\tilde{B}_s$ has law ν, see Lemma 6.3.2.

In order to transfer the result for the auxiliary Brownian motion $\tilde{B}$ to the given Brownian motion B, we need to control the growth of the gradient process $u_x(s, X_s)$, $s \in [0, T]$, and for this purpose have to describe this process by an intrinsic higher dimensional system of FBSDE. Provided g and δ are twice continuously differentiable with bounded derivative and both with Lipschitz continuous second derivative, cf. Lemma 6.3.10, the weak solution carries over to the originally given Gaussian process G. This finally solves the Skorokhod embedding problem for the Gaussian process G in the classical sense.

2. Pathwise stochastic integrals for model free finance

In this chapter we use Vovk's [Vov12] game-theoretic approach to develop two different techniques of stochastic integration in frictionless model free financial mathematics. A priori the integration problem is highly non-trivial in the model free context since we do not want to assume any probabilistic respectively semimartingale structure. Therefore, we do not have access to Itô integration and most known techniques completely break down. There are only two general solutions to the integration problem in a non-probabilistic continuous time setting that we are aware of. One was proposed by [DS14], who simply restrict themselves to trading strategies (integrands) of bounded variation. While this already allows to solve many interesting problems, it is not a very natural assumption to make in a frictionless market model. Indeed, in [DS14] a general duality approach is developed for pricing path-dependent derivatives that are Lipschitz continuous in the supremum norm, but so far their approach does not allow to treat derivatives depending on the volatility.

Another interesting solution was proposed by [DOR14] (using an idea which goes back to [Lyo95b]). They restrict the set of "possible price paths" to those admitting a quadratic variation. This allows them to apply Föllmer's pathwise Itô calculus [Fö81] to define pathwise stochastic integrals of the form $\int \nabla F(S)\,\mathrm{d}S$. In [Lyo95b] that approach was used to derive prices for American and European options under volatility uncertainty. In [DOR14] the given data is a finite number of European call and put prices and the derivative to be priced is a weighted variance swap. The restriction to the set of paths with quadratic variation is justified by referring to Vovk [Vov12], who proved that "typical price paths" (to be defined below) admit a quadratic variation.

In our first approach we do not restrict the set of paths and work on the space Ω of d-dimensional continuous paths (which represent all possible asset price trajectories). We follow Vovk in introducing an outer measure on Ω which is defined as the pathwise minimal superhedging price (in a suitable sense), and therefore has a purely financial interpretation and does not come from an artificially imposed probabilistic structure. Our first observation is that Vovk's outer measure allows us to define a topology on processes on Ω, and that the "natural Itô integral" on step functions is in a certain sense continuous in that topology. This allows us to extend the integral to càdlàg adapted integrands, and we call the resulting integral "model free Itô integral". We stress that the entire construction is based only on financial arguments.

Let us also stress that it is the *continuity* of our integral which is the most important aspect. Without reference to any topology the construction would certainly not be very useful, since already in the classical probabilistic setting virtually all applications of the Itô integral (SDEs, stochastic optimization, duality theory, ...) are based on the fact that it is a continuous operator.

This also motivates our second approach, which is more in the spirit of [Lyo95b, DOR14, DS14]. While in the first approach we do have a continuous operator, it is only continuous with respect to a sequence of pseudometrics and it seems impossible to find a Banach space structure that is compatible with it. This is a pity since Banach space theory is one of the key tools in the classical theory of financial mathematics, as emphasized for example in [DS01]. However, using the model free Itô integral we are able to show that every "typical price path" has a natural Itô rough path associated to it. Since in financial applications we can always restrict ourselves to typical price paths, this observation opens the door for the application of the controlled rough path integral [Lyo98, Gub04] in model free finance. Controlled rough path integration has the advantage of being an entirely linear Banach space theory which simultaneously extends

- the Riemann-Stieltjes integral of S against functions of bounded variation which was used by [DS14];
- the Young integral [You36]: typical price paths have finite p-variation for every $p > 2$, and therefore for every F of finite q-variation for $1 \leq q < 2$ (so that $1/p + 1/q > 1$), the integral $\int F \,\mathrm{d}S$ is defined as limit of non-anticipating Riemann sums;
- Föllmer's [Fö81] pathwise Itô integral, which was used by [Lyo95b, DOR14]. That this last integral is a special case of the controlled rough path integral is, to the best of our knowledge, proved rigorously for the first time in this chapter, although also [FH14] contains some observations in that direction.

In other words, our second approach covers all previously known techniques of integration in model free financial mathematics, while the first approach is much more general but at the price of leaving the Banach space world.

There is only one pitfall: the rough path integral is usually defined as a limit of compensated Riemann sums which have no obvious financial interpretation. This sabotages our entire philosophy of only using financial arguments. That is why we show that under some weak condition every rough path integral $\int F \,\mathrm{d}S$ is given as limit of non-anticipating Riemann sums that do not need to be compensated – the first time that such a statement is shown for general rough path integrals. Of course, this will not change anything in concrete applications, but it is of utmost importance from a philosophical point of view. Indeed, the justification for using the Itô integral in classical financial mathematics is crucially based on the fact that it is the limit of non-anticipating Riemann sums, even if in "every day applications" one never makes reference to that; see for example the discussion in [Lyo95b].

The chapter is organized as follows. Below we present a very incomplete list of solutions to the stochastic integration problem under model uncertainty and in a discrete time model free context (both a priori much simpler problems than the continuous time model free case), and we introduce some notations and conventions that will be used throughout the chapter. In Section 2.1 we briefly recall Vovk's game-theoretic approach to mathematical finance and introduce our outer measure. We also construct a topology on processes which is induced by the outer measure.

Section 2.2 is devoted to the construction of the model free Itô integral. Section 2.3 recalls some basic results from rough path theory, and continues by constructing rough paths associated to typical price paths. Here we also prove that the rough path integral is given as a limit of non-anticipating Riemann sums. Furthermore, we compare Föllmer's pathwise Itô integral with the rough path integral and prove that the latter is an extension of the former. Appendix A.1 recalls Vovk's pathwise Hoeffding inequality. In Appendix A.2 we show that a result of Davie which also allows to calculate rough path integrals as limits of Riemann sums is a special case of our results in Section 2.3.

Stochastic integration under model uncertainty

The first works which studied the option pricing problem under model uncertainty were [ALP95] and [Lyo95b], both considering the case of volatility uncertainty. As described above, [Lyo95b] is using Föllmer's pathwise Itô integral. In [ALP95] the problem is reduced to the classical setting by deriving a "worst case" model for the volatility.

A powerful tool in financial mathematics under model uncertainty is Karandikar's pathwise construction of the Itô integral [Kar95, Bic81], which allows to construct the Itô integral of a càdlàg integrand simultaneously under all semimartingale measures. The crucial point that makes the construction useful is that the Itô integral is a continuous operator under every semimartingale measure. While its pathwise definition would allow us to use the same construction also in a model free setting, it is not even clear what the output should signify in that case (for example the construction depends on a certain sequence of partitions and changing the sequence will change the output). Certainly it is not obvious whether the Karandikar integral is continuous in any topology once we dispose of semimartingale measures. A more general pathwise construction of the Itô integral was given in [Nut12], but it suffers from the same drawbacks with respect to applications in model free finance.

A general approach to stochastic analysis under model uncertainty was put forward in [DM06], and it is based on quasi sure analysis. This approach is extremely helpful when working under model uncertainty, but it also does not allow us to define stochastic integrals in a model free context.

In a related but slightly different direction, in [CDR11] non-semimartingale models are studied (which do not violate arbitrage assumptions if the set of admissible strategies is restricted). While the authors work under one fixed probability measure, the fact that their price process is not a semimartingale prevents them from using Itô integrals, a difficulty which is overcome by working with the Russo-Vallois integral [RV93].

Of course all these technical problems disappear if we restrict ourselves to discrete time, and indeed in that case [BHLP13] develop an essentially fully satisfactory duality theory for the pricing of derivatives under model uncertainty.

Notation and conventions

Throughout the chapter we fix $T \in (0,\infty)$ and we write $\Omega := C([0,T],\mathbb{R}^d)$ for the space of d-dimensional continuous paths. The coordinate process on Ω is denoted by $S_t(\omega) = \omega(t)$, $t \in [0,T]$. For $i \in \{1,\dots,d\}$, we also write $S^i_t(\omega) := \omega^i(t)$, where $\omega = (\omega^1,\dots,\omega^d)$. The filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is defined as $\mathcal{F}_t := \sigma(S_s : s \le t)$, and we set $\mathcal{F} := \mathcal{F}_T$. Stopping times τ and the associated σ-algebras $\mathcal{F}_\tau$ are defined as usual.

Unless explicitly stated otherwise, inequalities of the type $F_t \ge G_t$, where F and G are processes on Ω, are supposed to hold for all $\omega \in \Omega$, and not modulo null sets, as it is usually assumed in stochastic analysis.

The indicator function of a set A is denoted by $\mathbf{1}_A$.

A *partition* π of $[0,T]$ is a finite set of time points, $\pi = \{0 = t_0 < t_1 < \dots < t_m = T\}$. Occasionally, we will identify π with the set of intervals $\{[t_0,t_1],[t_1,t_2],\dots,[t_{m-1},t_m]\}$, and write expressions like $\sum_{[s,t]\in\pi}$.

For $f\colon [0,T] \to \mathbb{R}^n$ and $t_1,t_2 \in [0,T]$, denote $f_{t_1,t_2} := f(t_2) - f(t_1)$ and define the p-variation of f restricted to $[s,t] \subseteq [0,T]$ as

$$\|f\|_{p\text{-var},[s,t]} := \sup\left\{\left(\sum_{k=0}^{m-1} |f_{t_k,t_{k+1}}|^p\right)^{1/p} : s = t_0 < \dots < t_m = t, m \in \mathbb{N}\right\}, \quad p > 0, \tag{2.1}$$

(possibly taking the value $+\infty$). We set $\|f\|_{p\text{-var}} := \|f\|_{p\text{-var},[0,T]}$. We write $\Delta_T := \{(s,t) : 0 \le s \le t \le T\}$ for the simplex and define the p-variation of a function $g\colon \Delta_T \to \mathbb{R}^n$ in the same manner, replacing $f_{t_k,t_{k+1}}$ in (2.1) by $g(t_k,t_{k+1})$.

For $\alpha > 0$ and $\lfloor\alpha\rfloor := \max\{z \in \mathbb{Z} : z \le \alpha\}$, the space C^α consists of those functions that are $\lfloor\alpha\rfloor$ times continuously differentiable, with $(\alpha-\lfloor\alpha\rfloor)$-Hölder continuous partial derivatives of order $\lfloor\alpha\rfloor$ (and with continuous partial derivatives of order α in case $\alpha = \lfloor\alpha\rfloor$). The space C^α_b consists of those functions in C^α that are bounded, together with their partial derivatives, and we define the norm $\|\cdot\|_{C^\alpha_b}$ by setting

$$\|f\|_{C^\alpha_b} := \sum_{k=0}^{\lfloor\alpha\rfloor} \|D^k f\|_\infty + \mathbf{1}_{\alpha>\lfloor\alpha\rfloor}\|D^{\lfloor\alpha\rfloor} f\|_{\alpha-\lfloor\alpha\rfloor}, \tag{2.2}$$

where $\|\cdot\|_\beta$ denotes the β-Hölder norm for $\beta \in (0,1)$, and $\|\cdot\|_\infty$ denotes the supremum norm.

For $x,y \in \mathbb{R}^d$, we write $xy := \sum_{i=1}^d x_i y_i$ for the usual inner product. However, often we will encounter terms of the form $\int S\,\mathrm{d}S$ or $S_s S_{s,t}$ for $s,t \in [0,T]$, where we recall that S denotes the coordinate process on Ω. Those expressions are to be understood as the matrix $(\int S^i\,\mathrm{d}S^j)_{1\le i,j\le d}$, and similarly for $S_s S_{s,t}$. The interpretation will be usually clear from the context, otherwise we will make a remark to clarify things.

We use the notation $a \lesssim b$ if there exists a constant $c > 0$, independent of the variables under consideration, such that $a \le c \cdot b$. If we want to emphasize the dependence of c on the variable x, then we write $a(x) \lesssim_x b(x)$.

We make the convention that $0/0 := 0 \cdot \infty := 0$, $1 \cdot \infty := \infty$ and $\inf \emptyset := \infty$.

2.1. Superhedging and typical price paths

2.1.1. The outer measure and its basic properties

In a recent series of papers, Vovk [Vov08, Vov11a, Vov12] has introduced a model free, hedging based approach to mathematical finance that uses arbitrage considerations to examine which properties are satisfied by "typical price paths". This is achieved with the help of an outer measure given by the cheapest superhedging price.

Recall that $T \in (0, \infty)$ and $\Omega = C([0,T], \mathbb{R}^d)$ is the space of continuous paths, with coordinate process S, natural filtration $(\mathcal{F}_t)_{t\in[0,T]}$, and $\mathcal{F} = \mathcal{F}_T$. A process $H \colon \Omega \times [0,T] \to \mathbb{R}^d$ is called a *simple strategy* if there exist stopping times $0 = \tau_0 < \tau_1 < \dots$, and $\mathcal{F}_{\tau_n}$-measurable bounded functions $F_n \colon \Omega \to \mathbb{R}^d$, such that for every $\omega \in \Omega$ we have $\tau_n(\omega) = \infty$ for all but finitely many n, and such that

$$H_t(\omega) = \sum_{n=0}^{\infty} F_n(\omega) \mathbf{1}_{(\tau_n(\omega), \tau_{n+1}(\omega)]}(t).$$

In that case, the integral

$$(H \cdot S)_t(\omega) := \sum_{n=0}^{\infty} F_n(\omega)(S_{\tau_{n+1}\wedge t}(\omega) - S_{\tau_n \wedge t}(\omega)) = \sum_{n=0}^{\infty} F_n(\omega) S_{\tau_n\wedge t, \tau_{n+1}\wedge t}(\omega)$$

is well defined for all $\omega \in \Omega$, $t \in [0,T]$. Here $F_n(\omega) S_{\tau_n\wedge t, \tau_{n+1}\wedge t}(\omega)$ denotes the usual inner product on $\mathbb{R}^d$. For $\lambda > 0$, a simple strategy H is called λ-*admissible* if $(H \cdot S)_t(\omega) \geq -\lambda$ for all $\omega \in \Omega$, $t \in [0,T]$. The set of λ-admissible simple strategies is denoted by $\mathcal{H}_\lambda$.

Definition 2.1.1. The *outer measure* of $A \subseteq \Omega$ is defined as the cheapest superhedging price for $\mathbf{1}_A$, that is

$$\overline{P}(A) := \inf \Big\{ \lambda > 0 : \exists\, (H^n) \subseteq \mathcal{H}_\lambda \text{ s.t. } \liminf_{n\to\infty} (\lambda + (H^n \cdot S)_T(\omega)) \geq \mathbf{1}_A(\omega)\, \forall \omega \in \Omega \Big\}.$$

A set of paths $A \subseteq \Omega$ is called a *null set* if it has outer measure zero.

The term outer measure will be justified by Lemma 2.1.3 below. Our definition of $\overline{P}$ is very similar to the one used by Vovk [Vov12], but not quite the same. For a discussion see Section 2.1.3 below.

By definition, every Itô stochastic integral is the limit of stochastic integrals against simple strategies. Therefore, our definition of the cheapest superhedging price is essentially the same as in the classical setting, with one important difference: we require superhedging for *all* $\omega \in \Omega$, and not just almost surely.

Remark 2.1.2 ([Vov12], p. 564). *An equivalent definition of $\overline{P}$ would be*

$$\widetilde{P}(A) := \inf \Big\{ \lambda > 0 : \exists\, (H^n)_{n\in\mathbb{N}} \subseteq \mathcal{H}_\lambda \text{ s.t. } \liminf_{n\to\infty} \sup_{t\in[0,T]} (\lambda + (H^n \cdot S)_t(\omega)) \geq \mathbf{1}_A(\omega)\, \forall \omega \in \Omega \Big\}.$$

Clearly $\widetilde{P} \leq \overline{P}$. *To see the opposite inequality, let* $\widetilde{P}(A) < \lambda$. *Let* $(H^n)_{n\in\mathbb{N}} \subset \mathcal{H}_\lambda$ *be a sequence of simple strategies such that* $\liminf_{n\to\infty} \sup_{t\in[0,T]}(\lambda + (H^n \cdot S)_t) \geq \mathbf{1}_A$, *and let* $\varepsilon > 0$. *Define* $\tau_n := \inf\{t \in [0,T] : \lambda + \varepsilon + (H^n \cdot S)_t \geq 1\}$. *Then the stopped strategy* $G^n_t(\omega) := H^n_t(\omega)\mathbf{1}_{[0,\tau_n(\omega))}(t)$ *is in* $\mathcal{H}_\lambda \subseteq \mathcal{H}_{\lambda+\varepsilon}$ *and*

$$\liminf_{n\to\infty}(\lambda + \varepsilon + (G^n \cdot S)_T(\omega)) \geq \liminf_{n\to\infty} \mathbf{1}_{\{\lambda+\varepsilon+\sup_{t\in[0,T]}(H^n\cdot S)_t\geq 1\}}(\omega) \geq \mathbf{1}_A(\omega).$$

Therefore $\overline{P}(A) \leq \lambda + \varepsilon$, *and since* $\varepsilon > 0$ *was arbitrary* $\overline{P} \leq \widetilde{P}$, *and thus* $\overline{P} = \widetilde{P}$.

Lemma 2.1.3 ([Vov12], Lemma 4.1)**.** $\overline{P}$ *is in fact an outer measure, i.e. a nonnegative function defined on the subsets of* Ω *such that*

- $\overline{P}(\emptyset) = 0$;
- $\overline{P}(A) \leq \overline{P}(B)$ *if* $A \subseteq B$;
- *if* $(A_n)_{n\in\mathbb{N}}$ *is a sequence of subsets of* Ω, *then* $\overline{P}(\bigcup_n A_n) \leq \sum_n \overline{P}(A_n)$.

Proof. Monotonicity and $\overline{P}(\emptyset) = 0$ are obvious. So let (A_n) be a sequence of subsets of Ω. Let $\varepsilon > 0$, $n \in \mathbb{N}$, and let $(H^{n,m})_{m\in\mathbb{N}}$ be a sequence of $(\overline{P}(A_n)+\varepsilon 2^{-n-1})$-admissible simple strategies such that $\liminf_{m\to\infty}(\overline{P}(A_n)+\varepsilon 2^{-n-1}+(H^{n,m}\cdot S)_T) \geq \mathbf{1}_{A_n}$. Define for $m \in \mathbb{N}$ the $(\sum_n \overline{P}(A_n) + \varepsilon)$-admissible simple strategy $G^m := \sum_{n=0}^m H^{n,m}$. Then by Fatou's lemma

$$\begin{aligned}\liminf_{m\to\infty}\left(\sum_{n=0}^{\infty} \overline{P}(A_n) + \varepsilon + (G^m \cdot S)_T\right) &\geq \sum_{n=0}^{k} \left(\overline{P}(A_n) + \varepsilon 2^{-n-1} + \liminf_{m\to\infty}(H^{n,m} \cdot S)_T\right) \\ &\geq \mathbf{1}_{\bigcup_{n=0}^k A_n}\end{aligned}$$

for all $k \in \mathbb{N}$. Since the left hand side does not depend on k, we can replace $\mathbf{1}_{\bigcup_{n=0}^k A_n}$ by $\mathbf{1}_{\bigcup_n A_n}$ and the proof is complete. □

Maybe the most important property of $\overline{P}$ is that there exists an arbitrage interpretation for sets with outer measure zero:

Lemma 2.1.4. *A set* $A \subseteq \Omega$ *is a null set if and only if there exists a sequence of 1-admissible simple strategies* $(H^n)_n \subset \mathcal{H}_1$ *such that*

$$\liminf_{n\to\infty}(1 + (H^n \cdot S)_T(\omega)) \geq \infty \cdot \mathbf{1}_A(\omega), \tag{2.3}$$

where we use the convention $0 \cdot \infty = 0$ *and* $1 \cdot \infty := \infty$.

Proof. If such a sequence exists, then we can scale it down by an arbitrary factor $\varepsilon > 0$ to obtain a sequence of strategies in $\mathcal{H}_\varepsilon$ that superhedge $\mathbf{1}_A$, and therefore $\overline{P}(A) = 0$.

If conversely $\overline{P}(A) = 0$, then for every $n \in \mathbb{N}$ there exists a sequence of simple strategies $(H^{n,m})_{m\in\mathbb{N}} \subset \mathcal{H}_{2^{-n-1}}$ such that $2^{-n-1} + \liminf_{m\to\infty}(H^{n,m} \cdot \omega)_T \geq \mathbf{1}_A(\omega)$ for all $\omega \in \Omega$. Define $G^m := \sum_{n=0}^m H^{n,m}$, so that $G^m \in \mathcal{H}_1$. For every $k \in \mathbb{N}$ we obtain

$$\liminf_{m\to\infty}\left(1 + (G^m \cdot S)_T\right) \geq \sum_{n=0}^{k}\left(2^{-n-1} + \liminf_{m\to\infty}(H^{n,m} \cdot S)_T\right) \geq (k+1)\mathbf{1}_A.$$

Since the left hand side does not depend on k, the sequence (G^m) satisfies (2.3). □

In other words, if a set A has outer measure 0, then we can make infinite profit by investing in the paths from A, without ever risking to lose more than the initial capital 1.

This motivates the following definition:

Definition 2.1.5. We say that a property (P) holds for *typical price paths* if the set A where (P) is violated is a null set.

The basic idea of Vovk, which we shall adopt in the following, is that we only need to concentrate on typical price paths. Indeed, "non-typical price paths" can be excluded since they are in a certain sense "too good to be true": they would allow investors to realize infinite profit while at the same time taking essentially no risk.

2.1.2. Arbitrage notions and link to classical mathematical finance

Before we continue, let us discuss different notions of arbitrage and link our outer measure to classical mathematical finance. We start by observing that $\overline{P}$ is an outer measure which simultaneously dominates all local martingale measures on Ω.

Proposition 2.1.6 ([Vov12], Lemma 6.3)**.** *Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$, such that the coordinate process S is a $\mathbb{P}$-local martingale, and let $A \in \mathcal{F}$. Then $\mathbb{P}(A) \leq \overline{P}(A)$.*

Proof. Let $\lambda > 0$ and let $(H^n)_{n\in\mathbb{N}} \subseteq \mathcal{H}_\lambda$ be such that $\liminf_n(\lambda + (H^n \cdot S)_T) \geq \mathbf{1}_A$. Then

$$\mathbb{P}(A) \leq \mathbb{E}_\mathbb{P}[\liminf_n(\lambda + (H^n \cdot S)_T)] \leq \liminf_n \mathbb{E}_\mathbb{P}[\lambda + (H^n \cdot S)_T] \leq \lambda,$$

where in the last step we used that $\lambda + (H^n \cdot S)$ is a nonnegative $\mathbb{P}$-local martingale and thus a $\mathbb{P}$-supermartingale. □

This already indicates that $\overline{P}$-null sets are quite degenerate, in the sense that they are null sets under all local martingale measures. However, if that was the only reason for our definition of typical price paths, then a definition based on model free arbitrage opportunities would be equally valid. A map $X \colon \Omega \to [0, \infty)$ is a *model free arbitrage opportunity* if X is not identically 0 and if there exists $c > 0$ and a sequence $(H^n) \subseteq \mathcal{H}_c$ such that $\liminf_{n\to\infty}(H^n \cdot S)_T(\omega) = X(\omega)$ for all $\omega \in \Omega$. See [DH07, ABPS13] where (a similar) definition is used in the discrete time setting.

It might then appear more natural to say that a property holds for typical price paths if the indicator function of its complement is a model free arbitrage opportunity, rather than working with Definition 2.1.5. This "arbitrage definition" would also imply that any property which holds for typical price paths is almost surely satisfied under every local martingale measure. Nonetheless we decidedly claim that our definition is "the correct one". First of all the arbitrage definition would make our life much more difficult since it seems not very easy to work with. But of course this is only a convenience and cannot serve as justification of our approach. Instead, we argue by relating the two notions to classical mathematical finance.

For that purpose recall the fundamental theorem of asset pricing [DS94]: If $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$ under which S is a semimartingale, then there exists an equivalent measure $\mathbb{Q}$ such that S is a $\mathbb{Q}$-local martingale if and only if S admits *no free lunch with vanishing risk (NFLVR)*. But (NFLVR) is equivalent to the two conditions *no arbitrage (NA)* (intuitively: no profit without risk) and *no arbitrage opportunities of the first kind (NA1)* (intuitively: no very large profit with a small risk). The (NA) property holds if for every $c > 0$ and every sequence $(H^n) \subseteq \mathcal{H}_c$ for which $\lim_{n\to\infty}(H^n \cdot S)_T(\omega)$ exists for all ω we have $\mathbb{P}(\lim_{n\to\infty}(H^n \cdot S)_T < 0) > 0$ or $\mathbb{P}(\lim_{n\to\infty}(H^n \cdot S)_T = 0) = 1$. The (NA1) property holds if $\{1 + (H \cdot S)_T : H \in \mathcal{H}_1\}$ is bounded in $\mathbb{P}$-probability, i.e. if

$$\lim_{c\to\infty} \sup_{H\in\mathcal{H}_1} \mathbb{P}(1 + (H \cdot S)_T \geq c) = 0.$$

Strictly speaking this is (NA1) with simple strategies, but as observed by [KP11] (NA1) and (NA1) with simple strategies are equivalent; see also [Ank05, IP11].

It turns out that the arbitrage definition of typical price paths corresponds to (NA), while our definition corresponds to (NA1):

Proposition 2.1.7. *Let $A \in \mathcal{F}$ be a null set, and let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$ such that the coordinate process satisfies (NA1). Then $\mathbb{P}(A) = 0$.*

Proof. Let $(H^n)_{n\in\mathbb{N}} \subseteq \mathcal{H}_1$ be such that $1 + \liminf_n (H^n \cdot S)_T \geq \infty \cdot \mathbf{1}_A$. Then for every $c > 0$

$$\mathbb{P}(A) = \mathbb{P}(A \cap \{\liminf_{n\to\infty}(H^n \cdot S)_T > c\}) \leq \sup_{H\in\mathcal{H}_1} \mathbb{P}(\{(H \cdot S)_T > c\}).$$

By assumption, the right hand side converges to 0 as $c \to \infty$ and thus $\mathbb{P}(A) = 0$. □

Remark 2.1.8. *Proposition 2.1.7 is actually a consequence of Proposition 2.1.6, because if S satisfies (NA1) under $\mathbb{P}$, then there exists a dominating measure $\mathbb{Q} \gg \mathbb{P}$, such that S is a $\mathbb{Q}$-local martingale. See [Ruf13] for the case of continuous S, and [IP11] for the general case.*

The crucial point is that (NA1) is *the* essential property which every sensible market model has to satisfy, whereas (NA) is nice to have but not strictly necessary. Indeed, (NA1) is equivalent to the existence of an unbounded utility function such that the maximum expected utility is finite [KK07, IP11]. (NA) is what is needed in addition to (NA1) in order to obtain equivalent local martingale measures [DS94]. But there are perfectly viable models which violate (NA), for example the three dimensional Bessel process [DS95, KK07]. By working with the arbitrage definition of typical price paths, we would in a certain sense ignore these models.

2.1.3. Relation to Vovk's outer measure

Our definition of the outer measure $\overline{P}$ is not exactly the same as Vovk's [Vov12]. We find our definition more intuitive and it also seems to be easier to work with. However, since we rely on some of the results established by Vovk, let us compare the two notions.

For $\lambda > 0$, Vovk defines the set of processes

$$\mathcal{S}_\lambda := \Big\{ \sum_{k=0}^{\infty} H^k : H^k \in \mathcal{H}_{\lambda_k}, \lambda_k > 0, \sum_{k=0}^{\infty} \lambda_k = \lambda \Big\}. \tag{2.4}$$

For every $G = \sum_{k \geq 0} H^k \in \mathcal{S}_\lambda$, every $\omega \in \Omega$ and every $t \in [0,T]$, the integral

$$(G \cdot S)_t(\omega) := \sum_{k \geq 0} (H^k \cdot S)_t(\omega) = \sum_{k \geq 0} (\lambda_k + (H^k \cdot S)_t(\omega)) - \lambda$$

is well defined and takes values in $[-\lambda, \infty]$. Vovk then defines for $A \subseteq \Omega$ the cheapest superhedging price as

$$\overline{Q}(A) := \inf \{\lambda > 0 : \exists\, G \in \mathcal{S}_\lambda \text{ s.t. } \lambda + (G \cdot S)_T \geq \mathbf{1}_A\}. \tag{2.5}$$

This definition corresponds to the usual construction of an outer measure from an outer content (i.e. an outer measure which is only finitely subadditive and not countably subadditive); see [Fol99], Chapter 1.4, or [Tao11], Chapter 1.7. Here, the outer content is given by the cheapest superhedging price using only simple strategies. It is easy to see that $\overline{P}$ is dominated by $\overline{Q}$:

Lemma 2.1.9. *Let $A \subseteq \Omega$. Then $\overline{P}(A) \leq \overline{Q}(A)$.*

Proof. Let $G = \sum_k H^k$, with $H^k \in \mathcal{H}_{\lambda_k}$ and $\sum_k \lambda_k = \lambda$, and assume that $\lambda + (G \cdot S)_T \geq \mathbf{1}_A$. Then $(\sum_{k=0}^n H^k)_{n \in \mathbb{N}}$ defines a sequence of simple strategies in $\mathcal{H}_\lambda$, such that

$$\liminf_{n \to \infty} \Big(\lambda + \Big(\Big(\sum_{k=0}^n H^k \Big) \cdot S \Big)_T \Big) = \lambda + (G \cdot S)_T \geq \mathbf{1}_A.$$

So if $\overline{Q}(A) < \lambda$, then also $\overline{P}(A) \leq \lambda$, and therefore $\overline{P}(A) \leq \overline{Q}(A)$. □

Corollary 2.1.10. *For every $p > 2$, the set $A_p := \{\omega \in \Omega : \|S(\omega)\|_{p\text{-var}} = \infty\}$ has outer measure zero, that is $\overline{P}(A_p) = 0$.*

Proof. Theorem 1 of Vovk [Vov08] states that $\overline{Q}(A_p) = 0$, so $\overline{P}(A_p) = 0$ by Lemma 2.1.9. □

It is a remarkable result of [Vov12] that if $\Omega = C([0,\infty), \mathbb{R})$ (i.e. if the asset price process is one-dimensional), and if $A \subseteq \Omega$ is "invariant under time changes" and such that $S_0(\omega) = 0$ for all $\omega \in A$, then $A \in \mathcal{F}$ and $\overline{Q}(A) = \mathbb{P}(A)$, where $\mathbb{P}$ denotes the Wiener measure. This can be interpreted as a pathwise Dambis Dubins-Schwarz theorem.

2.1.4. A topology on path-dependent functionals

It will be very useful to introduce a topology on functionals on Ω. For that purpose let us identify $X, Y \colon \Omega \to \mathbb{R}$ if $X = Y$ for typical price paths. Clearly this defines an equivalence relation, and we write $\overline{L}_0$ for the space of equivalence classes. We then introduce the analogue of convergence in probability in our context: (X_n) *converges in outer measure* to X if

$$\lim_{n\to\infty} \overline{P}(|X_n - X| > \varepsilon) = 0 \quad \text{for all } \varepsilon > 0.$$

We follow [Vov12] in defining an expectation operator. If $X\colon \Omega \to [0,\infty]$, then

$$\overline{E}[X] := \inf\Big\{\lambda > 0 : \exists\, (H^n)_{n\in\mathbb{N}} \subseteq \mathcal{H}_\lambda \text{ s.t. } \liminf_{n\to\infty}(\lambda + (H^n \cdot S)_T(\omega)) \geq X(\omega)\, \forall \omega \in \Omega\Big\}. \tag{2.6}$$

In particular, $\overline{P}(A) = \overline{E}[\mathbf{1}_A]$. The expectation $\overline{E}$ is countably subadditive, monotone, and positively homogeneous. It is an easy exercise to verify that

$$d(X,Y) := \overline{E}[|X - Y| \wedge 1]$$

defines a metric on $\overline{L}_0$.

Lemma 2.1.11. *The distance d metrizes the convergence in outer measure. More precisely, a sequence (X_n) converges to X in outer measure if and only if*

$$\lim_n d(X_n, X) = 0.$$

Moreover, $(\overline{L}_0, d)$ is a complete metric space.

Proof. The arguments are the same as in the classical setting. Using subadditivity and monotonicity of the expectation operator, we have

$$\varepsilon \overline{P}(|X_n - X| \geq \varepsilon) \leq \overline{E}[|X_n - X| \wedge 1] \leq \overline{P}(|X_n - X| > \varepsilon) + \varepsilon$$

for all $\varepsilon \in (0,1]$, showing that convergence in outer measure is equivalent to convergence with respect to d.

As for completeness, let (X_n) be a Cauchy sequence with respect to d. Then there exists a subsequence (X_{n_k}) such that $d(X_{n_k}, X_{n_{k+1}}) \leq 2^{-k}$ for all k, so that

$$\overline{E}\Big[\sum_k (|X_{n_k} - X_{n_{k+1}}| \wedge 1)\Big] \leq \sum_k \overline{E}[|X_{n_k} - X_{n_{k+1}}| \wedge 1] = \sum_k d(X_{n_k}, X_{n_{k+1}}) < \infty,$$

which means that (X_{n_k}) converges for typical price paths. Define $X := \liminf_k X_{n_k}$. Then we have for all n and k

$$\begin{aligned} d(X_n, X) &\leq d(X_n, X_{n_k}) + d(X_{n_k}, X) \leq d(X_n, X_{n_k}) + \sum_{\ell \geq k} d(X_{n_\ell}, X_{n_{\ell+1}}) \\ &\leq d(X_n, X_{n_k}) + 2^{-k}. \end{aligned}$$

Choosing n and k large, we see that $d(X_n, X)$ tends to 0. $\square$

2.2. Model free Itô integration

The present section is devoted to the construction of a model free Itô integral. The main ingredient is a (weak) type of model free Itô isometry, which allows us to estimate the integral against a step function in terms of the amplitude of the step function and the quadratic variation of the price path. Using the topology introduced in Section 2.1.4, it is then easy to extend the integral to càdlàg integrands by a continuity argument.

Since we are in an unusual setting, let us spell out the following standard definitions:

Definition 2.2.1. A process $F\colon \Omega \times [0,T] \to \mathbb{R}^d$ is called *adapted* if the random variable $\omega \mapsto F_t(\omega)$ is $\mathcal{F}_t$-measurable for all $t \in [0,T]$.

The process F is said to be *càdlàg* if the sample path $t \mapsto F_t(\omega)$ is càdlàg for all $\omega \in \Omega$.

To prove our weak Itô isometry, we will need an appropriate sequence of stopping times: Let $n \in \mathbb{N}$. For each $i = 1, \dots, d$ define inductively

$$\sigma_0^{n,i} := 0, \qquad \sigma_{k+1}^{n,i} := \inf \{t \geq \sigma_k^{n,i} : |S_t^i - S^i_{\sigma_k^{n,i}}| \geq 2^{-n}\}, \quad k \in \mathbb{N}.$$

Since we are working with continuous paths and we are considering entrance times into closed sets, the maps $(\sigma^{n,i})$ are indeed stopping times, despite the fact that $(\mathcal{F}_t)$ is neither complete nor right-continuous. Denote $\pi^{n,i} := \{\sigma_k^{n,i} : k \in \mathbb{N}\}$. To obtain an increasing sequence of partitions, we take the union of the $(\pi^{n,i})$, that is we define $\sigma_0^n := 0$ and then

$$\sigma_{k+1}^n(\omega) := \min \left\{ t > \sigma_k^n(\omega) : t \in \bigcup_{i=1}^d \pi^{n,i}(\omega) \right\}, \quad k \in \mathbb{N}, \tag{2.7}$$

and we write $\pi^n := \{\sigma_k^n : k \in \mathbb{N}\}$ for the corresponding partition.

Lemma 2.2.2 ([Vov11a], Theorem 4.1)**.** *For typical price paths $\omega \in \Omega$, the quadratic variation along $(\pi^{n,i}(\omega))_{n\in\mathbb{N}}$ exists. That is,*

$$V_t^{n,i}(\omega) := \sum_{k=0}^{\infty} \big(S^i_{\sigma_{k+1}^{n,i} \wedge t}(\omega) - S^i_{\sigma_k^{n,i} \wedge t}(\omega)\big)^2, \quad t \in [0,T], \quad n \in \mathbb{N},$$

converges uniformly to a function $\langle S^i \rangle(\omega) \in C([0,T],\mathbb{R})$ for all $i \in \{1,\dots,d\}$.

For later reference, let us estimate $N_t^n := \max\{k \in \mathbb{N} : \sigma_k^n \leq t \text{ and } \sigma_k^n \neq 0\}$, the number of stopping times $\sigma_k^n \neq 0$ in π^n with values in $[0,t]$:

Lemma 2.2.3. *For all $\omega \in \Omega$, $n \in \mathbb{N}$, and $t \in [0,T]$, we have*

$$2^{-2n} N_t^n(\omega) \leq \sum_{i=1}^d V_t^{n,i}(\omega) =: V_t^n(\omega).$$

Proof. For $i \in \{1, \dots, d\}$ define $N_t^{n,i} := \max\{k \in \mathbb{N} : \sigma_k^{n,i} \leq t \text{ and } \sigma_k^{n,i} \neq 0\}$. Since S^i is continuous, we have $|S^i_{\sigma^{n,i}_{k+1}} - S^i_{\sigma^{n,i}_k}| = 2^{-n}$ as long as $\sigma^{n,i}_{k+1} \leq T$. Therefore, we obtain

$$N_t^n(\omega) \leq \sum_{i=1}^d N_t^{n,i}(\omega) = \sum_{i=1}^d \sum_{k=0}^{N_t^{n,i}(\omega)-1} \frac{1}{2^{-2n}} \left(S_{\sigma^{n,i}_{k+1}}(\omega) - S_{\sigma^{n,i}_k}(\omega)\right)^2 \leq 2^{2n} \sum_{i=1}^d V_t^{n,i}(\omega).$$

□

We will start by constructing the integral against step functions, which are defined similarly as simple strategies, except possibly unbounded: A process $F\colon \Omega \times [0,T] \to \mathbb{R}^d$ is called a *step function* if there exist stopping times $0 = \tau_0 < \tau_1 < \dots$, and $\mathcal{F}_{\tau_n}$-measurable functions $F_n\colon \Omega \to \mathbb{R}^d$, such that for every $\omega \in \Omega$ we have $\tau_n(\omega) = \infty$ for all but finitely many n, and such that

$$F_t(\omega) = \sum_{n=0}^{\infty} F_n(\omega) \mathbf{1}_{[\tau_n(\omega), \tau_{n+1}(\omega))}(t).$$

For notational convenience we are now considering the interval $[\tau_n(\omega), \tau_{n+1}(\omega))$ which is closed on the left-hand side. This allows us define the integral

$$(F \cdot S)_t := \sum_{n=0}^{\infty} F_n S_{\tau_n \wedge t, \tau_{n+1} \wedge t} = \sum_{n=0}^{\infty} F_{\tau_n} S_{\tau_n \wedge t, \tau_{n+1} \wedge t}, \quad t \in [0,T].$$

The following lemma will be the main building block in the construction of our integral.

Lemma 2.2.4 (Model free version of Itô's isometry)**.** *Let F be a step function. Then for all $a, b, c > 0$ we have*

$$\overline{P}\left(\{\|(F \cdot S)\|_\infty \geq ab\sqrt{c}\} \cap \{\|F\|_\infty \leq a\} \cap \{\langle S \rangle_T \leq c\}\right) \leq 2\exp(-b^2/(2d)),$$

where the set $\{\langle S \rangle_T \leq c\}$ should be read as

$$\{\langle S \rangle_T = \lim_n V_T^n \textit{ exists and satisfies } \langle S \rangle_T \leq c\}.$$

Proof. Assume $F_t = \sum_{n=0}^\infty F_n \mathbf{1}_{[\tau_n, \tau_{n+1})}(t)$ and set $\tau_a := \inf\{t > 0 : |F_t| \geq a\}$. Let $n \in \mathbb{N}$ and define $\rho_0^n := 0$ and then for $k \in \mathbb{N}$

$$\rho_{k+1}^n := \min\left\{t > \rho_k^n : t \in \pi^n \cup \{\tau_m : m \in \mathbb{N}\}\right\},$$

where we recall that $\pi^n = \{\sigma_k^n : k \in \mathbb{N}\}$ is the n-th generation of the dyadic partition generated by S. For $t \in [0,T]$, we have $(F \cdot S)_{\tau_a \wedge t} = \sum_k F_{\rho_k^n} S_{\tau_a \wedge \rho_k^n \wedge t, \tau_a \wedge \rho_{k+1}^n \wedge t}$, and by the definition of $\pi^n(\omega)$ and τ_a we get

$$\sup_{t \in [0,T]} \left|F_{\rho_k^n} S_{\tau_a \wedge \rho_k^n \wedge t, \tau_a \wedge \rho_{k+1}^n \wedge t}\right| \leq a\sqrt{d} 2^{-n}.$$

Hence, the pathwise Hoeffding inequality, Lemma A.1.1 in Appendix A.1, yields for every $\lambda \in \mathbb{R}$ the existence of a 1-admissible simple strategy $H^{\lambda,n} \in \mathcal{H}_1$ such that

$$1 + (H^{\lambda,n} \cdot S)_t \geq \exp\left(\lambda(F \cdot S)_{\tau_a \wedge t} - \frac{\lambda^2}{2}(N_t^{(\rho^n)} + 1)2^{-2n}a^2 d\right) =: \mathcal{E}^{\lambda,n}_{\tau_a \wedge t}$$

for all $t \in [0,T]$, where

$$N_t^{(\rho^n)} := \max\{k : \rho_k^n \leq t\} \leq N_t^n + N_t^{(\tau)} := N_t^n + \max\{k : \tau_k \leq t\}.$$

By Lemma 2.2.3, we have $N_t^n \leq 2^{2n} V_t^n$, so that

$$\mathcal{E}^{\lambda,n}_{\tau_a \wedge t} \geq \exp\left(\lambda(F \cdot S)_t - \frac{\lambda^2}{2} V_T^n a^2 d - \frac{\lambda^2}{2}(N_T^{(\tau)} + 1)2^{-2n}a^2 d\right).$$

If now $\|(F \cdot S)\|_\infty \geq ab\sqrt{c}$, $\|F(\omega)\|_\infty \leq a$ and $\langle S \rangle_T \leq c$, then

$$\liminf_{n\to\infty} \sup_{t\in[0,T]} \frac{\mathcal{E}_t^{\lambda,n} + \mathcal{E}_t^{-\lambda,n}}{2} \geq \frac{1}{2}\exp\left(\lambda ab\sqrt{c} - \frac{\lambda^2}{2} c a^2 d\right).$$

The argument inside the exponential is maximized for $\lambda = b/(a\sqrt{c}d)$, in which case we obtain $1/2 \exp(b^2/(2d))$. The statement now follows from Remark 2.1.2. □

Of course, we did not actually establish an isometry but only an upper bound for the integral. But this estimate is the key ingredient which allows us to extend the model free Itô integral to more general integrands, and it is this analogy to the classical setting that the terminology "model free version of Itô's isometry" alludes to.

Let us extend the topology of Section 2.1.4 to processes: we identify $X, Y : \Omega \times [0,T] \to \mathbb{R}^m$ if for typical price paths we have $X_t = Y_t$ for all $t \in [0,T]$, and we write $\overline{L}_0([0,T], \mathbb{R}^m)$ for the resulting space of equivalence classes which we equip with the distance

$$d_\infty(X,Y) := \overline{E}[\|X - Y\|_\infty \wedge 1]. \tag{2.8}$$

Ideally, we would like the stochastic integral on step functions to be continuous with respect to d_∞. However, using Proposition 2.1.6 it is easy to see that $\overline{P}(\|((1/n) \cdot S)\|_\infty > \varepsilon) = 1$ for all $n \in \mathbb{N}$ and $\varepsilon > 0$. This is why we also introduce for $c > 0$ the pseudometric

$$d_c(X,Y) := \overline{E}[(\|X - Y\|_\infty \wedge 1)\mathbf{1}_{\langle S \rangle_T \leq c}] \leq d_\infty(X,Y), \tag{2.9}$$

and then

$$d_{\mathrm{loc}}(X,Y) := \sum_{n=1}^\infty 2^{-n} d_{2^n}(X,Y) \leq d_\infty(X,Y). \tag{2.10}$$

The distance d_{loc} is somewhat analogous to the distance used to metrize the topology of uniform convergence on compacts, except that we do not localize in time but instead we control the size of the quadratic variation. For step functions F and G, we get from Lemma 2.2.4

$$\begin{aligned} d_c((F\cdot S),&(G\cdot S)) \\ &\le \overline{P}(\{\|((F-G)\cdot S)\|_\infty \ge ab\sqrt{c}\}\cap\{\|F-G\|_\infty\le a\}\cap\{\langle S\rangle_T\le c\}) \\ &\quad + \frac{d_c(F,G)}{a} + ab\sqrt{c} \\ &\le 2\exp\Big(-\frac{b^2}{2d}\Big) + \frac{d_c(F,G)}{a} + ab\sqrt{c} \end{aligned}$$

whenever $a,b>0$. Setting $a := \sqrt{d_c(F,G)}$ and $b := \sqrt{d|\log a|}$, we deduce that

$$d_c((F\cdot S),(G\cdot S)) \lesssim (1+\sqrt{c})d_c(F,G)^{1/2-\varepsilon} \tag{2.11}$$

for all $\varepsilon>0$, and in particular

$$d_{\mathrm{loc}}((F\cdot S),(G\cdot S)) \lesssim \sum_{n=1}^\infty 2^{-n/2} d_{2^n}(F,G)^{1/2-\varepsilon} \lesssim d_\infty(F,G)^{1/2-\varepsilon}.$$

Theorem 2.2.5. *Let F be an adapted, càdlàg process with values in $\mathbb{R}^d$. Then there exists $\int F\,\mathrm{d}S \in \overline{L}_0([0,T],\mathbb{R})$ such that for every sequence of step functions (F^n) satisfying $\lim_n d_\infty(F^n,F)=0$ we have $\lim_n d_{\mathrm{loc}}((F^n\cdot S),\int F\,\mathrm{d}S)=0$. The integral process $\int F\,\mathrm{d}S$ is continuous for typical price paths, and there exists a representative $\int F\,\mathrm{d}S$ which is adapted, although it may take the values $\pm\infty$. We usually write $\int_0^t F_s\,\mathrm{d}S_s := \int F\,\mathrm{d}S(t)$, and we call $\int F\,\mathrm{d}S$ the* model free Itô integral *of F with respect to S.*

The map $F\mapsto \int F\,\mathrm{d}S$ is linear, satisfies

$$d_{\mathrm{loc}}\Big(\int F\,\mathrm{d}S, \int G\,\mathrm{d}S\Big) \lesssim d_\infty(F,G)^{1/2-\varepsilon}$$

for all $\varepsilon>0$, and the model free version of Itô's isometry extends to this setting:

$$\overline{P}\Big(\Big\{\|\int F\,\mathrm{d}S\|_\infty \ge ab\sqrt{c}\Big\}\cap\{\|F\|_\infty\le a\}\cap\{\langle S\rangle_T\le c\}\Big) \le 2\exp(-b^2/(2d))$$

for all $a,b,c>0$.

Proof. Everything follows in a straightforward way from (2.11) in combination with Lemma 2.1.11. We have to use the fact that F is adapted and càdlàg in order to approximate it uniformly by step functions. □

Another simple consequence of our model free version of Itô's isometry is a strengthened version of Karandikar's [Kar95] pathwise Itô integral which works for all typical price paths and not just quasi surely under the local martingale measures.

Corollary 2.2.6. *In the setting of Theorem 2.2.5, let $(F^m)_{m\in\mathbb{N}}$ be a sequence of step functions with $\|F^m(\omega)-F(\omega)\|_\infty \le c_m$ for all $\omega\in\Omega$ and all $m\in\mathbb{N}$. Then for typical price paths ω there exists a constant $C(\omega)>0$ such that*

$$\Big\|(F^m\cdot S)(\omega) - \int F\,\mathrm{d}S(\omega)\Big\|_\infty \le C(\omega)c_m\sqrt{\log m} \tag{2.12}$$

for all $m\in\mathbb{N}$. So, if $c_m = o((\log m)^{-1/2})$, then for typical price paths $(F^m\cdot S)$ converges to $\int F\,\mathrm{d}S$.

Proof. For $c > 0$ the model free Itô isometry gives

$$\overline{P}\left(\left\{\|(F^m \cdot S) - \int F \,\mathrm{d}S\|_\infty \geq c_m \sqrt{4d \log m}\sqrt{c}\right\} \cap \{\langle S \rangle_T \leq c\}\right) \leq \frac{1}{m^2}.$$

Since this is summable in m, the claim follows from Borel Cantelli (which only requires countable subadditivity and can thus be applied for the outer measure $\overline{P}$). □

Remark 2.2.7. *The speed of convergence* (2.12) *is better than the one that can be obtained using the arguments in [Kar95], where the summability of* (c_m) *is needed.*

Remark 2.2.8. *It would be desirable to extend the robust Itô integral obtained in Theorem 2.2.5 to general locally square integrable integrands, that is adapted processes* H *with measurable trajectories and such that* $\int_0^t H_s^2(\omega)\,\mathrm{d}\langle S\rangle_s(\omega) < \infty$ *for all* t *and for all* ω *which have a continuous quadratic variation* $\langle S\rangle(\omega)$ *up to time* t*. The reason why our methods break down in this setting is that our "model free version of Itô's isometry" requires as input a uniform bound on the integrand. However, even with the restriction to càdlàg integrands our robust Itô integral is suitable for all (financial) applications which use Karandikar's pathwise stochastic integral [Kar95], with the great advantage of being a "model free" and not just a "quasi sure" object.*

Similarly, it would be nice to have an extension of Theorem 2.2.5 to càdlàg integrators. Unfortunately, neither the outer measure $\overline{P}$ *nor Vovk's outer measure* $\overline{Q}$ *have an obvious reasonable extension to the space* $D([0,T],\mathbb{R}^d)$ *of all càdlàg functions. The problem is that on this space there are no non-zero admissible strategies. As initiated in [Vov11a], it is possible to consider* $\overline{P}$ *or* $\overline{Q}$ *on the subspace of all paths in* $D([0,T],\mathbb{R}^d)$ *whose jump size at time* $t > 0$ *is bounded by a function of their supremum up to time* t*. However, it would be necessary to develop new techniques to obtain Theorem 2.2.5 in this setting since for instance the pathwise Hoeffding inequality (Lemma A.1.1) would not be applicable anymore.*

2.3. Rough path integration for typical price paths

Our second approach to model free stochastic integration is based on the rough path integral, which has the advantage of being a continuous linear operator between Banach spaces. The disadvantage is that we have to restrict the set of integrands to those "locally looking like S", modulo a smoother remainder. Our two main results in this section are that every typical price path has a naturally associated Itô rough path, and that the rough path integral can be constructed as limit of Riemann sums.

Let us start by recalling the basic definitions and results of rough path theory.

2.3.1. The Lyons-Gubinelli rough path integral

Here we follow more or less the lecture notes [FH14], to which we refer for a gentle introduction to rough paths. More advanced monographs are [LQ02, LCL07, FV10b]. The main difference to [FH14] in the derivation below is that we use p-variation to describe the regularity, and not Hölder continuity, because it is not true that all typical price paths are Hölder continuous. Also, we make an effort to give reasonably sharp results, whereas in [FH14] the focus lies more on the pedagogical presentation of the material. We stress that in this subsection we are merely collecting classical results.

Definition 2.3.1. A *control function* is a continuous map $c \colon \Delta_T \to [0, \infty)$ with $c(t,t) = 0$ for all $t \in [0,T]$ and such that $c(s,u) + c(u,t) \leq c(s,t)$ for all $0 \leq s \leq u \leq t \leq T$.

Observe that if $f \colon [0,T] \to \mathbb{R}^d$ satisfies $|f_{s,t}|^p \leq c(s,t)$ for all $(s,t) \in \Delta_T$, then the p-variation of f is bounded from above by $c(0,T)^{1/p}$.

Definition 2.3.2. Let $p \in (2,3)$. A p-rough path is a map $\mathbb{S} = (S, A) \colon \Delta_T \to \mathbb{R}^d \times \mathbb{R}^{d\times d}$ such that *Chen's relation*

$$S^i(s,t) = S^i(s,u) + S^i(u,t) \quad \text{and} \quad A^{i,j}(s,t) = A^{i,j}(s,u) + A^{i,j}(u,t) + S^i(s,u)S^j(u,t)$$

holds for all $1 \leq i,j \leq d$ and $0 \leq s \leq u \leq t \leq T$ and such that there exists a control function c with

$$|S(s,t)|^p + |A(s,t)|^{p/2} \leq c(s,t)$$

(in other words S has finite p-variation and A has finite $p/2$-variation). In that case we call A the *area* of S.

Remark 2.3.3. *Chen's relation simply states that S is the increment of a function, that is $S(s,t) = S(0,t) - S(0,s) = S_{s,t}$ for $S_t := S(0,t)$, and that for all i,j there exists a function $f^{i,j} \colon [0,T] \to \mathbb{R}$ such that $A^{i,j}(s,t) = f^{i,j}(t) - f^{i,j}(s) - S^i_s S^j_{s,t}$. Indeed, it suffices to set $f^{i,j}(t) := A^{i,j}(0,t) + S^i_0 S^j_{0,t}$.*

Remark 2.3.4. *The (strictly speaking incorrect) name "area" stems from the fact that if*
$S \colon [0,T] \to \mathbb{R}^2$ is a two-dimensional smooth function and if

$$A^{i,j}(s,t) = \int_s^t \int_s^{r_2} \mathrm{d}S^i_{r_1} \, \mathrm{d}S^j_{r_2} = \int_s^t S^i_{s,r_2} \, \mathrm{d}S^j_{r_2},$$

then the antisymmetric part of $A(s,t)$ corresponds to the algebraic area enclosed by the curve $(S_r)_{r\in[s,t]}$. It is a deep insight of Lyons [Lyo98], proving a conjecture of Föllmer, that the area is exactly the additional information which is needed to solve differential equations driven by S in a pathwise continuous manner, and to construct stochastic integrals as continuous maps. Actually, [Lyo98] solves a much more general problem and proves that if the driving signal is of finite p-variation for some $p > 1$, then it has to be equipped with the iterated integrals up to order $\lfloor p \rfloor - 1$ to obtain a continuous integral map. The for us relevant case $p \in (2,3)$ was already treated in [Lyo95a].

Example 2.3.5. If S is a continuous semimartingale and if we set $S(s,t) := S_{s,t}$ as well as

$$A^{i,j}(s,t) := \int_s^t \int_s^{r_2} \mathrm{d}S^i_{r_1} \, \mathrm{d}S^j_{r_2} = \int_s^t S^i_{s,r_2} \, \mathrm{d}S^j_{r_2},$$

where the integral can be understood either in the Itô or in the Stratonovich sense, then almost surely $\mathbb{S} = (S, A)$ is a p-rough path for all $p \in (2, 3)$. This is shown in [CL05], and we will give a simplified model free proof below (indeed we will show that every typical price path together with its model free Itô integral is a p-rough path for all $p \in (2, 3)$, from where the statement about continuous semimartingales easily follows).

From now on we fix $p \in (2, 3)$ and we assume that $\mathbb{S}$ is a p-rough path. Gubinelli [Gub04] observed that for every rough path there is a naturally associated Banach space of integrands, the space of *controlled paths*. Heuristically, a path F is controlled by S, if it locally "looks like S", modulo a smooth remainder. The precise definition is:

Definition 2.3.6. Let $p \in (2, 3)$ and $q > 0$ be such that $2/p+1/q > 1$. Let $\mathbb{S} = (S, A)$ be a p-rough path and let $F\colon [0, T] \to \mathbb{R}^n$ and $F'\colon [0, T] \to \mathbb{R}^{n\times d}$. We say that the pair (F, F') is *controlled* by S if the *derivative* F' has finite q-variation, and the *remainder* $R_F\colon \Delta_T \to \mathbb{R}^n$, defined by

$$R_F(s, t) := F_{s,t} - F'_s S_{s,t},$$

has finite r-variation for $1/r = 1/p + 1/q$. In this case, we write $(F, F') \in \mathscr{C}^q_{\mathbb{S}}$, and define

$$\|(F, F')\|_{\mathscr{C}^q_{\mathbb{S}}} := \|F'\|_{q\text{-}var} + \|R_F\|_{r\text{-var}}.$$

Equipped with the norm $|F_0| + |F'_0| + \|(F, F')\|_{\mathscr{C}^q_{\mathbb{S}}}$, the space $\mathscr{C}^q_{\mathbb{S}}$ is a Banach space.

Naturally, the function F' should be interpreted as the derivative of F with respect to S. The reason for considering pairs (F, F') and not just functions F is that the regularity requirement on the remainder R_F usually does not determine F' uniquely for a given path F. For example, if F and S both have finite r-variation rather than just finite p-variation, then for every F' of finite q-variation we have $(F, F') \in \mathscr{C}^q_{\mathbb{S}}$.

Note that we do not require F or F' to be continuous. We will point out in Remark 2.3.10 below why this does not pose any problem.

To gain a more "quantitative" feeling for the condition on q, let us assume for the moment that we can choose $p > 2$ arbitrarily close to 2 (which is the case in the example of a continuous semimartingale rough path). Then $2/p + 1/q > 1$ as long as $q > 0$, so that the derivative F' may essentially be as irregular as we want. The remainder R_F has to be of finite r-variation for $1/r = 1/p + 1/q$, so in other words it should be of finite r-variation for some $r < 2$ and thus slightly more regular than the sample path of a continuous local martingale.

Example 2.3.7. Let $\varepsilon \in (0, 1]$ be such that $(2 + \varepsilon)/p > 1$. Let $\varphi \in C^{1+\varepsilon}_b$ and define $F_s := \varphi(S_s)$ and $F'_s := \varphi'(S_s)$. Then $(F, F') \in \mathscr{C}^{p/\varepsilon}_{\mathbb{S}}$: Clearly F' has finite p/ε-variation. For the remainder, we have

$$|R_F(s, t)|^{p/(1+\varepsilon)} = |\varphi(S_t) - \varphi(S_s) - \varphi'(S_s)S_{s,t}|^{p/(1+\varepsilon)} \le \|\varphi\|_{C^{1+\varepsilon}_b} c(s, t),$$

where c is a control function for S. As the image of the continuous path S is compact, it is not actually necessary to assume that φ is bounded. We may always consider a $C^{1+\varepsilon}$ function ψ of compact support, such that ψ agrees with φ on the image of S.

This example shows that in general $R_F(s,t)$ is not a path increment of the form $R_F(s,t) = G(t) - G(s)$ for some function G defined on $[0,T]$, but really a function of two variables.

Example 2.3.8. Let G be a path of finite r-variation for some r with $1/p + 1/r > 1$. Setting $(F, F') = (G, 0)$, we obtain a controlled path in $\mathscr{C}^q_{\mathbb{S}}$, where $1/q = 1/r - 1/p$. In combination with Theorem 2.3.9 below, this example shows in particular that the controlled rough path integral extends the Young integral and the Riemann-Stieltjes integral.

The basic idea of rough path integration is that if we already know how to define $\int S \,\mathrm{d}S$, and if F looks like S on small scales, then we should be able to define $\int F \,\mathrm{d}S$ as well. The precise result is given by the following theorem:

Theorem 2.3.9 (Theorem 4.9 in [FH14], see also [Gub04], Theorem 1)**.** *Let $p \in (2,3)$ and $q > 0$ be such that $2/p + 1/q > 1$. Let $\mathbb{S} = (S, A)$ be a p-rough path and let $(F, F') \in \mathscr{C}^q_{\mathbb{S}}$. Then there exists a unique function $\int F \,\mathrm{d}S \in C([0,T], \mathbb{R}^n)$ which satisfies*

$$\Big| \int_s^t F_u \,\mathrm{d}S_u - F_s S_{s,t} - F'_s A(s,t) \Big| \lesssim \|S\|_{p\text{-var},[s,t]} \|R_F\|_{r\text{-var},[s,t]} + \|A\|_{p/2\text{-var},[s,t]} \|F'\|_{q\text{-var},[s,t]}$$

for all $(s,t) \in \Delta_T$. The integral is given as limit of the compensated Riemann sums

$$\int_0^t F_u \,\mathrm{d}S_u = \lim_{m \to \infty} \sum_{[s_1,s_2] \in \pi^m} [F_{s_1} S_{s_1,s_2} + F'_{s_1} A(s_1,s_2)], \tag{2.13}$$

where (π^m) is any sequence of partitions of $[0,t]$ with mesh size going to 0.

The map $(F, F') \mapsto (G, G') := (\int F_u \,\mathrm{d}S_u, F)$ is continuous from $\mathscr{C}^q_{\mathbb{S}}$ to $\mathscr{C}^p_{\mathbb{S}}$ and satisfies

$$\|(G, G')\|_{\mathscr{C}^p_{\mathbb{S}}} \lesssim \|F\|_{p\text{-var}} + (\|F'\|_\infty + \|F'\|_{q\text{-var}}) \|A\|_{p/2\text{-var}} + \|S\|_{p\text{-var}} \|R_F\|_{r\text{-var}}.$$

Remark 2.3.10. *To the best of our knowledge, there is no publication in which the controlled path approach to rough paths is formulated using p-variation regularity. The references on the subject all work with Hölder continuity. But in the p-variation setting, all the proofs work exactly as in the Hölder setting, and it is a simple exercise to translate the proof of Theorem 4.9 in [FH14] (which is based on Young's maximal inequality which we will encounter below) to obtain Theorem 2.3.9.*

There is only one small pitfall: We did not require F or F' to be continuous. The rough path integral for discontinuous functions is somewhat tricky, see [Wil01, FS14]. But here we do not run into any problems, because the integrand $\mathbb{S} = (S, A)$ is continuous. The construction based on Young's maximal inequality works as long as integrand and integrator have no common discontinuities, see the Theorem on page 264 of [You36].

If now $\varphi \in C_b^{1+\varepsilon}$ for some $\varepsilon > 0$, then using a Taylor expansion one can show that there exist $p > 2$ and $q > 0$ with $2/p + 1/q > 0$, such that $(F, F') \mapsto (\varphi(F), \varphi'(F)F')$ is a locally bounded map from $\mathscr{C}_{\mathbb{S}}^p$ to $\mathscr{C}_{\mathbb{S}}^q$. Combining this with the fact that the rough path integral is a bounded map from $\mathscr{C}_{\mathbb{S}}^q$ to $\mathscr{C}_{\mathbb{S}}^p$, it is not hard to prove the *existence* of solutions to the rough differential equation

$$X_t = x_0 + \int_0^t \varphi(X_s)\,\mathrm{d}S_s, \tag{2.14}$$

$t \in [0, T]$, where $X \in \mathscr{C}_{\mathbb{S}}^p$, $\int \varphi(X_s)\,\mathrm{d}S_s$ denotes the rough path integral, and S is a typical price path. Similarly, if $\varphi \in C_b^{2+\varepsilon}$, then the map $(F, F') \mapsto (\varphi(F), \varphi'(F)F')$ is locally Lipschitz continuous from $\mathscr{C}_{\mathbb{S}}^p$ to $\mathscr{C}_{\mathbb{S}}^q$, and this yields the *uniqueness* of the solution to (2.14) – at least among the functions in the Banach space $\mathscr{C}_{\mathbb{S}}^p$. See Section 5.3 of [Gub04] for details.

A remark is in order about the stringent regularity requirements on φ. In the classical Itô theory of SDEs, the function φ is only required to be Lipschitz continuous. But to solve a Stratonovich SDE, we need better regularity of φ. This is natural, because the Stratonovich SDE can be rewritten as an Itô SDE with a Stratonovich correction term: the equations

$$\mathrm{d}X_t = \varphi(X_t) \circ \mathrm{d}W_t \quad \text{and}$$
$$\mathrm{d}X_t = \varphi(X_t)\,\mathrm{d}W_t + \frac{1}{2}\varphi'(X_t)\varphi(X_t)\,\mathrm{d}t$$

are equivalent (where W is a standard Brownian motion, $\mathrm{d}W_t$ denotes Itô integration, and $\circ\,\mathrm{d}W_t$ denotes Stratonovich integration). To solve the second equation, we need $\varphi'\varphi$ to be Lipschitz continuous, which is always satisfied if $\varphi \in C_b^2$. But rough path theory cannot distinguish between Itô and Stratonovich integrals: If we define the area of W using Itô (respectively Stratonovich) integration, then the rough path solution of the equation will coincide with the Itô (respectively Stratonovich) solution. So in the rough path setting, the function φ should satisfy at least the same conditions as in the Stratonovich setting. The regularity requirements on φ are essentially sharp, see [Dav07], but the boundedness assumption can be relaxed, see [Lej12]. See also Section 10.5 of [FV10b] for a slight relaxation of the regularity requirements in the Brownian case.

Of course, the most interesting result of rough path theory is that the solution to a rough differential equation depends continuously on the driving signal. This is a consequence of the following observation:

Proposition 2.3.11 (Proposition 9.1 of [FH14])**.** *Let $p \in (2, 3)$ and $q > 0$ with $2/p + 1/q > 0$. Let $\mathbb{S} = (S, A)$ and $\tilde{\mathbb{S}} = (\tilde{S}, \tilde{A})$ be two p-rough paths, let $(F, F') \in \mathscr{C}_{\mathbb{S}}^q$ and $(\tilde{F}, \tilde{F}') \in \mathscr{C}_{\tilde{\mathbb{S}}}^q$. Then for every $M > 0$ there exists $C_M > 0$ such that*

$$\begin{aligned}\Big\| \int_0^{\cdot} F_s\,\mathrm{d}S_s - \int_0^{\cdot} \tilde{F}_s\,\mathrm{d}\tilde{S}_s \Big\|_{p\text{-var}} \leq & C_M \Big(|F_0 - \tilde{F}_0| + |F'_0 - \tilde{F}'_0| + \|F' - \tilde{F}'\|_{q\text{-var}} \\ & + \|R_F - R_{\tilde{F}}\|_{r\text{-var}} + \|S - \tilde{S}\|_{p\text{-var}} + \|A - \tilde{A}\|_{p/2\text{-var}} \Big),\end{aligned}$$

as long as

$$\max\{|F_0'| + \|(F,F')\|_{\mathscr{C}^q_{\mathbb{S}}}, |\tilde{F}_0'| + \|(\tilde{F},\tilde{F}')\|_{\mathscr{C}^q_{\tilde{\mathbb{S}}}}, \|S\|_{p\text{-var}}, \|A\|_{p/2\text{-var}}, \|\tilde{S}\|_{p\text{-var}}, \|\tilde{A}\|_{p/2\text{-var}}\}$$

is smaller or equal M.

In other words, the rough path integral depends on integrand and integrator in a locally Lipschitz continuous way, and therefore it is no surprise that the solutions to differential equations driven by rough paths depend continuously on the signal.

2.3.2. Typical price paths as rough paths

Our second approach to stochastic integration in model free financial mathematics is based on the rough path integral. Here we show that for every typical price path, the pair (S, A) is a p-rough path for all $p \in (2,3)$, where A corresponds to the model free Itô integral $\int S \, \mathrm{d}S$ which we constructed in Section 2.2. We also show that many Riemann sum approximations to $\int S \, \mathrm{d}S$ uniformly satisfy a certain coarse grained regularity condition, which we will use in the following section to prove that in our setting rough path integrals can be calculated as limits of Riemann sums (and not compensated Riemann sums as in Theorem 2.3.9). The main ingredient in the proofs will be our speed of convergence (2.12).

Theorem 2.3.12. *For $(s,t) \in \Delta_T$, $\omega \in \Omega$, and $i,j \in \{1,\dots,d\}$ define*

$$A^{i,j}_{s,t}(\omega) := \int_s^t S^i_r \, \mathrm{d}S^j_r(\omega) - S^i_s(\omega)S^j_{s,t}(\omega) := \int_0^t S^i_r \, \mathrm{d}S^j_r(\omega) - \int_0^s S^i_r \, \mathrm{d}S^j_r(\omega) - S^i_s(\omega)S^j_{s,t}(\omega),$$

where $\int S^i \, \mathrm{d}S^j$ is the integral constructed in Theorem 2.2.5. If $p > 2$, then for typical price paths $A = (A^{i,j})_{1 \le i,j \le d}$ has finite $p/2$-variation, and in particular $\mathbb{S} = (S, A)$ is a p-rough path.

Proof. Define the dyadic stopping times $(\tau^n_k)_{n,k\in\mathbb{N}}$ by $\tau^n_0 := 0$ and

$$\tau^n_{k+1} := \inf\{t \ge \tau^n_k : |S_t - S_{\tau^n_k}| = 2^{-n}\},$$

and set $S^n_t := \sum_k S_{\tau^n_k} \mathbf{1}_{[\tau^n_k, \tau^n_{k+1})}(t)$, so that $\|S^n - S\|_\infty \le 2^{-n}$. Accorcing to (2.12), for typical price paths ω there exists $C(\omega) > 0$ such that

$$\Big\|(S^n \cdot S)(\omega) - \int S \, \mathrm{d}S(\omega)\Big\|_\infty \le C(\omega) 2^{-n} \sqrt{\log n}.$$

Fix such a typical price path ω, which is also of finite q-variation for all $q > 2$ (recall from Corollary 2.1.10 that this is satisfied by typical price paths). Let us show that for such ω, the process A is of finite $p/2$-variation for all $p > 2$.

We have for $(s,t) \in \Delta_T$, omitting the argument ω of the processes under consideration,

$$\begin{aligned} |A_{s,t}| &\le \Big| \int_s^t S_r \, \mathrm{d}S_r - (S^n \cdot S)_{s,t} \Big| + |(S^n \cdot S)_{s,t} - S_s S_{s,t}| \\ &\le C 2^{-n} \sqrt{\log n} + |(S^n \cdot S)_{s,t} - S_s S_{s,t}| \lesssim_\varepsilon C 2^{-n(1-\varepsilon)} + |(S^n \cdot S)_{s,t} - S_s S_{s,t}| \end{aligned}$$

for every $n \in \mathbb{N}$, $\varepsilon > 0$. The second term on the right hand side can be estimated, using an argument based on Young's maximal inequality (see [LCL07], Theorem 1.16), by

$$|(S^n \cdot S)_{s,t} - S_s S_{s,t}| \lesssim \max\{2^{-n} c(s,t)^{1/q}, (\#\{k : \tau_k^n \in [s,t]\})^{1-2/q} c(s,t)^{2/q} + c(s,t)^{2/q}\}, \tag{2.15}$$

where $c(s,t)$ is a control function with $|S_{s,t}|^q \le c(s,t)$ for all $(s,t) \in \Delta_T$. Indeed, if there exists no k with $\tau_k^n \in [s,t]$, then $|(S^n \cdot S)_{s,t} - S_s S_{s,t}| \le 2^{-n} c(s,t)^{1/q}$, using that $|S_{s,t}| \le c(s,t)^{1/q}$. This corresponds to the first term in the maximum in (2.15).

Otherwise, note that at the price of adding $c(s,t)^{2/q}$ to the right hand side, we may suppose that $s = \tau_{k_0}^n$ for some k_0. Let now $\tau_{k_0}^n, \dots, \tau_{k_0+N-1}^n$ be those $(\tau_k^n)_k$ which are in $[s,t)$. Without loss of generality we may suppose $N \ge 2$, because otherwise $(S^n \cdot S)_{s,t} = S_s S_{s,t}$. Abusing notation, we write $\tau_{k_0+N}^n = t$. The idea is now to successively delete points $(\tau_{k_0+\ell}^n)$ from the partition, in order to pass from $(S^n \cdot S)$ to $S_s S_{s,t}$. By super-additivity of c, there must exist $\ell \in \{1, \dots, N-1\}$, for which

$$c(\tau_{k_0+\ell-1}^n, \tau_{k_0+\ell+1}^n) \le \frac{2}{N-1} c(s,t).$$

Deleting $\tau_{k_0+\ell}^n$ from the partition and subtracting the resulting integral from $(S^n \cdot S)_{s,t}$, we get

$$\begin{aligned}
&|S_{\tau_{k_0+\ell-1}^n} S_{\tau_{k_0+\ell-1}^n, \tau_{k_0+\ell}^n} + S_{\tau_{k_0+\ell}^n} S_{\tau_{k_0+\ell}^n, \tau_{k_0+\ell+1}^n} - S_{\tau_{k_0+\ell-1}^n} S_{\tau_{k_0+\ell-1}^n, \tau_{k_0+\ell+1}^n}| \\
&\qquad = |S_{\tau_{k_0+\ell-1}^n, \tau_{k_0+\ell}^n} S_{\tau_{k_0+\ell}^n, \tau_{k_0+\ell+1}^n}| \le c(\tau_{k_0+\ell-1}^n, \tau_{k_0+\ell+1}^n)^{2/q} \le \Big(\frac{2}{N-1} c(s,t)\Big)^{2/q}.
\end{aligned}$$

Successively deleting all the points except $\tau_{k_0}^n = s$ and $\tau_{k_0+N}^n = t$ from the partition gives

$$|(S^n \cdot S)_{s,t} - S_s S_{s,t}| \le \sum_{k=2}^{N} \Big(\frac{2}{k-1} c(s,t)\Big)^{2/q} \lesssim N^{1-2/q} c(s,t)^{2/q},$$

and therefore (2.15). Now it is easy to see that $\#\{k : \tau_k^n \in [s,t]\} \le 2^{nq} c(s,t)$ (compare also the proof of Lemma 2.2.3), and thus

$$\begin{aligned}
|A_{s,t}| &\lesssim_\varepsilon C 2^{-n(1-\varepsilon)} + \max\{2^{-n} c(s,t)^{1/q}, (2^{nq} c(s,t))^{1-2/q} c(s,t)^{2/q} + c(s,t)^{2/q}\} \\
&= C 2^{-n(1-\varepsilon)} + \max\{2^{-n} c(s,t)^{1/q}, 2^{-n(2-q)} c(s,t) + c(s,t)^{2/q}\}.
\end{aligned} \tag{2.16}$$

This holds for all $n \in \mathbb{N}$, $\varepsilon > 0$, $q > 2$. Let us suppose for the moment that $c(s,t) \le 1$ and let $\alpha > 0$ to be determined later. Then there exists $n \in \mathbb{N}$ for which $2^{-n-1} < c(s,t)^{1/\alpha(1-\varepsilon)} \le 2^{-n}$. Using this n in (2.16), we get

$$\begin{aligned}
|A_{s,t}|^\alpha &\lesssim_{\varepsilon,\omega,\alpha} c(s,t) + \max\Big\{c(s,t)^{1/(1-\varepsilon)} c(s,t)^{\alpha/q}, c(s,t)^{(2-q)/(1-\varepsilon)+\alpha} + c(s,t)^{2\alpha/q}\Big\} \\
&= c(s,t) + \max\Big\{c(s,t)^{\frac{q+\alpha(1-\varepsilon)}{q(1-\varepsilon)}}, c(s,t)^{\frac{2-q+\alpha(1-\varepsilon)}{1-\varepsilon}} + c(s,t)^{2\alpha/q}\Big\}.
\end{aligned}$$

We would like all the exponents in the maximum on the right hand side to be larger or equal to 1. For the first term, this is satisfied as long as $\varepsilon < 1$. For the third term, we need $\alpha \ge q/2$. For the second term, we need $\alpha \ge (q-1-\varepsilon)/(1-\varepsilon)$. Since $\varepsilon > 0$

can be chosen arbitrarily close to 0, it suffices if $\alpha > q - 1$. Now, since $q > 2$ can be chosen arbitrarily close to 2, we see that α can be chosen arbitrarily close to 1. In particular, we may take $\alpha = p/2$ for any $p > 2$, and we obtain $|A_{s,t}|^{p/2} \lesssim_\omega c(s,t)$.

It remains to treat the case $c(s,t) > 1$, for which we simply estimate

$$|A_{s,t}|^{p/2} \lesssim_p \Big\| \int_0^\cdot S_r \,\mathrm{d}S_r \Big\|_\infty^{p/2} + \|S\|_\infty^p \le \Big(\Big\| \int_0^\cdot S_r \,\mathrm{d}S_r \Big\|_\infty^{p/2} + \|S\|_\infty^p \Big) c(s,t).$$

So for every interval $[s,t]$ we can estimate $|A_{s,t}|^{p/2} \lesssim_{\omega,p} c(s,t)$, and the proof is complete. □

Remark 2.3.13. *To the best of our knowledge, this is one of the first times that a non-geometric rough path is constructed in a non-probabilistic setting, and certainly we are not aware of any works where rough paths are constructed using financial arguments.*

We also point out that, thanks to Proposition 2.1.6, we gave a simple, model free, and pathwise proof for the fact that a local martingale together with its Itô integral defines a rough path. While this seems intuitively clear, the only other proof that we know of is somewhat involved: it relies on a strong version of the Burkholder-Davis-Gundy inequality, a time change, and Kolmogorov's continuity criterion; see [CL05] or Chapter 14 of [FV10b].

The following auxiliary result will allow us to obtain the rough path integral as a limit of Riemann sums, rather than compensated Riemann sums, which are usually used to define it.

Lemma 2.3.14. *Let $(c_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers such that $c_n = o((\log n)^{-c})$ for all $c > 0$. For $n \in \mathbb{N}$ define $\tau_0^n := 0$ and $\tau_{k+1}^n := \inf\{t \ge \tau_k^n : |S_t - S_{\tau_k^n}| = c_n\}$, $k \in \mathbb{N}$, and set $S_t^n := \sum_k S_{\tau_k^n} \mathbf{1}_{[\tau_k^n, \tau_{k+1}^n)}(t)$. Then for typical price paths, $((S^n \cdot S))$ converges uniformly to $\int S \,\mathrm{d}S$ defined in Theorem 2.2.5. Moreover, for $p > 2$ and for typical price paths there exists a control function $c = c(p,\omega)$ such that*

$$\sup_n \sup_{k<\ell} \frac{|(S^n \cdot S)_{\tau_k^n,\tau_\ell^n}(\omega) - S_{\tau_k^n}(\omega) S_{\tau_k^n,\tau_\ell^n}(\omega)|^{p/2}}{c(\tau_k^n, \tau_\ell^n)} \le 1.$$

Proof. The uniform convergence of $((S^n \cdot S))$ to $\int S \,\mathrm{d}S$ follows from Corollary 2.2.6. For the second claim, fix $n \in \mathbb{N}$ and $k < \ell$ such that $\tau_\ell^n \le T$. Then

$$\begin{aligned} |(S^n \cdot S)_{\tau_k^n,\tau_\ell^n} - S_{\tau_k^n} S_{\tau_k^n,\tau_\ell^n}| &\lesssim \Big\| (S^n \cdot S) - \int_0^\cdot S_s \,\mathrm{d}S_s \Big\|_\infty + \Big| A_{\tau_k^n,\tau_\ell^n} \Big| \\ &\lesssim_\omega c_n \sqrt{\log n} + v_{p/2}(\tau_k^n, \tau_\ell^n)^{2/p} \lesssim_\varepsilon c_n^{1-\varepsilon} + v_{p/2}(\tau_k^n, \tau_\ell^n)^{2/p}, \end{aligned} \tag{2.17}$$

where $\varepsilon > 0$ and the last estimate holds by our assumption on the sequence (c_n), and where $v_{p/2}(s,t) := \|A\|_{p/2\text{-var},[s,t]}^{p/2}$ for $(s,t) \in \Delta_T$. Of course, this inequality only holds for typical price paths and not for all $\omega \in \Omega$.

On the other side, the same argument as in the proof of Theorem 2.3.12 (using Young's maximal inequality and successively deleting points from the partition) shows that

$$|(S^n \cdot S)_{\tau_k^n,\tau_\ell^n} - S_{\tau_k^n} S_{\tau_k^n,\tau_\ell^n}| \lesssim c_n^{2-q} v_q(\tau_k^n,\tau_\ell^n), \tag{2.18}$$

where $v_q(s,t) := \|S\|^q_{q\text{-var},[s,t]}$ for $(s,t) \in \Delta_T$.

Let us define the control function $\tilde{c} := v_q + v_{p/2}$. Take $\alpha > 0$ to be determined below. If $c_n > \tilde{c}(s,t)^{1/\alpha(1-\varepsilon)}$, then we use (2.18) and the fact that $2-q < 0$, to obtain

$$|(S^n \cdot S)_{\tau_k^n,\tau_\ell^n} - S_{\tau_k^n} S_{\tau_k^n,\tau_\ell^n}|^\alpha \lesssim (\tilde{c}(\tau_k^n,\tau_\ell^n))^{\frac{2-q}{(1-\varepsilon)}} v_q(\tau_k^n,\tau_\ell^n)^\alpha \le \tilde{c}(\tau_k^n,\tau_\ell^n)^{\frac{2-q+\alpha(1-\varepsilon)}{(1-\varepsilon)}}.$$

The exponent is larger or equal to 1 as long as $\alpha \ge (q-1-\varepsilon)/(1-\varepsilon)$. Since q and ε can be chosen arbitrarily close to 2 and 0 respectively, we can take $\alpha = p/2$, and get

$$|(S^n \cdot S)_{\tau_k^n,\tau_\ell^n} - S_{\tau_k^n} S_{\tau_k^n,\tau_\ell^n}|^{p/2} \lesssim \tilde{c}(\tau_k^n,\tau_\ell^n)(1+\tilde{c}(0,T)^\delta)$$

for a suitable $\delta > 0$.

On the other side, if $c_n \le \tilde{c}(s,t)^{1/\alpha(1-\varepsilon)}$, then we use (2.17) to obtain

$$|(S^n \cdot S)_{\tau_k^n,\tau_\ell^n} - S_{\tau_k^n} S_{\tau_k^n,\tau_\ell^n}|^\alpha \lesssim \tilde{c}(\tau_k^n,\tau_\ell^n) + \tilde{c}(\tau_k^n,\tau_\ell^n)^{2\alpha/p},$$

so that also in this case we may take $\alpha = p/2$, and thus we have in both cases

$$|(S^n \cdot S)_{\tau_k^n,\tau_\ell^n} - S_{\tau_k^n} S_{\tau_k^n,\tau_\ell^n}|^{p/2} \le c(\tau_k^n,\tau_\ell^n),$$

where c is a suitable (ω-dependent) multiple of $\tilde{c}$. □

2.3.3. The rough path integral as limit of Riemann sums

Theorem 2.3.12 shows that we can apply the controlled rough path integral in model free financial mathematics since every typical price path is a rough path. But there remains a philosophical problem: As we have seen in Theorem 2.3.9, the rough path integral $\int F \,\mathrm{d}S$ is given as limit of the compensated Riemann sums

$$\int_0^t F_s \,\mathrm{d}S_s = \lim_{m\to\infty} \sum_{[r_1,r_2]\in\pi^m} [F_{r_1} S_{r_1,r_2} + F'_{r_1} A(r_1,r_2)],$$

where (π^m) is an arbitrary sequence of partitions of $[0,t]$ with mesh size going to 0. The term $F_{r_1} S_{r_1,r_2}$ has an obvious financial interpretation as profit made by buying F_{r_1} units of the traded asset at time r_1 and by selling them at time r_2. However, for the "compensator" $F'_{r_1} A(r_1,r_2)$ there seems to be no financial interpretation, and therefore it is not clear whether the rough path integral can be understood as profit obtained by investing in S.

However, we observed in Section 2.2 that along suitable stopping times $(\tau_k^n)_{n,k}$, we have

$$\int_0^t S_s \,\mathrm{d}S_s = \lim_{n\to\infty} \sum_k S_{\tau_k^n} S_{\tau_k^n\wedge t,\tau_{k+1}^n\wedge t}.$$

By the philosophy of controlled paths, we expect that also for F which looks like S on small scales we should obtain

$$\int_0^t F_s \,\mathrm{d}S_s = \lim_{n\to\infty} \sum_k F_{\tau_k^n} S_{\tau_k^n \wedge t, \tau_{k+1}^n \wedge t},$$

without having to introduce the compensator $F'_{\tau_k^n} A(\tau_k^n \wedge t, \tau_{k+1}^n \wedge t)$ in the Riemann sum. With the results we have at hand, this statement is actually relatively easy to prove. Nonetheless, it seems not to have been observed before.

For the remainder of this section we fix $S \in C([0,T], \mathbb{R}^d)$, and we work under the following assumption:

Assumption (Rie)**.** Let $\pi^n = \{0 = t_0^n < t_1^n < \dots < t_{N_n}^n = T\}$, $n \in \mathbb{N}$, be a given sequence of partitions such that $\sup\{|S_{t_k^n, t_{k+1}^n}| : k = 0, \dots, N_n - 1\}$ converges to 0, and let $p \in (2,3)$. Set

$$S_t^n := \sum_{k=0}^{N_n - 1} S_{t_k^n} \mathbf{1}_{[t_k^n, t_{k+1}^n)}(t).$$

We assume that the Riemann sums $(S^n \cdot S)$ converge uniformly to a limit that we denote by $\int S \,\mathrm{d}S$, and that there exists a control function c for which

$$\sup_{(s,t)\in\Delta_T} \frac{|S_{s,t}|^p}{c(s,t)} + \sup_n \sup_{0\le k<\ell\le N_n} \frac{|(S^n \cdot S)_{t_k^n, t_\ell^n} - S_{t_k^n} S_{t_k^n, t_\ell^n}|^{p/2}}{c(t_k^n, t_\ell^n)} \le 1. \tag{2.19}$$

Remark 2.3.15. *We expect that "coarse-grained" regularity conditions as in* (2.19) *have been used for a long time, but were only able to find quite recent references: condition* (2.19) *was previously used in [Per14], see also [GIP14], and has also appeared independently in [Kel14]. In our setting this is quite a natural relaxation of a uniform p-variation bound since say for $s,t \in [t_k^n, t_{k+1}^n]$ with $|t-s| \ll |t_{k+1}^n - t_k^n|$ the increment of the discrete integral $(S^n \cdot S)_{s,t}$ is not a good approximation of $\int_s^t S_r \,\mathrm{d}S_r$, and therefore we cannot expect it to be close to $S_s S_{s,t}$.*

Remark 2.3.16. *Every typical price path satisfies* (Rie) *if we choose (t_k^n) to be a partition of stopping times such as the (τ_k^n) in Lemma 2.3.14.*

It is not hard to see that if S satisfies (Rie) and if we define $A(s,t) := \int_s^t S_r \,\mathrm{d}S_r - S_s S_{s,t}$, then (S, A) is a p-rough path. This means that we can calculate the rough path integral $\int F \,\mathrm{d}S$ whenever (F, F') is controlled by S, and the aim of the remainder of this section is to show that this integral is given as limit of (uncompensated) Riemann sums. Our proof is somewhat indirect. We translate everything from Itô type integrals to related Stratonovich type integrals, for which the convergence follows from the continuity of the rough path integral, Proposition 2.3.11. Then we translate everything back to our Itô type integrals. To go from Itô to Stratonovich, we need the quadratic variation:

Lemma 2.3.17. *Under Assumption* (Rie)*, let $1 \le i, j \le d$, and define*

$$\langle S^i, S^j \rangle_t := S_t^i S_t^j - S_0^i S_0^j - \int_0^t S_r^i \,\mathrm{d}S_r^j - \int_0^t S_r^j \,\mathrm{d}S_r^i.$$

Then $\langle S^i, S^j\rangle$ is a continuous function and

$$\langle S^i, S^j\rangle_t = \lim_{n\to\infty} \langle S^i, S^j\rangle^n_t = \lim_{n\to\infty} \sum_{k=0}^{N_n-1} (S^i_{t^n_{k+1}\wedge t} - S^i_{t^n_k\wedge t})(S^j_{t^n_{k+1}\wedge t} - S^j_{t^n_k\wedge t}). \tag{2.20}$$

The sequence $(\langle S^i, S^j\rangle^n)_n$ is of uniformly bounded total variation, and in particular $\langle S^i, S^j\rangle$ is of bounded variation. We write $\langle S\rangle = \langle S, S\rangle = (\langle S^i, S^j\rangle)_{1\le i,j\le d}$, and call $\langle S\rangle$ the quadratic variation *of S.*

Proof. The function $\langle S^i, S^j\rangle$ is continuous by definition. The specific form (2.20) of $\langle S^i, S^j\rangle$ follows from two simple observations:

$$S^i_t S^j_t - S^i_0 S^j_0 = \sum_{k=0}^{N_n-1} \left(S^i_{t^n_{k+1}\wedge t} S^j_{t^n_{k+1}\wedge t} - S^i_{t^n_k\wedge t} S^j_{t^n_k\wedge t} \right)$$

for every $n \in \mathbb{N}$, and

$$\begin{aligned} &S^i_{t^n_{k+1}\wedge t} S^j_{t^n_{k+1}\wedge t} - S^i_{t^n_k\wedge t} S^j_{t^n_k\wedge t} \\ &\qquad = S^i_{t^n_k\wedge t} S^j_{t^n_k\wedge t, t^n_{k+1}\wedge t} + S^j_{t^n_k\wedge t} S^i_{t^n_k\wedge t, t^n_{k+1}\wedge t} + S^i_{t^n_k\wedge t, t^n_{k+1}\wedge t} S^j_{t^n_k\wedge t, t^n_{k+1}\wedge t}, \end{aligned}$$

so that the convergence in (2.20) is a consequence of the convergence of $(S^n \cdot S)$ to $\int S \,\mathrm{d}S$.

To see that $\langle S^i, S^j\rangle$ is of bounded variation, note that

$$S^i_{t^n_k\wedge t, t^n_{k+1}\wedge t} S^j_{t^n_k\wedge t, t^n_{k+1}\wedge t} = \frac{1}{4}\left(\left((S^i + S^j)_{t^n_k\wedge t, t^n_{k+1}\wedge t} \right)^2 - \left((S^i - S^j)_{t^n_k\wedge t, t^n_{k+1}\wedge t} \right)^2 \right)$$

(read $\langle S^i, S^j\rangle = 1/4(\langle S^i + S^j\rangle - \langle S^i - S^j\rangle)$). In other words, the n-th approximation of $\langle S^i, S^j\rangle$ is the difference of two increasing functions, and its total variation is bounded from above by

$$\begin{aligned} &\sum_{k=0}^{N_n-1} \left(\left((S^i + S^j)_{t^n_k, t^n_{k+1}} \right)^2 + \left((S^i - S^j)_{t^n_k, t^n_{k+1}} \right)^2 \right) \\ &\qquad\qquad \lesssim \sup_m \sum_{k=0}^{N_m-1} \left((S^i_{t^m_k, t^m_{k+1}})^2 + (S^j_{t^m_k, t^m_{k+1}})^2 \right). \end{aligned}$$

Since the right hand side is finite, also the limit $\langle S^i, S^j\rangle$ is of bounded variation. □

Given the quadratic variation, the existence of the Stratonovich integral is straightforward:

Lemma 2.3.18. *Under Assumption* (Rie)*, define $\tilde{S}^n|_{[t^n_k, t^n_{k+1}]}$ as the linear interpolation of $S_{t^n_k}$ and $S_{t^n_{k+1}}$ for $k = 0, \dots N_n - 1$. Then $(\int \tilde{S}^n \,\mathrm{d}\tilde{S}^n)$ converges uniformly to*

$$\int_s^t S_r \circ \mathrm{d}S_r := \int_s^t S_r \,\mathrm{d}S_r + \frac{1}{2}\langle S\rangle_{s,t}. \tag{2.21}$$

Moreover, setting $\tilde{A}^n(s,t) := \int_s^t \tilde{S}^n_{s,r} \,\mathrm{d}\tilde{S}^n_r$ for $(s,t) \in \Delta_T$, we have $\sup_n \|\tilde{A}^n\|_{p/2\text{-var}} < \infty$.

Proof. Let $n \in \mathbb{N}$ and $k \in \{0, \dots, N_n - 1\}$. Then for $t \in [t_k^n, t_{k+1}^n]$ we have

$$\tilde{S}_t^n = S_{t_k^n} + \frac{t - t_k^n}{t_{k+1}^n - t_k^n} S_{t_k^n, t_{k+1}^n},$$

so that

$$\int_{t_k^n}^{t_{k+1}^n} \tilde{S}_r^n \,\mathrm{d}\tilde{S}_r^n = S_{t_k^n} S_{t_k^n, t_{k+1}^n} + \frac{1}{2} S_{t_k^n, t_{k+1}^n} S_{t_k^n, t_{k+1}^n}, \tag{2.22}$$

from where the uniform convergence and the representation (2.21) follow by Lemma 2.3.17.

To prove that $\tilde{A}^n$ has uniformly bounded $\frac{p}{2}$-variation, consider $(s,t) \in \Delta_T$. If there exists k such that $t_k^n \le s < t \le t_{k+1}^n$, then we estimate

$$\begin{aligned} |\tilde{A}^n(s,t)|^{p/2} &= \Big| \int_s^t \tilde{S}_{s,r}^n \,\mathrm{d}\tilde{S}_r^n \Big|^{p/2} \le \Big| \int_s^t (r - s) \frac{|S_{t_k^n, t_{k+1}^n}|^2}{|t_{k+1}^n - t_k^n|^2} \,\mathrm{d}r \Big|^{p/2} \\ &= \frac{1}{2^{p/2}} |t - s|^p \frac{|S_{t_k^n, t_{k+1}^n}|^p}{|t_{k+1}^n - t_k^n|^p} \le \frac{|t-s|}{|t_{k+1}^n - t_k^n|} \|S\|_{p\text{-var},[t_k^n, t_{k+1}^n]}^p. \end{aligned} \tag{2.23}$$

Otherwise, let k_0 be the smallest k such that $t_k^n \in (s,t)$, and let k_1 be the largest such k. We decompose

$$\tilde{A}^n(s,t) = \tilde{A}^n(s, t_{k_0}^n) + \tilde{A}^n(t_{k_0}^n, t_{k_1}^n) + \tilde{A}^n(t_{k_1}^n, t) + \tilde{S}_{s,t_{k_0}^n}^n \tilde{S}_{t_{k_0}^n, t_{k_1}^n}^n + \tilde{S}_{s,t_{k_1}^n}^n \tilde{S}_{t_{k_1}^n, t}^n.$$

We get from (2.22) that

$$|\tilde{A}^n(t_{k_0}^n, t_{k_1}^n)|^{p/2} \lesssim |(S^n \cdot S)_{t_{k_0}^n, t_{k_1}^n} - S_{t_{k_0}^n} S_{t_{k_0}^n, t_{k_1}^n}|^{p/2} + (\langle S \rangle_{t_{k_0}^n, t_{k_1}^n}^n)^{p/2},$$

where $\langle S \rangle^n$ denotes the n-th approximation of the quadratic variation. By the assumption (Rie) and Lemma 2.3.17, there exists a control function $\tilde{c}$ so that the right hand side is bounded from above by $\tilde{c}(t_{k_0}^n, t_{k_1}^n)$. Combining this with (2.23) and a simple estimate for the terms $\tilde{S}_{s,t_{k_0}^n}^n \tilde{S}_{t_{k_0}^n, t_{k_1}^n}^n$ and $\tilde{S}_{s,t_{k_1}^n}^n \tilde{S}_{t_{k_1}^n, t}^n$, we deduce that $\|\tilde{A}^n\|_{p/2\text{-var}} \lesssim \tilde{c}(0,T) + \|S\|_{p\text{-var}}^2$, and the proof is complete. $\square$

We are now ready to prove the main result of this section.

Theorem 2.3.19. *Under Assumption* (Rie), *let $q > 0$ be such that $2/p + 1/q > 1$. Let $(F, F') \in \mathscr{C}_{\mathbb{S}}^q$ be a controlled path such that F is continuous. Then the rough path integral $\int F \,\mathrm{d}S$ which was defined in Theorem 2.3.9 is given by*

$$\int_0^t F_s \,\mathrm{d}S_s = \lim_{n \to \infty} \sum_{k=0}^{N_n - 1} F_{t_k^n} S_{t_k^n \wedge t, t_{k+1}^n \wedge t},$$

where the convergence is uniform in t.

Proof. For $n \in \mathbb{N}$ define $\tilde{F}^n$ as the linear interpolation of F between the points in π^n. Then $(\tilde{F}^n, F')$ is controlled by $\tilde{S}^n$: Clearly $\|\tilde{F}^n\|_{q\text{-var}} \le \|F\|_{q\text{-var}}$. The remainder

$\tilde{R}^n_{\tilde{F}^n}$ of $\tilde{F}^n$ with respect to $\tilde{S}^n$ is given by $\tilde{R}^n_{\tilde{F}^n}(s,t) = \tilde{F}^n_{s,t} - F'_s \tilde{S}^n_{s,t}$ for $(s,t) \in \Delta_T$. We need to show that $\tilde{R}^n_{\tilde{F}^n}$ has finite r-variation for $1/r = 1/p + 1/q$.

If $t^n_k \leq s \leq t \leq t^n_{k+1}$, we have

$$
\begin{aligned}
|\tilde{R}^n_{\tilde{F}^n}(s,t)|^r &= \Big| \frac{t-s}{t^n_{k+1} - t^n_k} F_{t^n_k, t^n_{k+1}} - F'_s \frac{t-s}{t^n_{k+1} - t^n_k} S_{t^n_k, t^n_{k+1}} \Big|^r \\
&\leq \Big| \frac{t-s}{t^n_{k+1} - t^n_k} \Big|^r \big(\|R_F\|_{r\text{-var},[t^n_k, t^n_{k+1}]} + \|F'\|^{r/q}_{q\text{-var},[t^n_k, s]} \|S\|^{r/p}_{p\text{-var},[t^n_k, t^n_{k+1}]} \big) \\
&\leq \frac{|t-s|}{|t^n_{k+1} - t^n_k|} \big(\|R_F\|_{r\text{-var},[t^n_k, t^n_{k+1}]} + \|F'\|_{q\text{-var},[t^n_k, t^n_{k+1}]} + \|S\|_{p\text{-var},[t^n_k, t^n_{k+1}]} \big),
\end{aligned}
\tag{2.24}
$$

where in the last step we used that $1/r = 1/p + 1/q$, and thus $r/q + r/p = 1$.

Otherwise, there exists $k \in \{1, \ldots, N_n - 1\}$ with $t^n_k \in (s,t)$. Let k_0 and k_1 be the smallest and largest such k, respectively. Then

$$
\begin{aligned}
|\tilde{R}^n_{\tilde{F}^n}(s,t)|^r \lesssim_r & |\tilde{R}^n_{\tilde{F}^n}(s, t^n_{k_0})|^r + |\tilde{R}^n_{\tilde{F}^n}(t^n_{k_0}, t^n_{k_1})|^r \\
& + |\tilde{R}^n_{\tilde{F}^n}(t^n_{k_1}, t)|^r + |F'_{s, t^n_{k_0}} S_{t^n_{k_0}, t^n_{k_1}}|^r + |F'_{s, t^n_{k_1}} S_{t^n_{k_1}, t}|^r.
\end{aligned}
$$

Now $\tilde{R}^n_{\tilde{F}^n}(t^n_{k_0}, t^n_{k_1}) = R_F(t^n_{k_0}, t^n_{k_1})$, and therefore we can use (2.24), the assumption on R_F, and the fact that $1/r = 1/p + 1/q$ (which is needed to treat the last two terms on the right hand side), to obtain

$$
\|\tilde{R}^n_{\tilde{F}^n}\|_{r\text{-var}} \lesssim_r \|R_F\|_{r\text{-var}} + \|F'\|_{q\text{-var}} + \|S\|_{p\text{-var}}.
$$

On the other side, since F and R_F are continuous, $(\tilde{F}^n, \tilde{R}^n_{\tilde{F}^n})$ converges uniformly to (F, R_F). Now for continuous functions, uniform convergence with uniformly bounded p-variation implies convergence in p'-variation for every $p' > p$. See Exercise 2.8 in [FH14] for the case of Hölder continuous functions.

Thus, using Lemma 2.3.18, we see that if $p' > p$ and $q' > q$ are such that $2/p' + 1/q' > 0$, then $((\tilde{S}^n, \tilde{A}^n)_n)$ converges in $(p', p'/2)$-variation to (S, A°), where $A^\circ(s,t) = A(s,t) + 1/2 \langle S \rangle_{s,t}$. Similarly, $((\tilde{F}^n, F', \tilde{R}^n_{\tilde{F}^n}))$ converges in (q', p', r')-variation to (F, F', R_F), where $1/r' = 1/p' + 1/q'$.

Proposition 2.3.11 now yields the uniform convergence of $\int \tilde{F}^n \,\mathrm{d}\tilde{S}^n$ to $\int F \circ \mathrm{d}S$, by which we denote the rough path integral of the controlled path (F, F') against the rough path (S, A°). But for every $t \in [0,T]$ we have

$$
\begin{aligned}
\lim_{n\to\infty} \int_0^t \tilde{F}^n_s \,\mathrm{d}\tilde{S}^n_s &= \lim_{n\to\infty} \sum_{k: t^n_{k+1} \leq t} \frac{1}{2} (F_{t^n_k} + F_{t^n_{k+1}}) S_{t^n_k, t^n_{k+1}} \\
&= \lim_{n\to\infty} \Big(\sum_{k: t^n_{k+1} \leq t} F_{t^n_k} S_{t^n_k, t^n_{k+1}} + \frac{1}{2} \sum_{k: t^n_{k+1} \leq t} F_{t^n_k, t^n_{k+1}} S_{t^n_k, t^n_{k+1}} \Big).
\end{aligned}
$$

Using that F is controlled by S, it is easy to see that the second term on the right hand side converges uniformly to $1/2 \int_0^t F'_s \,\mathrm{d}\langle S \rangle_s$, $t \in [0,T]$. Thus, the Riemann sums $\sum_k F_{t^n_k} S_{t^n_k \wedge \cdot, t^n_{k+1} \wedge \cdot}$ converge uniformly to $\int F \circ \mathrm{d}S - 1/2 \int F' \,\mathrm{d}\langle S \rangle$, and from the representation of the rough path integral as limit of compensated Riemann sums (2.13), it is easy to see that $\int F \circ \mathrm{d}S = \int F \,\mathrm{d}S + 1/2 \int F' \,\mathrm{d}\langle S \rangle$, which completes the proof. $\square$

Remark 2.3.20. *Given Theorem 2.3.19 it is natural to conjecture that if* (S, A) *is the rough path which we constructed in Theorem 2.3.12 and Lemma 2.3.14, then for typical price paths and for adapted, controlled, and continuous integrands* F *the rough path integral agrees with the model free integral of Section 2.2. This seems not very easy to show, but what can be verified is that if* $F \in C^{1+\varepsilon}$*, then for the integrand* $F(S)$ *both integrals coincide – simply take Riemann sums along the dyadic stopping times defined in* (2.7).

Theorem 2.3.19 is reminiscent of Föllmer's pathwise Itô integral [Fö81]. Föllmer assumes that the quadratic variation $\langle S\rangle$ of S exists along a given sequence of partitions and is continuous, and uses this to prove an Itô formula for S: if $F \in C^2$, then

$$F(S_t) = F(S_0) + \int_0^t \nabla F(S_s)\, \mathrm{d}S_s + \frac{1}{2}\int_0^t \mathrm{D}^2 F(S_s)\, \mathrm{d}\langle S\rangle_s, \qquad (2.25)$$

where the integral $\int_0^\cdot \nabla F(S_s)\, \mathrm{d}S_s$ is given as limit of Riemann sums along that same sequence of partitions. Friz and Hairer [FH14] observe that if for $p \in (2,3)$ the function S is of finite p-variation and $\langle S\rangle$ is an arbitrary continuous function of finite $p/2$-variation, then setting

$$\mathrm{Sym}(A)(s,t) := \frac{1}{2}(S_{s,t}S_{s,t} + \langle S\rangle_{s,t})$$

one obtains a "reduced rough path" $(S, \mathrm{Sym}(A))$. They continue to show that if F is controlled by S with *symmetric* derivative F', then it is possible to define the rough path integral $\int F\, \mathrm{d}S$. This is not surprising since then we have $F'_s A_{s,t} = F'_s \mathrm{Sym}(A)_{s,t}$ for the compensator term in the definition of the rough path integral. They also derive an Itô formula for reduced rough paths, which takes the same form as (2.25), except that now $\int \nabla F(S)\, \mathrm{d}S$ is a rough path integral (and therefore defined as limit of compensated Riemann sums).

So both the assumption and the result of [FH14] are slightly different from the ones in [Fö81], and while it seems intuitively clear, it is still not shown rigorously that Föllmer's pathwise Itô integral is a special case of the rough path integral. We will now show that Föllmer's result is a special case of Theorem 2.3.19. For that purpose we only need to prove that Föllmer's condition on the convergence of the quadratic variation is a special case of the assumption in Theorem 2.3.19, at least as long as we only need the symmetric part of the area.

Definition 2.3.21. Let $f \in C([0,T], \mathbb{R})$ and let $\pi^n = \{0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T\}$, $n \in \mathbb{N}$ be such that $\sup\{|f_{t_k^n, t_{k+1}^n}| : k = 0, \ldots, N_n - 1\}$ converges to 0. We say that f has *quadratic variation along* (π^n) *in the sense of Föllmer* if the sequence of discrete measures (μ^n) on $([0,T], \mathcal{B}[0,T])$, defined by

$$\mu_n := \sum_{k=0}^{N_n - 1} |f_{t_k^n, t_{k+1}^n}|^2 \delta_{t_k^n}, \qquad (2.26)$$

converges weakly to a non-atomic measure μ. We write $[f]_t$ for the "distribution function" of μ (in general μ will not be a probability measure). The function $f =$

$(f^1, \dots, f^d) \in C([0,T], \mathbb{R}^d)$ has *quadratic variation along* (π^n) *in the sense of Föllmer* if this holds for all f^i and $f^i + f^j$, $1 \le i, j \le d$. In this case, we set

$$[f^i, f^j]_t := \frac{1}{2}([f^i + f^j]_t - [f^i]_t - [f^j]_t), \quad t \in [0,T].$$

Lemma 2.3.22 (see also [Vov11a], Proposition 6.1)**.** *Let* $p \in (2,3)$*, and let* $S = (S^1, \dots, S^d) \in C([0,T], \mathbb{R}^d)$ *have finite* p*-variation. Let* $\pi^n = \{0 = t_0^n < t_1^n < \dots < t_{N_n}^n = T\}$*,* $n \in \mathbb{N}$*, be a sequence of partitions such that* $\sup\{|S_{t_k^n, t_{k+1}^n}| : k = 0, \dots, N_n - 1\}$ *converges to 0. Then the following conditions are equivalent:*

(i) The function S *has quadratic variation along* (π^n) *in the sense of Föllmer.*

(ii) For all $1 \le i, j \le d$*, the discrete quadratic variation*

$$\langle S^i, S^j \rangle_t^n := \sum_{k=0}^{N_n - 1} S^i_{t_k^n \wedge t, t_{k+1}^n \wedge t} S^j_{t_k^n \wedge t, t_{k+1}^n \wedge t}$$

converges uniformly in $C([0,T], \mathbb{R})$ *to a limit* $\langle S^i, S^j \rangle$*.*

(iii) For $S^{n,i} := \sum_{k=0}^{N_n-1} S^i_{t_k^n} \mathbf{1}_{[t_k^n, t_{k+1}^n)}$*,* $i \in \{1, \dots, d\}$*,* $n \in \mathbb{N}$*, the Riemann sums* $(S^{n,i} \cdot S^j) + (S^{n,j} \cdot S^i)$ *converge uniformly to a limit* $\int S^i \,\mathrm{d}S^j + \int S^j \,\mathrm{d}S^i$*. Moreover, the symmetric part of the approximate area,*

$$\mathrm{Sym}(A^n)^{i,j}(s,t) = \frac{1}{2}\big((S^{n,i} \cdot S^j)_{s,t} + (S^{n,j} \cdot S^i)_{s,t} - S^i_s S^j_{s,t} - S^j_s S^i_{s,t}\big), (s,t) \in \Delta_T,$$

for $1 \le i, j \le d$*, has uniformly bounded* $p/2$*-variation along* (π^n)*, in the sense of* (2.19)*.*

If these conditions hold, then $[S^i, S^j] = \langle S^i, S^j \rangle$ *for all* $1 \le i, j \le d$*.*

Proof. Assume (i) and note that

$$S^i_{t_k^n \wedge t, t_{k+1}^n \wedge t} S^j_{t_k^n \wedge t, t_{k+1}^n \wedge t} = \frac{1}{2}\big(((S^i + S^j)_{t_k^n \wedge t, t_{k+1}^n \wedge t})^2 - (S^i_{t_k^n \wedge t, t_{k+1}^n \wedge t})^2 - (S^j_{t_k^n \wedge t, t_{k+1}^n \wedge t})^2\big).$$

Thus, the uniform convergence of $\langle S^i, S^j \rangle^n$ and the fact that $\langle S^i, S^j \rangle = [S^i, S^j]$ follow once we show that Föllmer's weak convergence of the measures (2.26) implies the uniform convergence of their distribution functions. But since the limiting distribution is continuous by assumption, this is a standard result.

Next, assume (ii) The uniform convergence of the Riemann sums $(S^{n,i} \cdot S^j) + (S^{n,j} \cdot S^i)$ is shown as in Lemma 2.3.17. To see that $\mathrm{Sym}(A^n)$ has uniformly bounded $p/2$-variation along (π^n), note that for $0 \le k \le \ell \le N_n$ and $1 \le i, j \le d$ we have

$$\begin{aligned} |(S^{n,i} \cdot S^j)_{t_k^n, t_\ell^n} + (S^{n,j} \cdot S^i)_{t_k^n, t_\ell^n} - S^i_s S^j_{t_k^n, t_\ell^n} - S^j_s S^i_{t_k^n, t_\ell^n}|^{p/2} \\ = |S^i_{t_k^n, t_\ell^n} S^j_{t_k^n, t_\ell^n} - \langle S^i, S^j \rangle^n_{t_k^n, t_\ell^n}|^{p/2} \\ \le \|S\|_{p\text{-var},[t_k^n, t_\ell^n]} + \|\langle S^i, S^j \rangle^n\|_{1\text{-var},[t_k^n, t_\ell^n]}. \end{aligned}$$

That $\|\langle S^i, S^j \rangle^n\|_{1\text{-var}}$ is uniformly bounded in n is shown in Lemma 2.3.17.

That (iii) implies (i) is also shown in Lemma 2.3.17. $\square$

Remark 2.3.23. *With Theorem 2.3.19 we can only derive an Itô formula for* $F \in C^{2+\varepsilon}$*, since we are only able to integrate* $\nabla F(S)$ *if* $\nabla F \in C^{1+\varepsilon}$*. But this only seems to be due to the fact that our analysis is not sharp. We expect that typical price paths have an associated rough path of finite* 2*-variation, up to logarithmic corrections. For such rough paths, the integral extends to integrands* $F \in C^1$*, see Chapter 10.5 of [FV10b]. For typical price paths (but not for the area), it is shown in [Vov12], Section 4.3, that they are of finite* 2*-variation up to logarithmic corrections.*

3. Local times for typical price paths and pathwise Tanaka formulas

This chapter uses Vovk's [Vov12] game-theoretic approach to mathematical finance to construct local times for "typical price paths". Vovk's approach is based on an outer measure, which is given by the cheapest pathwise superhedging price, and it does not presume any probabilistic structure.

In the last chapter we proved that in a multidimensional setting every typical price path has a natural Itô rough path in the sense of Lyons [Lyo98] associated to it. Based on this, we set up a theory of pathwise integration which was motivated by possible applications in model free financial mathematics. With the techniques of Chapter 2, we are able to treat integrands that are not necessarily functions of the integrator. But if we want to construct $\int f(S)\,\mathrm{d}S$, then we need $f \in C^{1+\varepsilon}$. The aim of the current chapter is to show that for one-dimensional price processes this assumption can be essentially relaxed.

We define discrete versions of the local time and prove that outside a set of outer measure zero they converge to a continuous limit. Roughly speaking, this means that it should be possible to make an arbitrarily large profit by investing in those paths where the convergence of the discrete local times fails. A nice consequence is that the convergence takes place quasi surely under all semimartingale measures for which the coordinate process satisfies the classical condition of "no arbitrage opportunities of the first kind", i.e. for which the drift has a square integrable density with respect to the quadratic variation of the local martingale part.

Using these pathwise local times, we derive various pathwise change of variable formulas which generalize Föllmer's pathwise Itô formula [Fö81] in the same way that the classical Tanaka formula generalizes the classical Itô formula. In particular, we can integrate $f(S)$ against a typical price path S whenever f has finite q-variation for some $q < 2$.

For a more detailed discussion about pathwise integration in mathematical finance we refer back to Chapter 2. However, for the present chapter some additional motivation comes amongst others from [DOR14], where pathwise local times and a pathwise generalized Itô formula are used to derive arbitrage free price bounds for weighted variance swaps in a model free setting. The techniques of [DOR14] allow to handle integrands in the Sobolev space H^1. Here we extend this to not necessarily continuous integrands of finite q-variation for some $q < 2$. Further motivations can be found in the survey paper [FS13] which emphasizes possible applications of pathwise integration to robust hedging problems, or in [CJ90] and [Son06], where local times appear naturally in a financial context and are used to resolve the so-called "stop-loss start-gain paradox".

This chapter is organized as follows: In Section 3.1 we present various extensions of Föllmer's pathwise Itô formula under suitable assumptions on the local time. In Section 3.2 we show that typical price paths possess local times which satisfy all the assumptions of Section 3.1.

3.1. Pathwise Tanaka formulas

A first non-probabilistic approach to stochastic calculus was introduced by Föllmer in [Fö81], where an Itô formula was developed for a class of real-valued functions with quadratic variation. This builds our starting point for a pathwise version of Tanaka's formula and a generalized Itô formula, respectively. Let us start by recalling Föllmer's definition of quadratic variation.

A *partition* π is an increasing sequence $0 = t_0 < t_1 < \ldots$ without accumulation points, possibly taking the value ∞. For $T > 0$ we denote by $\pi[0,T] := \{t_j : t_j \in [0,T)\} \cup \{T\}$ the partition π restricted to $[0,T]$, and if $S\colon [0,\infty) \to \mathbb{R}$ is a continuous function we write

$$m(S, \pi[0,T]) := \max_{t_j \in \pi[0,T] \setminus \{t_0\}} |S(t_j) - S(t_{j-1})|$$

for the mesh size of π along S on the interval $[0,T]$. We denote by $\mathcal{B}([0,\infty))$ the Borel σ-algebra on $[0,\infty)$.

Definition 3.1.1. Let (π^n) be a sequence of partitions and let $S \in C([0,\infty),\mathbb{R})$ be such that $\lim_{n\to\infty} m(S,\pi^n[0,T]) = 0$ for all $T > 0$. We say that S has *quadratic variation* along (π^n) if the sequence of measures

$$\mu_n := \sum_{t_j \in \pi^n \setminus \{\infty\}} (S(t_{j+1}) - S(t_j))^2 \delta_{t_j}, \qquad n \in \mathbb{N},$$

on $([0,\infty), \mathcal{B}([0,\infty)))$ converges vaguely to a nonnegative Radon measure μ without atoms, where δ_t denotes the Dirac measure at $t \in [0,\infty)$. We write $\langle S\rangle(t) := \mu([0,t])$ for the continuous "distribution function" of μ and $\mathcal{Q}(\pi^n)$ for the set of all continuous functions having quadratic variation along (π^n).

The reason for only requiring $\lim_n m(S,\pi^n[0,T]) = 0$ rather than assuming that the mesh size of (π^n) goes to zero is that later we will work with Lebesgue partitions and paths with piecewise constant parts, in which case only the first assumption holds.

We stress the fact that $\mathcal{Q}(\pi^n)$ depends on the sequence (π^n) and that for a given path the quadratic variation along two different sequences of partitions can be different, even if both exist. This is very unpleasant and might lead the reader to question the usefulness of our results. But quite remarkably there is a large class of paths which have a natural pathwise quadratic variation that is independent of the specific partition used to calculate it. More precisely, in the master's thesis [Lem83], see also [CLPT81], the notion of *quadratic arc length* is introduced. Roughly speaking, a path S has quadratic arc length A if the quadratic variation of S along *any sequence of Lebesgue partitions* is equal to A. It is shown in [Lem83], Theorem III.3.3, that almost every sample path $S(\omega)$ of a continuous semimartingale has a quadratic arc length which is equal to the semimartingale quadratic variation $\langle S\rangle(\omega)$. The same theorem also shows that almost every sample path of a continuous semimartingale has a natural local time which can be obtained by counting interval upcrossings.

For $k \in \mathbb{N}$ let us write $C^k = C^k(\mathbb{R}, \mathbb{R})$ for the space of k times continuously differentiable functions, and $C^k_b = C^k_b(\mathbb{R}, \mathbb{R})$ for the space of functions in C^k that are bounded with bounded derivatives, equipped with the usual norm $\|\cdot\|_{C^k_b}$.

Theorem 3.1.2 ([Fö81])**.** *Let (π^n) be a sequence of partitions and let $S \in \mathcal{Q}(\pi^n)$ and $f \in C^2$. Then the pathwise Itô formula*

$$f(S(t)) = f(S(0)) + \int_0^t f'(S(s))\,\mathrm{d}S(s) + \frac{1}{2}\int_0^t f''(S(s))\,\mathrm{d}\langle S\rangle(s)$$

holds with

$$\int_0^t f'(S(s))\,\mathrm{d}S(s) := \lim_{n\to\infty} \sum_{t_j\in\pi^n} f'(S(t_j))(S(t_{j+1}\wedge t) - S(t_j \wedge t)), \quad t \in [0,\infty), \tag{3.1}$$

where the series in (3.1) *is absolutely convergent.*

In particular, the integral $\int_0^\cdot g(S(s))\,\mathrm{d}S(s)$ is defined for all $g \in C^1$, and for all $T > 0$ the map $C^1_b \ni g \mapsto \int_0^\cdot g(S(s))\,\mathrm{d}S(s) \in C([0,T],\mathbb{R})$ defines a bounded linear operator and we have

$$\Big|\int_0^t g(S(s))\,\mathrm{d}S(s)\Big| \le |S(t) - S(0)| \times \|g\|_{L^\infty(\mathrm{supp}(S|_{[0,t]}))} + \frac{1}{2}\langle S\rangle(t)\|g'\|_{L^\infty(\mathrm{supp}(S|_{[0,t]}))}$$

for all $t \ge 0$, where $\mathrm{supp}(S|_{[0,t]}])$ *denotes the support of S restricted to the interval* $[0,t]$.

Föllmer actually requires the mesh size $\max_{t_j\in\pi^n\setminus\{t_0\},\, t_j\le T} |t_j - t_{j-1}|$ to converge to zero for all $T > 0$, but he also considers càdlàg functions S. For continuous S, the proof only uses that $m(S, \pi^n[0,T])$ converges to zero.

The continuity of the Itô integral is among its most important properties: if we approximate the integrand in a suitable topology, then the approximate integrals converge in probability to the correct limit. This is absolutely crucial in applications, for example when solving stochastic optimization problems or SDEs. Here we are arguing for one fixed path, so the statement in Theorem 3.1.2 is a natural formulation of the continuity properties in our context.

In the theory of continuous semimartingales, Itô's formula can be extended further to a generalized Itô rule for convex functions, see for instance Theorem 6.22 in [KS88]. In the spirit of Föllmer, a generalized Itô rule for functions in suitable Sobolev spaces was derived in the unpublished diploma thesis of Wuermli [Wue80]. We briefly recall here the idea for this pathwise version as presented in [Wue80] or [DOR14].

Let f' be right-continuous and of locally bounded variation, and we set $f(x) := \int_{(0,x]} f'(y)\,\mathrm{d}y$ for $x \ge 0$ and $f(x) := -\int_{(x,0]} f'(y)\,\mathrm{d}y$ for $x < 0$. Then we get for $b \ge a$ that

$$f(b) - f(a) = f'(a)(b-a) + \int_{(a,b]} (f'(x) - f'(a))\,\mathrm{d}x = f'(a)(b-a) + \int_{(a,b]} (b-t)\,\mathrm{d}f'(t),$$

where we used integration by parts, and where the integral on the right hand side is to be understood in the Riemann-Stieltjes sense. For $b < a$, we get $f(b) - f(a) = f'(a)(b-a) + \int_{(b,a]}(t-b)\,\mathrm{d}f'(t)$. Therefore, for any $S \in C([0,\infty),\mathbb{R})$ and any partition π we have

$$f(S(t)) - f(S(0)) = \sum_{t_j \in \pi} f'(S(t_j \wedge t))(S(t_{j+1} \wedge t) - S(t_j \wedge t)) + \int_{-\infty}^{\infty} \Big(\sum_{t_j \in \pi} \mathbf{1}_{(\!| S(t_j \wedge t), S(t_{j+1} \wedge t)]\!]}(u) |S(t_{j+1} \wedge t) - u| \Big) \,\mathrm{d}f'(u), \tag{3.2}$$

where we used the notation

$$(\!| u, v]\!] := \begin{cases} (u,v], & \text{if } u \leq v, \\ (v,u], & \text{if } u > v, \end{cases} \tag{3.3}$$

for $u, v \in \mathbb{R}$. Let us define a discrete local time by setting

$$L_t^\pi(S,u) := \sum_{t_j \in \pi} \mathbf{1}_{(\!| S(t_j \wedge t), S(t_{j+1} \wedge t)]\!]}(u) |S(t_{j+1} \wedge t) - u|, \quad u \in \mathbb{R},$$

and note that $L_t^\pi(S,u) = 0$ for $u \notin [\inf_{s \in [0,t]} S(s), \sup_{s \in [0,t]} S(s)]$. In the following we may omit the S and just write $L_t^\pi(u)$.

Definition 3.1.3. Let (π^n) be a sequence of partitions and let $S \in C([0,\infty),\mathbb{R})$. A function $L(S)\colon [0,\infty) \times \mathbb{R} \to \mathbb{R}$ is called L^2-*local time* of S along (π^n) if for all $t \in [0,\infty)$ it holds $\lim_{n\to\infty} m(S, \pi^n[0,t]) = 0$ and the discrete pathwise local times $L_t^{\pi^n}(S,\cdot)$ converge weakly in $L^2(\mathrm{d}u)$ to $L_t(S,\cdot)$ as $n \to \infty$. We write $\mathcal{L}_{L^2}(\pi^n)$ for the set of all continuous functions having an L^2-local time along (π^n).

Using this definition of the local time, Wuermli showed the following theorem, where we denote by $H^k = H^k(\mathbb{R},\mathbb{R})$ the Sobolev space of functions which are k times weakly differentiable in $L^2(\mathbb{R},\mathbb{R})$.

Theorem 3.1.4 ([Wue80], Satz 9 or [DOR14], Proposition B.4)**.** *Let (π^n) be a sequence of partitions and let $S \in \mathcal{L}_{L^2}(\pi^n)$. Then $S \in \mathcal{Q}(\pi^n)$, and for every $f \in H^2$ the generalized pathwise Itô formula*

$$f(S(t)) = f(S(0)) + \int_0^t f'(S(s))\,\mathrm{d}S(s) + \int_{-\infty}^{\infty} f''(u) L_t(S,u)\,\mathrm{d}u$$

holds with

$$\int_0^t f'(S(s))\,\mathrm{d}S(s) := \lim_{n\to\infty} \sum_{t_j \in \pi^n} f'(S(t_j))(S(t_{j+1} \wedge t) - S(t_j \wedge t)), \quad t \in [0,\infty).$$

(Note that f' is continuous for $f \in H^2$). In particular, the integral $\int_0^\cdot g(S(s))\,\mathrm{d}S(s)$ is defined for all $g \in H^1$, and for all $T > 0$, the map $H^1 \ni g \mapsto \int_0^\cdot g(S(s))\,\mathrm{d}S(s) \in$

$C([0,T],\mathbb{R})$ defines a bounded linear operator. Moreover, for $A \in \mathcal{B}(\mathbb{R})$ we have the occupation density formula

$$\int_A L_t(u)\,\mathrm{d}u = \frac{1}{2}\int_0^t \mathbf{1}_A(S(s))\,\mathrm{d}\langle S\rangle(s), \quad t \in [0,\infty).$$

In other words, for all $t \geq 0$ the occupation measure of S on $[0,t]$ is absolutely continuous with respect to the Lebesgue measure, with density $2L_t$.

Sketch of proof. Formula (3.2) in combination with the continuity of f and S yields

$$\begin{aligned} f(S(t)) - f(S(0)) = &\sum_{t_j\in\pi^n} f'(S(t_j))(S(t_{j+1}\wedge t) - S(t_j\wedge t)) \\ &+ \int_{-\infty}^{\infty}\Big(\sum_{t_j\in\pi^n} \mathbf{1}_{⟦S(t_j\wedge t),S(t_{j+1}\wedge t)⟧}(u)|S(t_{j+1}\wedge t)-u|\Big) f''(u)\,\mathrm{d}u. \end{aligned}$$

By assumption, the second term on the right hand side converges to

$$\int_{-\infty}^{\infty} f''(u)L_t(S,u)\,\mathrm{d}u$$

as n tends to ∞, so that also the Riemann sums have to converge.

The occupation density formula follows by approximating $\mathbf{1}_A$ with continuous functions. □

As already observed by Bertoin [Ber87], the key point of this extension of Föllmer's pathwise stochastic integral is again that it is given by a *continuous* linear operator on H^1. Since $L_t(S,\cdot)$ is compactly supported for all $t \geq 0$, the same arguments also work for functions f that are locally in H^2, i.e. such that $f|_{(a,b)} \in H^2((a,b),\mathbb{R})$ for all $-\infty < a < b < \infty$.

As we make stronger assumptions on the local times $L(S)$, it is natural to expect that we can extend Wuermli's generalized Itô formula to larger spaces of functions.

Definition 3.1.5. Let (π^n) be a sequence of partitions and let $S \in \mathcal{L}_{L^2}(\pi^n)$. We say that S has a *continuous local time* along (π^n) if for all $t \in [0,\infty)$ the discrete pathwise local times $L_t^{\pi^n}(S,\cdot)$ converge uniformly to a continuous limit $L_t(S,\cdot)$ as $n \to \infty$ and if $(t,u) \mapsto L_t(S,u)$ is jointly continuous. We write $\mathcal{L}_c(\pi^n)$ for the set of all S having a continuous local time along (π^n).

In the following theorem, $\mathrm{BV} = \mathrm{BV}(\mathbb{R},\mathbb{R})$ denotes the space of right-continuous bounded variation functions, equipped with the total variation norm.

Theorem 3.1.6. *Let (π^n) be a sequence of partitions and let $S \in \mathcal{L}_c(\pi^n)$. Let $f\colon \mathbb{R} \to \mathbb{R}$ be absolutely continuous with right-continuous Radon-Nikodym derivative f' of locally bounded variation. Then we have the generalized change of variable formula*

$$f(S(t)) = f(S(0)) + \int_0^t f'(S(s))\,\mathrm{d}S(s) + \int_{-\infty}^{\infty} L_t(u)\,\mathrm{d}f'(u)$$

for all $t \in [0,\infty)$, where

$$\int_0^t f'(S(s))\,\mathrm{d}S(s) := \lim_{n\to\infty} \sum_{t_j\in\pi^n} f'(S(t_j))(S(t_{j+1}\wedge t) - S(t_j \wedge t)), \quad t \in [0,\infty). \tag{3.4}$$

In particular, the integral $\int_0^\cdot g(S(s))\,\mathrm{d}S(s)$ is defined for all g of locally bounded variation, and for all $T > 0$ the map $\mathrm{BV} \ni g \mapsto \int_0^\cdot g(S(s))\,\mathrm{d}S(s) \in C([0,T],\mathbb{R})$ defines a bounded linear operator.

Proof. From (3.2) we get

$$f(S(t)) - f(S(0)) = \sum_{t_j\in\pi^n} f'(S(t_j))(S(t_{j+1}\wedge t) - S(t_j\wedge t)) + \int_{-\infty}^{\infty} L_t^{\pi^n}(u)\,\mathrm{d}f'(u)$$

for all $t \ge 0$. Since $L_t^{\pi^n}$ converges uniformly to L_t, our claim immediately follows. □

Observe that f satisfies the assumptions of Theorem 3.1.6 if and only if it is the difference of two convex functions. For such f, Sottinen and Viitasaari [SV14] prove a generalized change of variable formula for a class of Gaussian processes. They make the very nice observation that for a suitable Gaussian process X one can control the fractional Besov regularity of $f'(X)$, and they use this insight to construct $\int_0^\cdot f'(X_t)\,\mathrm{d}X_t$ as a fractional integral. Such a regularity result is somewhat surprising since in general $f'(X)$ is not even làdlàg, so in particular not of finite p-variation for any $p > 0$. But since regularity of $f'(X)$ is shown using probabilistic arguments, the integral of Sottinen and Viitasaari is not directly a pathwise object: the null set outside of which it exists may depend on f. Moreover, they can only handle Gaussian processes that are Hölder continuous of order $\alpha > 1/2$, and their approach breaks down when considering processes with non-trivial quadratic variation. Here we have a completely different focus, since we are interested in pathwise results for paths with non-trivial quadratic variation.

As an immediate consequence of Theorem 3.1.6 we obtain a pathwise version of the classical Tanaka formula.

Corollary 3.1.7. *Let (π^n) be a sequence of partitions and let $S \in \mathcal{L}_c(\pi^n)$. The pathwise Tanaka-Meyer formula*

$$L_t(u) = (S(t)-u)^- - (S(0)-u)^- + \int_0^t \mathbf{1}_{(-\infty,u)}(S(s))\,\mathrm{d}S(s)$$

is valid for all $(t,u) \in [0,\infty)\times\mathbb{R}$, with the notation $(\cdot - u)^- := \max\{0, u - \cdot\}$. The analogous formulas for $\mathbf{1}_{[u,\infty)}(\cdot)$ and $\operatorname{sgn}(\cdot - u)$ hold as well.

At this point we see a picture emerge: the more regularity the local time has, the larger the space of functions is to which we can extend our pathwise stochastic integral. Indeed, the previous examples are all based on duality between the derivative of the integrand and the occupation measure. In the classical Föllmer-Itô case and for fixed time $T \ge 0$, the occupation measure is just a finite measure on a compact interval $[a,b]$, and certainly the continuous functions belong to the dual space of the finite measures on $[a,b]$. In the Wuermli setting, the occupation measure has a density in L^2 and therefore defines a bounded functional on L^2. If the local time is continuous, then we can even integrate Radon measures against it.

So if we can quantify the continuity of the local time, then the dual space further increases and we can extend the pathwise Itô formula to a bigger class of functions. To this end we introduce for a given sequence of partitions (π^n) and $p \geq 1$ the set $\mathcal{L}_{c,p}(\pi^n) \subseteq \mathcal{L}_c(\pi^n)$ consisting of those $S \in \mathcal{L}_c(\pi^n)$ for which the discrete local times $(L_t^{\pi^n})$ have uniformly bounded p-variation, uniformly in $t \in [0,T]$ for all $T > 0$, i.e. for which

$$\sup_{n\in\mathbb{N}} \|L^{\pi^n}\|_{C_T\mathcal{V}^p} := \sup_{n\in\mathbb{N}} \sup_{t\in[0,T]} \|L_t^{\pi^n}(\cdot)\|_{p\text{-var}} < \infty$$

for all $T > 0$, where we write for any $f : \mathbb{R} \to \mathbb{R}$

$$\|f\|_{p\text{-var}} := \sup\Big\{\Big(\sum_{k=1}^{n} |f(u_k) - f(u_{k-1})|^p\Big)^{1/p} : -\infty < u_0 < \ldots < u_n < \infty,\, n \in \mathbb{N}\Big\}.$$

We also write $\mathcal{V}^p$ for the space of right-continuous functions of finite p-variation, equipped with the maximum of the p-variation seminorm and the supremum norm.

For $S \in \mathcal{L}_{c,p}(\pi^n)$ and using the Young integral it is possible to extend the pathwise Tanaka formula to an even larger class of integrands, allowing us to integrate $\int g(S)\,\mathrm{d}S$ provided that g has finite q-variation for some q with $1/p+1/q > 1$. This is similar in spirit to the Bouleau-Yor [BY81] extension of the classical Tanaka formula. Such an extension was previously derived by Feng and Zhao [FZ06], Theorem 2.2. But Feng and Zhao stay in a semimartingale setting, and they interpret the stochastic integral appearing in (3.6) as a usual Itô integral. Here we obtain a pathwise integral, which is given very naturally as a limit of Riemann sums.

Let us briefly recall the main concepts of Young integration. In [You36], Young showed that if $-\infty < a < b < \infty$, if f and g are two functions on $[a,b]$ of finite p- and q-variation respectively with $1/p + 1/q > 1$, and if π is a partition of $[a,b]$, then there exists a universal constant $C(p,q) > 0$ such that

$$\Big|\sum_{t_j,t_{j+1}\in\pi} f(t_j)(g(t_{j+1}) - g(t_j))\Big| \leq C(p,q)\|f\|_{p\text{-var},[a,b]}\|g\|_{q\text{-var},[a,b]},$$

where we wrote $\|f\|_{p\text{-var},[a,b]}$ for the p-variation of f on $[a,b]$ and similarly for g. In particular, if there exists a sequence of partitions (π^n) and if the Riemann sums of f against g along (π^n) converge to a limit which we denote by $\int_a^b f(s)\,\mathrm{d}g(s)$, then

$$\Big|\int_a^b f(s)\,\mathrm{d}g(s)\Big| \leq C(p,q)(|f(a)| + \|f\|_{p\text{-var},[a,b]})\|g\|_{q\text{-var},[a,b]}. \tag{3.5}$$

Moreover, Young showed that if f and g have no common points of discontinuity, then the Riemann sums along any sequence of partitions with mesh size going to zero converge to the same limit $\int_0^t f(s)\,\mathrm{d}g(s)$, independently of the specific sequence of partitions.

We therefore easily obtain the following theorem.

Theorem 3.1.8 (see also [FZ06], Theorem 2.2)**.** *Let $p, q \geq 1$ be such that $\frac{1}{p} + \frac{1}{q} > 1$. Let (π^n) be a sequence of partitions and let $S \in \mathcal{L}_{c,p}(\pi^n)$. Let $f\colon \mathbb{R} \to \mathbb{R}$ be absolutely continuous with right-continuous Radon-Nikodym derivative f' of locally finite q-variation. Then for all $t \in [0, \infty)$ the generalized change of variable formula*

$$f(S(t)) = f(S(0)) + \int_0^t f'(S(s))\,\mathrm{d}S(s) + \int_{-\infty}^{\infty} L_t(u)\,\mathrm{d}f'(u) \tag{3.6}$$

holds, where $\mathrm{d}f'(u)$ denotes Young integration and where

$$\int_0^t f'(S(s))\,\mathrm{d}S(s) := \lim_{n\to\infty} \sum_{t_j \in \pi^n} f'(S(t_j))(S(t_{j+1} \wedge t) - S(t_j \wedge t)), \quad t \in [0, \infty).$$

In particular, the integral $\int_0^\cdot g(S(s))\,\mathrm{d}S(s)$ is defined for all right-continuous g of locally finite q-variation, and for all $T > 0$ the map $\mathcal{V}^q \ni g \mapsto \int_0^\cdot g(S(s))\,\mathrm{d}S(s) \in C([0, T], \mathbb{R})$ defines a bounded linear operator.

Proof. Observe that for each $n \in \mathbb{N}$, the discrete local time $L_t^{\pi^n}$ is piecewise smooth and of bounded variation. Therefore, formula (3.2) holds for $L_t^{\pi^n}$ and f', and the integral on the right hand side of (3.2) is given as the limit of Riemann sums along an arbitrary sequence of partitions with mesh size going to zero – provided that every element of the sequence contains all jump points of $L_t^{\pi^n}$. Therefore, the integral must satisfy the bound (3.5). Since the p-variation of $(L_t^{\pi^n})$ is uniformly bounded, and the sequence converges uniformly to L_t, it is easy to see that it must converge in p'-variation for all $p' > p$. Choosing such a p' with $1/q + 1/p' > 1$ and combining the linearity of the Young integral with the bound (3.5), the result follows. □

Remark 3.1.9. *Theorem 2.2 in [FZ06] states* (3.6) *under the slightly weaker assumption that $f\colon \mathbb{R} \to \mathbb{R}$ is left-continuous and locally bounded with left-continuous and locally bounded left derivative $\mathrm{D}^- f$ of finite q-variation. But absolute continuity of f is clearly necessary: Consider the path $S(t) \equiv t$ for $t \in [0, \infty)$, for which $\langle S \rangle \equiv 0$ and thus $L \equiv 0$. In this case equation* (3.6) *would read*

$$f(t) = f(0) + \int_0^t \mathrm{D}^- f(u)\,\mathrm{d}u, \quad t \in [0, \infty),$$

a contradiction if f is not absolutely continuous.

In the following, we will show that any typical price path which might model an asset price trajectory must be in $\mathcal{L}_{c,p}(\pi^n)$ if (π^n) denotes the dyadic Lebesgue partition generated by S.

3.2. Local times for model free finance

3.2.1. Super-hedging and outer measure

In a recent series of papers [Vov11a, Vov11b, Vov12], Vovk introduced a hedging based, model free approach to mathematical finance. Roughly speaking, Vovk considers the set of real-valued continuous functions as price paths and introduces an outer measure on this set which is given by the cheapest super-hedging price. A property (P) is said to hold for "typical price paths" if it is possible to make an arbitrarily large profit by investing in the paths where (P) is violated. We will see that in Vovk's framework it is possible to construct continuous local times for typical price paths, which gives an axiomatic justification for the use of our pathwise generalized Itô formulas from Section 3.1 in model free finance. While we worked in Chapter 2 on a finite time horizon and with multidimensional price paths, the price paths are now assumed to be one-dimensional but may live on an infinite time horizon. Let us briefly introduce this slightly modified stetting.

More precisely, we consider the (sample) space $\Omega = C([0,\infty),\mathbb{R})$ of all continuous functions $\omega\colon [0,\infty) \to \mathbb{R}$. The coordinate process on Ω is denoted by $S_t(\omega) := \omega(t)$. For $t \in [0,\infty)$ we define $\mathcal{F}_t := \sigma(S_s : s \leq t)$ and we set $\mathcal{F} := \bigvee_{t\geq 0} \mathcal{F}_t$. Stopping times τ and the associated σ-algebras $\mathcal{F}_\tau$ are defined as usual.

A process $H\colon \Omega \times [0,\infty) \to \mathbb{R}$ is called a *simple strategy* if there exist stopping times $0 = \tau_0(\omega) < \tau_1(\omega) < \dots$ such that for every $\omega \in \Omega$ and every $T \in (0,\infty)$ we have $\tau_n(\omega) \leq T$ for only finitely many n, and $\mathcal{F}_{\tau_n}$-measurable bounded functions $F_n\colon \Omega \to \mathbb{R}$ such that $H_t(\omega) = \sum_{n\geq 0} F_n(\omega)\mathbf{1}_{(\tau_n(\omega),\tau_{n+1}(\omega)]}(t)$. In that case the integral

$$(H\cdot S)_t(\omega) = \sum_{n=0}^{\infty} F_n(\omega)[S_{\tau_{n+1}(\omega)\wedge t} - S_{\tau_n(\omega)\wedge t}]$$

is well defined for every $\omega \in \Omega$ and every $t \in [0,\infty)$.

For $\lambda > 0$ a simple strategy H is called λ-*admissible* if $(H\cdot S)_t(\omega) \geq -\lambda$ for all $t \in [0,\infty)$ and all $\omega \in \Omega$. The set of λ-admissible simple strategies is denoted by $\mathcal{H}_\lambda$.

Definition 3.2.1. The *outer measure* $\overline{P}$ of $A \subseteq \Omega$ is defined as the cheapest super-hedging price for $\mathbf{1}_A$, that is

$$\overline{P}(A) := \inf\Big\{\lambda > 0\ : \exists (H^n)_{n\in\mathbb{N}} \subseteq \mathcal{H}_\lambda \text{ such that } \liminf_{t\to\infty}\liminf_{n\to\infty}(\lambda + (H^n\cdot S)_t(\omega)) \geq \mathbf{1}_A(\omega)\forall\omega\in\Omega\Big\}.$$

A set of paths $A \subseteq \Omega$ is called a *null set* if it has outer measure zero. A property (P) holds for *typical price paths* if the set A where (P) is violated is a null set.

Of course, it would be more natural to minimize over simple trading strategies rather than over the limit inferior along sequences of simple strategies. But then $\overline{P}$ would not be countably subadditive, and this would make it very difficult to work with. Let us just remark that in the classical definition of superhedging prices in semimartingale models we work with general admissible strategies, and the Itô integral against a general strategy is given as limit of integrals against simple strategies. So in that sense our definition is analogous to the classical one (apart from the fact that we do not require convergence and consider the lim inf instead).

For us, the most important property of $\overline{P}$ is the following arbitrage interpretation for null sets.

Lemma 3.2.2 (cf. Lemma 2.1.4). *A set $A \subseteq \Omega$ is a null set if and only if there exists a sequence of 1-admissible simple strategies $(H^n)_{n\in\mathbb{N}} \subseteq \mathcal{H}_1$, such that*

$$\liminf_{t\to\infty} \liminf_{n\to\infty} (1 + (H^n \cdot S)_t(\omega)) \geq \infty \cdot \mathbf{1}_A(\omega),$$

where we set $\infty \cdot 0 = 0$ and $\infty \cdot 1 = \infty$.

In other words, a null set is essentially a model free arbitrage opportunity of the first kind, and to only work with typical price paths is analogous to only considering models which satisfy (NA1) (no arbitrage opportunities of the first kind). The notion (NA1) has raised a lot of interest in recent years since it is the minimal condition which has to be satisfied by any reasonable asset price model; see for example [KK07, Ruf13, IP11]. If $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$, we say that S satisfies *(NA1)* under $\mathbb{P}$ if the set $\mathcal{W}_1^\infty := \{1 + \int_0^\infty H_s \,\mathrm{d}S_s : H \in \mathcal{H}_1\}$ is bounded in probability, that is if $\lim_{n\to\infty} \sup_{X\in\mathcal{W}_1^\infty} \mathbb{P}(X \geq n) = 0$. In the continuous setting this is equivalent to S being a semimartingale of the form $S = M + \int_0^\cdot \alpha_s \,\mathrm{d}\langle M\rangle_s$, where M is a local martingale and $\int_0^\infty \alpha_s^2 \,\mathrm{d}\langle M\rangle_s < \infty$.

In the next proposition we collect further properties of $\overline{P}$. For proofs (in finite time) see Section 2.1.1.

Proposition 3.2.3. *(i) $\overline{P}$ is an outer measure with $\overline{P}(\Omega) = 1$, i.e. $\overline{P}$ is nondecreasing, countably subadditive, and $\overline{P}(\emptyset) = 0$.*

(ii) Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$ such that the coordinate process S is a $\mathbb{P}$-local martingale, and let $A \in \mathcal{F}$. Then $\mathbb{P}(A) \leq \overline{P}(A)$.

(iii) Let $A \in \mathcal{F}$ be a null set, and let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$ such that the coordinate process S satisfies (NA1) under $\mathbb{P}$. Then $\mathbb{P}(A) = 0$.

The last statement says that every property which is satisfied by typical price paths holds quasi-surely for all probability measures which might be of interest in mathematical finance.

Lemma 3.2.2 and Proposition 3.2.3 are originally due to Vovk, but here and in Chapter 2 we consider a small modification of Vovk's outer measure, which in our opinion has a slightly more natural financial interpretation and with which it is easier to work.

3.2.2. Existence of local times for typical price paths

This subsection is devoted to the presentation and the proof of our main result (Theorem 3.2.5): every typical price path has a local time which satisfies all the requirements needed to apply our most general Itô-Tanaka formula, Theorem 3.1.8.

For this purpose recall that for every partition $\pi(\omega) = \{0 = t_0(\omega) < t_1(\omega) < \ldots < t_{K(\omega)}(\omega) < t_{(K+1)(\omega)}(\omega) = \infty\}$ of $[0, \infty)$ a discrete version of the local time is given by

$$L_t^\pi(S, u)(\omega) = \sum_{j=0}^{K(\omega)} \mathbf{1}_{(\!(S_{t_j\wedge t}(\omega), S_{t_{j+1}\wedge t}(\omega)]\!]}(u) |S_{t_{j+1}\wedge t}(\omega) - u|, \quad (t, u) \in [0, \infty) \times \mathbb{R}.$$

From (3.2) we get the following discrete version of Tanaka's formula, which can also be obtained by direct computation:

$$L_t^\pi(S,u)(\omega) = (S_t(\omega)-u)^- - (S_0(\omega)-u)^- + \sum_{j=0}^{K(\omega)} \mathbf{1}_{(-\infty,u)}(S_{t_j}(\omega))[S_{t_{j+1}\wedge t}(\omega) - S_{t_j\wedge t}(\omega)] \tag{3.7}$$

for all $(t,u) \in [0,\infty) \times \mathbb{R}$ and $\omega \in \Omega$. Taking a sequence of partitions with mesh size converging to zero, we see that at least formally the construction of the stochastic integral $\int_0^\cdot \mathbf{1}_{(-\infty,u)}(S_s)\,\mathrm{d}S_s(\omega)$ is equivalent to the construction of the local time $L(S,u)(\omega)$.

In the following we will work with a very natural sequence of partitions, namely the dyadic Lebesgue partitions generated by S: For each $n \in \mathbb{N}$ denote $\mathbb{D}^n := \{k2^{-n} : k \in \mathbb{Z}\}$ and define the sequence of stopping times

$$\tau_0^n(\omega) := 0, \quad \tau_{k+1}^n(\omega) := \inf\{t \geq \tau_k^n(\omega) : S_t(\omega) \in \mathbb{D}^n \setminus S_{\tau_k^n(\omega)}(\omega)\}, \quad k \in \mathbb{N}. \tag{3.8}$$

We set $\pi^n(\omega) := \{0 = \tau_0^n(\omega) < \tau_1^n(\omega) < \dots\}$. Note that the functions $\tau_k^n(\omega)$ are stopping times and that $(\pi^n(\omega))$ is increasing, i.e. it holds $\pi^n(\omega) \subset \pi^{n+1}(\omega)$ for all $n \in \mathbb{N}$. From now on we will mostly omit the ω and just write π^n and τ_k^n instead of $\pi^n(\omega)$ and $\tau_k^n(\omega)$, respectively.

A key ingredient for our construction of the local time is the following analysis of the number of interval crossings. Let $U_t(\omega,a,b)$ be the number of upcrossings of the closed interval $[a,b] \subseteq \mathbb{R}$ by $S(\omega)$ during the time interval $[0,t]$, where an upcrossing is a pair $(u,v) \in [0,t]^2$ with $u < v$ such that $S_u(\omega) = a$, $S_v(\omega) = b$ and $S_w(\omega) \in (a,b)$ for all $w \in (u,v)$. Downcrossings are defined analogously and we write $D_t(\omega,a,b)$ for the number of downcrossings by $\omega \in \Omega$ during the time interval $[0,t]$.

Lemma 3.2.4. *For typical price paths $\omega \in \Omega$, there exists $C(\omega)\colon (0,\infty) \to (0,\infty)$ such that*

$$\max_{k\in\mathbb{Z}} \left(U_T^n(\omega,k2^{-n}) + D_T^n(\omega,k2^{-n})\right) \leq C_T(\omega) n^2 2^n$$

for all $n \in \mathbb{N}$, $T > 0$, where $U_T^n(\omega,u) := U_T(\omega,u,u+2^{-n})$ for $u \in \mathbb{R}$, and similarly for the number of downcrossings.

Proof. Let $K,T > 0$. Without loss of generality we may restrict our considerations to the set $A_K := \{\omega \in \Omega : \sup_{t\in[0,T]} |S_t(\omega)| < K\}$. Let $k \in (-2^nK, 2^nK)$ and write $u = k2^{-n}$. The following strategy will make a large profit if $U_T^n(u) := U_T^n(\omega,u)$ is large: start with wealth 1, when S first hits u buy $1/(2K)$ numbers of shares. When S hits $-K$ sell and stop trading. Otherwise, when S hits $u+2^{-n}$ sell. This gives us wealth $1 + 2^{-n}/(2K)$ on the set $\{U_T^n(u) \geq 1\} \cap A_K$. Now we repeat this strategy: next time we hit u, we buy our current wealth times $1/(2K)$ shares of S, and sell when S hits $u+2^{-n}$ or $-K$. After $n^2 2^n$ upcrossings of $[u,u+2^{-n}]$, stop trading. On the set $\{U_T^n(u) \geq n^2 2^n\} \cap A_K$ we then have a wealth of

$$\left(1 + \frac{2^{-n}}{2K}\right)^{n^2 2^n} \geq \exp\left(\frac{1}{4K} n^2\right)$$

for all n that are large enough. Therefore

$$\bar{P}(\{U_T^n(u) \geq n^2 2^n\} \cap A_K) \leq \exp\Big(-\frac{n^2}{4K}\Big)$$

for all large n. Summing over all dyadic points $u = k2^{-n}$ in $(-K, K)$, we obtain

$$\begin{aligned}\overline{P}\Big(\Big\{\max_{k\in\mathbb{Z}} U_T^n(k2^{-n}) \geq n^2 2^n\Big\} \cap A_K\Big) &\leq K 2^{n+1} \exp\Big(-\frac{n^2}{4K}\Big)\\ &= K \exp\Big(-\frac{n^2}{8K} + (n+1)\log(2)\Big)\end{aligned}$$

for all large n. Since this is summable in n, the claimed bound for the upcrossings follows for all typical price paths. To bound the downcrossings, it suffices to note that up- and downcrossings of a given interval differ by at most 1. $\square$

The following construction is partly inspired by [MP10], Chapter 6.2.

Theorem 3.2.5. *Let $T > 0$, $\alpha \in (0, 1/2)$ and (π^n) as defined in (3.8). For typical price paths $\omega \in \Omega$, the discrete local time $L^{\pi^n}(S, \cdot)$ converges uniformly in $(t, u) \in [0, T] \times \mathbb{R}$ to a limit $L(S, \cdot) \in C([0, T], C^\alpha(\mathbb{R}))$, and there exists $C = C(\omega) > 0$ such that*

$$\sup_n \Big\{2^{n\alpha} ||L^{\pi^n}(S, \cdot) - L(S, \cdot)||_{L^\infty([0,T]\times\mathbb{R})}\Big\} \leq C. \tag{3.9}$$

Moreover, for all $p > 2$ we have $\sup_{n\in\mathbb{N}} ||L^{\pi^n}||_{C_T \mathcal{V}^p} < \infty$ for typical price paths.

Proof. By the identity (3.7) it suffices to prove the corresponding statements with the stochastic integrals $\int_0^t 1_{(-\infty,u)}(S_s)\, \mathrm{d}S_s$ replacing $L_t(S, u)$. Using Lemma 3.2.4, we may fix $K > 0$ and restrict our attention to the set

$$\begin{aligned}A_K := \Big\{\omega \in \Omega \,:\, &\sup_{t\in[0,T]} |S_t(\omega)| < K\\ &\text{and } \max_{k\in\mathbb{Z}} \left(U_T^n(\omega, k2^{-n}) + D_T(\omega, k2^{-n})\right) \leq K n^2 2^n \ \forall\, n\Big\}.\end{aligned}$$

Let $u \in (-K, K)$. For every $n \in \mathbb{N}$ we approximate $\mathbf{1}_{(-\infty,u)}(S)$ by the process

$$F_t^n(u) := \sum_{k=0}^\infty \mathbf{1}_{(-\infty,u)}(S_{\tau_k^n}) \mathbf{1}_{[\tau_k^n, \tau_{k+1}^n)}(t), \quad t \geq 0.$$

Then we write for the corresponding integral process

$$I_t^{\pi^n}(u) := \sum_{k=0}^\infty \mathbf{1}_{(-\infty,u)}(S_{\tau_k^n}(\omega))[S_{\tau_{k+1}^n \wedge t}(\omega) - S_{\tau_k^n \wedge t}(\omega)], \quad t \geq 0,$$

and since (π^n) is increasing, we get

$$I_t^{\pi^n}(u) - I_t^{\pi^{n-1}}(u) = \sum_{k=0}^\infty [F_{\tau_k^n}^n(u) - F_{\tau_k^n}^{n-1}(u)][S_{\tau_{k+1}^n \wedge t} - S_{\tau_k^n \wedge t}].$$

By the construction of our stopping times (τ_k^n), we have

$$\sup_{t\geq 0}\big|[F^n_{\tau_k^n}(u)-F^{n-1}_{\tau_k^n}(u)][S_{\tau_{k+1}^n\wedge t}(\omega)-S_{\tau_k^n\wedge t}(\omega)]\big|\leq 2^{-n+2}.$$

Hence, the pathwise Hoeffding inequality, Theorem 3 in [Vov12] or Lemma A.1.1, implies for every $\lambda\in\mathbb{R}$ the existence of a 1-admissible simple strategy $H^\lambda\in\mathcal{H}_1$, such that

$$1+(H^\lambda\cdot S)_t(\omega)\geq \exp\left(\lambda(I_t^{\pi^n}(u)-I_t^{\pi^{n-1}}(u))-\frac{\lambda^2}{2}N_t^n(u,\omega)2^{-2n+4}\right)=:\mathcal{E}_t^{\lambda,n}(\omega)$$

for all $t\in[0,T]$ and all $\omega\in\Omega$, where $N_t^n(u):=N_t^n(u,\omega)$ denotes the number of stopping times $\tau_k^n\leq t$ with $F^n_{\tau_k^n}(u)-F^{n-1}_{\tau_k^n}(u)\neq 0$. Now observe that F_t^n and F_t^{n-1} are constant on dyadic intervals of length 2^{-n}, which means that we may suppose without loss of generality that $u=k2^{-n}$ is a dyadic number. But we can estimate $N_T^n(k2^{-n})$ by the number of upcrossings of the interval $[(k-1)2^{-n},k2^{-n}]$ plus the number of the downcrossings of the interval $[k2^{-n},(k+1)2^{-n}]$, which means that on A_K we have $N_T^n(u)\leq 2K2^n n^2$. So considering $(H^\lambda+H^{-\lambda})/2$ for $\lambda>0$, we get

$$\overline{P}\Big(\Big\{\sup_{t\in[0,T]}|I_t^{\pi^n}(u)-I_t^{\pi^{n-1}}(u)|\geq 2^{-n\alpha}\Big\}\cap A_K\Big)\leq 2\exp(-\lambda 2^{-n\alpha}+\lambda^2K2^{-n+4}n^2)$$

for all $\lambda,\alpha>0$. Choose now $\lambda=2^{n/2}$ and $\alpha\in(0,1/2)$. Then we get the estimate

$$\overline{P}\Big(\Big\{\sup_{t\in[0,T]}|I_t^{\pi^n}(u)-I_t^{\pi^{n-1}}(u)|\geq 2^{-n\alpha}\Big\}\cap A_K\Big)\leq 2\exp(-2^{n(1/2-\alpha)}+16Kn^2).$$

Moreover, noting that for all $t>0$ the maps $u\mapsto I_t^{\pi^n}(u)$ and $u\mapsto I_t^{\pi^{n-1}}(u)$ are constant on dyadic intervals of length 2^{-n} and that there are $2K2^n$ such intervals in $[-K,K]$, we can simply estimate

$$\begin{aligned}\overline{P}\Big(\Big\{\sup_{(t,u)\in[0,T]\times\mathbb{R}}|I_t^{\pi^n}(u)-I_t^{\pi^{n-1}}(u)|\geq 2^{-n\alpha}\Big\}\cap A_K\Big)&\\ \leq 2K2^n\times 2\exp(-2^{n(1/2-\alpha)}+16Kn^2)&\\ =\exp(-2^{n(1/2-\alpha)}+16Kn^2+(n+2)\log 2+\log K).&\end{aligned}$$

Obviously, this is summable in n and thus the proof of the uniform convergence and of the speed of convergence is complete.

It remains to prove the uniform bound on the p-variation norm of I^{π^n} and the Hölder continuity of the limit. Let $p>2$ and write $\alpha=1/p$, so that $\alpha\in(0,1/2)$. First let $u=k2^{-n}\in(-K,K)$ and write $v=(k+1)2^{-n}$. Then

$$I_t^{\pi^n}(v)-I_t^{\pi^n}(u)=\sum_{k=0}^\infty(F^n_{\tau_k^n}(v)-F^n_{\tau_k^n}(u))(S_{\tau_k^n\wedge t}-S_{\tau_{k-1}^n\wedge t}),$$

and similarly as before we have $\sup_{t\geq 0}|(F^n_{\tau_k^n}(v)-F^n_{\tau_k^n}(u))(S_{\tau_k^n\wedge t}-S_{\tau_{k-1}^n\wedge t})|\leq 2^{-n+1}$. On A_K, the number of stopping times $(\tau_k^n)_k$ with $F^n_{\tau_k^n}(u)\neq F^n_{\tau_k^n}(v)$ is bounded from above by $2K2^n n^2+1$, and therefore we can estimate as before

$$\overline{P}\left(\left\{\sup_{t\in[0,T]}\sup_{u,v\in\mathbb{R}:|u-v|\leq 2^{-n}}|I_t^{\pi^n}(v)-I_t^{\pi^n}(u)|\geq 2^{-n\alpha}\right\}\cap A_K\right)\leq \exp(-2^{n(1/2-\alpha)}+Cn^2),$$

for some appropriate constant $C=C(K)>0$.

We conclude that for typical price paths $\omega\in\Omega$ there exists $C=C(\omega)>0$ such that

$$\sup_{t\in[0,T]}\sup_{|u-v|\leq 2^{-n}}|I_t^{\pi^n}(v)-I_t^{\pi^n}(u)|+\sup_{t\in[0,T]}\sup_{u\in\mathbb{R}}|I_t^{\pi^n}(u)-I_t^{\pi^{n-1}}(u)|\leq C2^{-n\alpha}$$

for all $n\in\mathbb{N}$. Let now $n\in\mathbb{N}$ and let $u,v\in\mathbb{R}$ with $1\geq|u-v|\geq 2^{-n}$. Let $m\leq n$ be such that $2^{-m-1}<|u-v|\leq 2^{-m}$. Then

$$\begin{aligned}
&||I^{\pi^n}(v)-I^{\pi^n}(u)||_\infty\\
&\qquad\leq ||I^{\pi^n}(v)-I^{\pi^m}(v)||_\infty+||I^{\pi^m}(v)-I^{\pi^m}(u)||_\infty+||I^{\pi^m}(u)-I^{\pi^n}(u)||_\infty\\
&\qquad\leq C\left(\sum_{k=m+1}^n 2^{-k\alpha}+2^{-m\alpha}+\sum_{k=m+1}^n 2^{-k\alpha}\right)\leq C2^{-m\alpha}\leq C|v-u|^\alpha,
\end{aligned}$$

possibly adapting the value of $C>0$ in every step. Since $I_t^{\pi^n}$ is constant on dyadic intervals of length 2^{-n}, this proves that $\sup_{t\in[0,T]}||I_t^{\pi^n}||_{p\text{-var}}\leq C$. The α-Hölder continuity of the limit is shown in the same way. $\square$

We reduced the problem of constructing L to the problem of constructing certain integrals. In Corollary 2.2.6, we gave a general pathwise construction of stochastic integrals. But this result does not apply here, because in general $\mathbf{1}_{(-\infty,u)}(S)$ is not càdlàg.

Remark 3.2.6. *Theorem 3.2.5 gives a simple, model free proof that local times exist and have nice properties. Let us stress again that by Proposition 3.2.3, all the statements of Theorem 3.2.5 hold quasi-surely for all probability measures on $(\Omega,\mathcal{F})$ under which S satisfies (NA1).*

Below, we sketch an alternative proof based on Vovk's pathwise Dambis Dubins-Schwarz theorem. While we are interested in a statement for typical price paths, which a priori is stronger than a quasi-sure result for all measures satisfying (NA1), the quasi-sure statement may also be obtained by observing that every process satisfying (NA1) admits a dominating local martingale measure, see [Ruf13, IP11]. Under the local martingale measure we can then perform a time change to turn the coordinate process into a Brownian motion, and then we can invoke standard results for Brownian motion for which all statements of Theorem 3.2.5 except one are well known: The only result we could not find in the literature is the uniform boundedness in p-variation of the discrete local times.

Remark 3.2.7. *Note that for $u=k2^{-n}$ with $k\in\mathbb{Z}$ we have $L_t^{\pi^n}(u)=2^{-n}D_t(u-2^{-n},u)+\varepsilon(n,t,u)$ for some $\varepsilon(n,t,u)\in[0,2^{-n}]$. Therefore, our proof also shows that the renormalized downcrossings converge uniformly to the local time, with speed at* *least* $2^{-n\alpha}$ *for* $\alpha < 1/2$. *For the Brownian motion this is well known, see [CLPT81]; see also [Kho94] for the exact speed of convergence. In the Brownian case, we actually know more: Outside of one fixed null set we have*

$$\lim_{\varepsilon\to 0}\sup_{x\in\mathbb{R}}\sup_{t\in[0,T]}|\varepsilon^{-1}D_t(x,x+\varepsilon)-L_t(x)|=0$$

for all $T > 0$. *It should be possible to recover this result also in our setting. It follows from pathwise estimates once we prove Theorem 3.2.5 for a sequence of partitions* $(\widetilde{\pi}^n)$ *of the following type: Let* (c_n) *be a sequence of strictly positive numbers converging to 0, such that* c_{n+1}/c_n *converges to 1. Define* $\widetilde{\mathbb{D}}^n := \{kc_n : k \in \mathbb{Z}\}$. *Now define* $\widetilde{\pi}^n$ *as* π^n, *replacing* $\mathbb{D}^n$ *by* $\widetilde{\mathbb{D}}^n$. *The only problem is that then we cannot expect the sequence* $(\widetilde{\pi}^n)$ *to be increasing, and this would complicate the presentation, which is why we prefer to work with the dyadic Lebesgue partition.*

Finally, we want to briefly indicate that Theorem 3.2.5 could also be partially proven by relying on the pathwise Dambis Dubins-Schwarz type theorem of Vovk [Vov12], which allows to transfer properties of the one-dimensional Wiener process to typical price paths. For a more detailed exposition of the time-change argument we refer to Appendix A.3.

As mentioned above, Vovk's outer measure $\overline{Q}$ is defined slightly differently than $\overline{P}$ but all results which hold true outside of a $\overline{Q}$-null set are also true outside of a $\overline{P}$-null set; see Section 2.1.3. To understand Vovk's pathwise Dambis Dubins-Schwarz theorem, we need to recall the definition of time-superinvariant sets.

Definition 3.2.8. A continuous non-decreasing function $f\colon [0,\infty) \to [0,\infty)$ satisfying $f(0) = 0$ is said to be a *time change*. A subset $A \subseteq \Omega$ is called *time-superinvariant* if for each $\omega \in \Omega$ and each time change f it is true that $\omega \circ f \in A$ implies $\omega \in A$.

Roughly speaking, Vovk proved in Theorem 3.1 of [Vov12] that the Wiener measure of a time-superinvariant set equals the outer measure $\overline{Q}$ of this set. It turns out that the sets

$$\begin{aligned} A_c &:= \{\omega \in \Omega : S(\omega) \in \mathcal{L}_c\} \quad \text{and} \\ A_{c,p} &:= \{\omega \in A_c : u \mapsto L_t(S,u)(\omega) \text{ has finite } p\text{-variation for all } t \in [0,\infty)\} \end{aligned}$$

are time-superinvariant. Based on this, one can rely on classical results for the Wiener process (see [KS88], Theorem 3.6.11 or [MP10], Theorem 6.19) to show that typical price paths have an absolutely continuous occupation measure $L_t(S,u)$ with jointly continuous density and that $L_t(S,\cdot)$ has finite p-variation which is uniformly bounded in $t \in [0,T]$ for all $T > 0$ and all $p > 2$ (see [MP10], Theorem 6.19).

However, to the best of our knowledge the alternative approach does not give us the uniform boundedness in p-variation of the approximating sequence (L^{π^n}): we were not able to find such a result in the literature on Brownian motion. Without this, we would only be able to prove an abstract version of Theorem 3.1.8, where the pathwise stochastic integral $\int_0^t g(S_s)\,\mathrm{d}S_s$ is defined by approximating g with smooth functions for which the Föllmer-Itô formula Theorem 3.1.2 holds (see [FZ06] for similar arguments in a semimartingale context). Since we are interested in the Riemann sum interpretation of the pathwise integral, we need Theorem 3.2.5 to make sure that all requirements of Theorem 3.1.8 are satisfied for typical price paths.

4. Existence of Lévy's area and pathwise integration

The theory of rough paths (see [LCL07, Lej09, FH14]) has established an analytical frame in which stochastic differential and integral calculus beyond Young's classical notions is traced back to properties of the trajectories of processes involved without reference to a particular probability measure. See Section 2.3.1 for a brief introduction. For instance, in the simplest non-trivial setting it provides a topology on the set of continuous functions enhanced with an "area", with respect to which the (Itô) map associating the trajectories of a solution process of a stochastic differential equation driven by trajectories of a continuous martingale is continuous. In this topology, convergence of a sequence of functions $X^n = (X^{1,n}, \dots, X^{d,n})_{n\in\mathbb{N}}$ defined on the time interval $[0,T]$ involves besides uniform convergence also the convergence of the Lévy areas associated to the vector of trajectories, formally given by

$$\mathbb{L}_t^{i,j,n} := \int_0^t (X_s^{i,n}\,\mathrm{d}X_s^{j,n} - X_s^{j,n}\,\mathrm{d}X_s^{i,n}), \quad 1 \le i,j \le d, \quad t \in [0,T].$$

In Chapter 2 and especially in Section 2.3, we proved that the iterated integrals of typical price paths exist and in particular Lévy's area always exists for typical price paths. In probability theory the concept of Lévy's area is much older and was already studied in the 1940s. It was first introduced by P. Lévy in [Lé40] for a two dimensional Brownian motion (B^1, B^2). For time T fixed and any trajectory of the process it is defined as the area enclosed by the trajectory (B^1, B^2) and the chord given by the straight line from $(0,0)$ to (B_T^1, B_T^2), and may be expressed formally by

$$\frac{1}{2}\Big(\int_0^T B_t^1\,\mathrm{d}B_t^2 - \int_0^T B_t^2\,\mathrm{d}B_t^1\Big),$$

provided the integrals make sense.

More recently, an alternative calculus with a more Fourier analytic touch has been designed (see [GIP14, Per14]) in which an older idea by Gubinelli [Gub04] is further developed. It is based on the concept of *controlled paths*. In this calculus, rough path integrals are described in terms of Fourier series for instance in the Haar-Schauder wavelet, and are seen to decompose into different parts, one of them representing Lévy's area. The existence of a stochastic integral in this approach is seen to be linked to the existence of the corresponding Lévy area, and both can be approximated along a Schauder development in which Hölder functions are limits of their finite degree Schauder expansions. In its simplest (one-dimensional) form a path of bounded variation Y on $[0,T]$ is *controlled by* another path X of bounded variation on $[0,T]$, if the associated signed measures μ_X, μ_Y on the Borel sets of $[0,T]$ satisfy that μ_Y is absolutely continuous with respect to μ_X. In its version relevant here two rough (vector valued) functions X and Y on $[0, T]$ are considered, both with finite p-variation for some $p \geq 1$. In the simplest setting, Y is *controlled by* X if there exists a function Y' of finite p-variation such that the first order Taylor expansion errors

$$R^Y_{s,t} = Y_t - Y_s - Y'_s(X_t - X_s)$$

are bounded in a suitable semi-norm, i.e. $\sum_{[s,t]\in\pi} |R^Y_{s,t}|^r$ is bounded over all possible partitions π of $[0, T]$. Here $\frac{1}{r} = \frac{2}{p}$. Since for a path X Hölder continuity of order $\frac{1}{p}$ is closely related to finite p-variation, the control relation can be seen as expressing a type of fractional Taylor expansion of first order: the first order Taylor expansion error of Y with respect to X - both of Hölder order $\frac{1}{p}$ and "derivative" Y' - is of double Hölder order $\frac{2}{p}$. In its *para-controlled* refinement as developed by Gubinelli et al. in [GIP12] this notion has been seen to give an alternative approach to classical rough path analysis, which we shall generalize in Chapter 5. In the comparison of the two approaches, to make the Itô map continuous, information stored in the Lévy areas of vector valued paths has to be complemented by information conveyed by path control or vice versa. This raises the problem about the relationship between the existence of Lévy's area and the control relationship between vector trajectories or the components of such. We shall deal with this fundamental problem in Section 4.1.

Based on this study we then decompose Riemann approximations of different versions of integrals into a symmetric and an antisymmetric component and prove that for the classical Stratonovich integral just the antisymmetric Riemann sums have to converge, while for more general Stratonovich or Itô type integrals the existence of limits for the symmetric part has to be guaranteed along fixed sequences of partitions, as in Föllmer's approach [Fö81]. Under this assumption we additionally derive a pathwise version of a functional Itô formula due to [Ahn97], where the functional has to be just defined on the space of continuous functions. At this point our Itô formula circumvents a technical problem of Dupire differentiability (see [Dup09, CF13]), where the functional has to be defined for càdlàg functions as well.

The chapter is organized as follows. In Section 4.1 we show that for a vector X of functions a particular version of control, which we will call self-control, is sufficient for the pathwise existence of the Lévy areas. An example of two functions is given which are not mutually controlled and for which consequently Lévy's area fails to exist. In Section 4.2 we study the question how control concepts and the existence of different kinds of integrals (Itô type, Stratonovich type) are related, and in particular in which way control leads to versions of Föllmer's pathwise Itô formula. Finally, provided the quadratic variation exists, we present a pathwise version of a functional Itô's formula in Section 4.2.1.

4.1. Lévy's area and controlled paths

It is well-known that both the control of a path Y with respect to another path X, as well as the existence of Lévy's area for X entails the existence of the rough path integral of Y with respect to X, as we have seen in Section 2.3. This raises the question about the relative power of the hypotheses leading to the existence of the integral.

This question will be answered here. We will show that control entails the existence of Lévy's area. The analysis we present, as usual, is based on d-dimensional irregular paths, and corresponding notion of areas. For a continuous path $X\colon [0,T] \to \mathbb{R}^d$, say $X = (X^1, \dots, X^d)^*$, we recall that *Lévy's area* $\mathbb{L}(X) = (\mathbb{L}^{i,j}(X))_{i,j}$ is given by

$$\mathbb{L}(X)^{i,j} := \int_0^T X_t^i \, \mathrm{d}X_t^j - \int_0^T X_t^j \, \mathrm{d}X_t^i, \quad 1 \le i, j \le d,$$

where X^* denotes the transpose of the vector X, if the respective integrals exist. There are pairs of Hölder continuous paths X^1 and X^2 for which Lévy's area does not exist (see Example 4.1.8 below). To answer this question, we need the basic setup of rough path analysis, which we briefly recall here for the convenience of the reader, starting with the notion of power variation.

A *partition* $\pi := \{[t_{i-1}, t_i] : i = 1, \dots, N\}$ of an interval $[0,T]$ is a family of essentially disjoint intervals such that $\bigcup_{i=1}^N [t_{i-1}, t_i] = [0,T]$. For any $1 \le p < \infty$, a continuous function $X\colon [0,T] \to \mathbb{R}^d$ is of finite *p-variation* if

$$||X||_p := \sup_{\pi \in \mathcal{P}} \Bigg(\sum_{[s,t] \in \pi} |X_{s,t}|^p \Bigg)^{\frac{1}{p}} < \infty,$$

where the supremum is taken over the set $\mathcal{P}$ of all partitions of $[0,T]$ and $X_{s,t} := X_t - X_s$ for $s,t \in [0,T]$, $s \le t$. We write $\mathcal{V}^p([0,T], \mathbb{R}^d)$ for the set (linear space) of continuous functions of finite p-variation. Let, more generally, $R\colon [0,T]^2 \to \mathbb{R}^{d \times d}$ be a continuous function. In this case we consider the functional

$$||R||_r := \sup_{\pi \in \mathcal{P}} \Bigg(\sum_{[s,t] \in \pi} |R_{s,t}|^r \Bigg)^{\frac{1}{r}}, \quad 1 \le r < \infty.$$

An equivalent way to characterize the property of finite p-variation is by the existence of a control function. Denoting by $\Delta_T := \{(s,t) \in [0,T]^2 : 0 \le s \le t \le T\}$, we call a continuous function $\omega\colon \Delta_T \to \mathbb{R}^+$ vanishing on the diagonal *control function* if it is superadditive, i.e. if for $(s,u,t) \in [0,T]^3$ one has $\omega(s,u) + \omega(u,t) \le \omega(s,t)$ for $0 \le s \le u \le t \le T$. Note that a function is of finite p-variation if and only if there exists a control function ω such that $|X_{s,t}|^p \le \omega(s,t)$ for $(s,t) \in \Delta_T$. For a more detailed discussion of p-variation and control functions see Chapter 1.2 in [LCL07]. For later reference we remark that all objects are analogously defined for general Banach spaces instead of $\mathbb{R}^d$.

A fundamental insight due to Gubinelli [Gub04] was that an integral $\int Y \, \mathrm{d}X$ exists if "Y looks like X in the small scale", cf. Section 2.3.1. This leads to the concept of controlled paths, which we recall in its general form.

Definition 4.1.1. Let $p,q,r \in \mathbb{R}^+$ be such that $2/p + 1/q > 1$ and $1/r = 1/p + 1/q$. Suppose $X \in \mathcal{V}^p([0,T], \mathbb{R}^d)$. We call $Y \in \mathcal{V}^p([0,T], \mathbb{R}^d)$ *controlled* by X if there exists $Y' \in \mathcal{V}^q([0,T], \mathbb{R}^{d \times d})$ such that the remainder term R^Y given by the relation $Y_{s,t} = Y_s' X_{s,t} + R_{s,t}^Y$ satisfies $||R^Y||_r < \infty$. In this case we write $Y \in \mathscr{C}_X^q$, and call Y' *Gubinelli derivative.*

See Theorem 1 in [Gub04] for the case of Hölder continuous paths, or Theorem 2.3.9 for precise existence results of $\int Y \,\mathrm{d}X$. Let us now modify this concept to a notion of control of a path by itself.

Definition 4.1.2. Let $p, q, r \in \mathbb{R}^+$ be such that $2/p + 1/q > 1$ and $1/r = 1/p + 1/q$. We call $X \in \mathcal{V}^p([0,T], \mathbb{R}^d)$ *self-controlled* if we have $X^i \in \mathscr{C}^q_{X^j}$ or $X^j \in \mathscr{C}^q_{X^i}$ for all $1 \leq i, j \leq d$ with $i \neq j$.

With this notion we are now able to deal with the main task of this section, the construction of the Lévy area of a self-controlled path X. In fact, the integrals arising in Lévy's area will be obtained via left-point Riemann sums as

$$\mathbb{L}(X)^{i,j} = \int_0^T X^i_t \,\mathrm{d}X^j_t - \int_0^T X^j_t \,\mathrm{d}X^i_t := \lim_{|\pi| \to 0} \sum_{[s,t] \in \pi} (X^i_s X^j_{s,t} - X^j_s X^i_{s,t}), \tag{4.1}$$

for $1 \leq i, j \leq d$, where $|\pi|$ denotes the mesh of a partition π. Our approach uses the abstract version of classical ideas due to Young [You36] comprised in the so-called *sewing lemma*.

Lemma 4.1.3. *[Corollary 2.3, Corollary 2.4 in [FD06]] Let $\Xi\colon \Delta_T \to \mathbb{R}^d$ be a continuous function and $K > 0$ some constant. Assume that there exist a control function ω and a constant $\vartheta > 1$ such that for all $(s,u,t) \in [0,T]^3$ with $0 \leq s \leq u \leq t \leq T$ we have*

$$|\Xi_{s,t} - \Xi_{s,u} - \Xi_{u,t}| \leq K\omega(s,t)^{\vartheta}. \tag{4.2}$$

Then there exists a unique function $\Phi\colon [0,T] \to \mathbb{R}^d$ such that $\Phi(0) = 0$ and

$$|\Phi(t) - \Phi(s) - \Xi_{s,t}| \leq C(\vartheta)\omega(s,t)^{\vartheta} \quad \text{and} \quad \lim_{|\pi(s,t)| \to 0} \sum_{[u,v] \in \pi(s,t)} \Xi_{u,v} = \Phi(t) - \Phi(s),$$

for $(s,t) \in \Delta_T$, where $C(\vartheta) := K(1 - 2^{1-\vartheta})^{-1}$ and $\pi(s,t)$ denotes a partition of $[s,t]$.

Remark 4.1.4. *For simplicity we state Lemma 4.1.3 only for a continuous function $\Xi\colon \Delta_T \to \mathbb{R}^d$. Yet, it still holds true without the continuity assumption and for a general Banach space replacing $\mathbb{R}^d$. See Theorem 1 and Remark 3 in [FDM08]. Consequently, all results of this section extend to general Banach spaces.*

With this tool we now derive the existence of Lévy's area for self-controlled paths of finite p-variation with $p \geq 1$.

Theorem 4.1.5. *Let $1 \leq p < \infty$ and suppose that $X \in \mathcal{V}^p([0,T], \mathbb{R}^d)$ is self-controlled, then Lévy's area as defined in (4.1) exists.*

Proof. Let $X \in \mathcal{V}^p([0,T], \mathbb{R}^d)$ for $1 \leq p < \infty$ be self-controlled and fix $1 \leq i, j \leq d$, $i \neq j$. We may assume without loss of generality that $X^i \in \mathscr{C}^q_{X^j}$, i.e. $X^i_{s,t} = X'_s(i,j) X^j_{s,t} + R^{i,j}_{s,t}$ and $||X'(i,j)||_q, ||R^{i,j}||_r < \infty$. In order to apply Lemma 4.1.3, we set $\Xi^{i,j}_{s,t} := X^i_s X^j_{s,t} - X^j_s X^i_{s,t}$ for $(s,t) \in \Delta_T$ and observe that for $(s,u,t) \in [0,T]^3$ with $0 \leq s \leq u \leq t \leq T$, we have

$$\begin{aligned}
\Xi^{i,j}_{s,t} - \Xi^{i,j}_{s,u} - \Xi^{i,j}_{u,t} &= X^j_{s,u} X^i_{u,t} - X^i_{s,u} X^j_{u,t} \\
&= X^j_{s,u}(X'_u(i,j) X^j_{u,t} + R^{i,j}_{u,t}) - (X'_s(i,j) X^j_{s,u} + R^{i,j}_{s,u}) X^j_{u,t} \\
&= X^j_{s,u} R^{i,j}_{u,t} - R^{i,j}_{s,u} X^j_{u,t} + (X'_u(i,j) - X'_s(i,j)) X^j_{s,u} X^j_{u,t}.
\end{aligned}$$

Since the finite sum of control functions is again a control function, we can choose the same control function ω for $X^j, X'(i,j)$ and $R^{i,j}$, and setting $\vartheta := \frac{2}{p} + \frac{1}{q} > 1$ we get

$$|\Xi^{i,j}_{s,t} - \Xi^{i,j}_{s,u} - \Xi^{i,j}_{u,t}| \leq \omega(s,t)^{\frac{1}{p}+\frac{1}{r}} + \omega(s,t)^{\frac{1}{p}+\frac{1}{r}} + \omega(s,t)^{\frac{2}{p}+\frac{1}{q}} \leq 3\omega(s,t)^{\vartheta}.$$

□

We will next show that Riemann sums with arbitrary choices of base points for the integrand functions lead to the same Lévy area as just constructed.

Lemma 4.1.6. *Let $X \in \mathcal{V}^p([0,T],\mathbb{R}^d)$ for some $1 \leq p < \infty$. Suppose that X is self-controlled. Denote by $s' \in [s,t]$ an arbitrary point chosen in a partition interval $[s,t] \in \pi$. Then Lévy's area from the preceding theorem is also given by*

$$\mathbb{L}(X)^{i,j} = \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} (X^i_{s'} X^j_{s,t} - X^j_{s'} X^i_{s,t}), \quad 1 \leq i,j \leq d.$$

Proof. For a self-controlled path $X \in \mathcal{V}^p([0,T],\mathbb{R}^d)$ with $1 \leq p < \infty$ we may assume without loss of generality that $X^i \in \mathscr{C}^q_{X^j}$ for $1 \leq i,j \leq d$, $i \neq j$. From Theorem 4.1.5 we already know that the left-point Riemann sums converge. Hence, we only need to show that

$$\sum_{[s,t]\in\pi_n} (X^i_s X^j_{s,t} - X^j_s X^i_{s,t}) - \sum_{[s,t]\in\pi_n} (X^i_{s'} X^j_{s,t} - X^j_{s'} X^i_{s,t}) \tag{4.3}$$

tends to zero along every sequence of partitions (π_n) such that the mesh $|\pi_n|$ converges to zero. Indeed, we may write for a partition interval $[s,t]$

$$\begin{aligned} X^i_s X^j_{s,t} - X^j_s X^i_{s,t} - (X^i_{s'} X^j_{s,t} - X^j_{s'} X^i_{s,t}) &= -X^i_{s,s'} X^j_{s,t} + X^j_{s,s'} X^i_{s,t} \\ &= -(X'_s(i,j) X^j_{s,s'} + R^{i,j}_{s,s'}) X^j_{s,t} + X^j_{s,s'}(X'_s(i,j) X^j_{s,t} + R^{i,j}_{s,t}) \\ &= -R^{i,j}_{s,s'} X^j_{s,t} + X^j_{s,s'} R^{i,j}_{s,t}. \end{aligned}$$

Taking the same control function ω for X^j and $R^{i,j}$, we estimate

$$|X^i_s X^j_{s,t} - X^j_s X^i_{s,t} - (X^i_{s'} X^j_{s,t} - X^j_{s'} X^i_{s,t})| = |-R^{i,j}_{s,s'} X^j_{s,t} + X^j_{s,s'} R^{i,j}_{s,t}| \leq 2\omega(s,t)^{\vartheta}$$

with $\vartheta := \frac{2}{p} + \frac{1}{p} > 1$. Recalling the superadditivity of ω, we get for $n \in \mathbb{N}$

$$\left| \sum_{[s,t]\in\pi_n} (X^i_{s,s'} X^j_{s,t} - X^j_{s,s'} X^i_{s,t}) \right| \leq \sum_{[s,t]\in\pi_n} \omega(s,t)^{\vartheta} \leq \max_{[s,t]\in\pi_n} \omega(s,t)^{\vartheta-1} \omega(0,T),$$

which means that (4.3) tends to zero as $n \to \infty$. □

Example 4.1.7. Let $(B_t\,;\, t \in [0,T])$ be a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $f \in C^1(\mathbb{R},\mathbb{R})$ be a continuously differentiable function with α-Hölder continuous derivative for $\alpha > 0$. The trajectories of B are of finite p-variation for all $p > 2$ outside a null set $\mathcal{N}$. Thus we can deduce from Theorem 4.1.5 that Lévy's area of $(B + g_1, f(B) + g_2)$ exists outside the same null set $\mathcal{N}$ whenever $g_1, g_2 \in \mathcal{V}^q([0,T],\mathbb{R}^d)$ for some $1 \leq q < 2$.

The following example illustrates that for $p \geq 2$ things are essentially different. It will in particular show that in this case self-control of a path is necessary for the existence of Lévy's area.

Example 4.1.8. Let us consider for $m \in \mathbb{N}$ the functions $X^m\colon [-1,1] \to \mathbb{R}^2$ with components given by

$$X_t^{1,m} := \sum_{k=1}^m a_k \sin(2^k \pi t) \quad \text{and} \quad X_t^{2,m} := \sum_{k=1}^m a_k \cos(2^k \pi t), \quad t \in [-1,1],$$

where $a_k := 2^{-\alpha k}$ and $\alpha \in (0,1)$. Set $X := \lim_{m\to\infty} X^m$. These functions are α-Hölder continuous uniformly in m. Indeed, let $s,t \in [-1,1]$ and choose $k \in \mathbb{N}$ such that $2^{-k-1} \leq |s-t| \leq 2^{-k}$. Then we can estimate as follows

$$\begin{aligned}
|X_t^{1,m} - X_s^{1,m}| &= \Big| \sum_{l=1}^m a_l 2 \cos(2^{l-1}\pi(s+t)) \sin(2^{l-1}\pi(s-t)) \Big| \\
&\leq 2 \sum_{l=1}^k |a_l| |\sin(2^{l-1}\pi(s-t))| + 2 \sum_{l=k+1}^\infty |a_l| \\
&\leq 2 \sum_{l=1}^k |a_l| 2^{l-1}\pi |s-t| + 2 \sum_{l=k+1}^\infty |a_l| \\
&\leq \sum_{l=1}^k 2^{l-\alpha l}\pi |s-t| + 2^{-\alpha(k+1)+1} \frac{1}{1-2^{-\alpha}} \\
&\leq \frac{2^{(k+1)(1-\alpha)}-1}{2^{1-\alpha}-1}\pi |s-t| + \frac{2^{1-\alpha}}{1-2^{-\alpha}} |s-t|^\alpha \\
&\leq \frac{2^{(k+1)(1-\alpha)}-1}{2^{1-\alpha}-1}\pi 2^{-k(1-\alpha)} |s-t|^\alpha + \frac{2^{1-\alpha}}{1-2^{-\alpha}} |s-t|^\alpha \leq C|s-t|^\alpha
\end{aligned}$$

for some constant $C > 0$ independent of $m \in \mathbb{N}$. Analogously, we can get the α-Hölder continuity of $X^{2,m}$. Furthermore, it can be seen with the same estimate that (X^m) converges uniformly to X and thus also in α-Hölder topology. The limit function X is not β-Hölder continuous for every $\beta > \alpha$. In order to see this, choose $s = 0$ and $t = t_n = 2^{-n}$ for $n \in \mathbb{N}$ and observe that

$$\frac{|X_{t_n}^1 - X_0^1|}{|t_n - 0|^\beta} = \sum_{k=1}^{n-1} 2^{-\alpha k + \beta n} \sin(2^{k-n}\pi) \geq 2^{(\beta-\alpha)n+\alpha},$$

which obviously tends to infinity as n tends to infinity. Since α-Hölder continuity is obviously related to finite $\frac{1}{\alpha}$-variation, we can conclude that $X \in \mathcal{V}^{\frac{1}{\alpha}}([-1,1],\mathbb{R}^2)$, and $X \notin \mathcal{V}^\gamma([-1,1],\mathbb{R}^2)$ for $\gamma < \frac{1}{\alpha}$. Let us now show that X possesses no Lévy area. For this purpose, fix $\alpha \in (0,1)$ and $m \in \mathbb{N}$. Then Lévy's area for X^m is given by

$$\begin{aligned}
&\int_{-1}^1 X_s^{1,m} \,\mathrm{d}X_s^{2,m} - \int_{-1}^1 X_s^{2,m} \,\mathrm{d}X_s^{1,m} \\
&\qquad = -\sum_{k,l=1}^m a_k a_l \int_{-1}^1 \big(\sin(2^k\pi s)\sin(2^l \pi s) 2^l \pi + \cos(2^l \pi s)\cos(2^k \pi s) 2^k \pi \big) \,\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&= -\sum_{k,l=1}^{m} a_k a_l \Big(2^l \pi \int_{-1}^{1} \frac{1}{2}\big(\cos((2^k-2^l)\pi s) - \cos((2^k+2^l)\pi s)\big)\,\mathrm{d}s \\
&\qquad\qquad + 2^k \pi \int_{-1}^{1} \frac{1}{2}\big(\cos((2^k-2^l)\pi s) + \cos((2^k+2^l)\pi s)\big)\,\mathrm{d}s \Big) \\
&= -2\sum_{k=1}^{m} a_k^2 2^k \pi = -2\sum_{k=1}^{m} 2^{(1-2\alpha)k}\pi.
\end{aligned}
$$

This quantity diverges as m tends to infinity for $\frac{1}{\alpha} \geq 2$. Since (X^m) converges to X in the α-Hölder topology, we can use this result to choose partition sequences of $[-1,1]$ along which Riemann sums approximating the Lévy area of X diverge as well. This shows that X possesses no Lévy area. In return Theorem 4.1.5 implies that X cannot be self-controlled. However, it is not to hard to see directly that no regularity is gained by controlling X^1 with X^2. For this purpose, note that for $-1 \leq s \leq t \leq 1$, and $0 \neq X'_s \in \mathbb{R}$, one has

$$
\begin{aligned}
|X^1_{s,t} - X'_s X^2_{s,t}| &= \Big| \sum_{k=1}^{\infty} a_k \big[(\sin(2^k\pi t) - \sin(2^k \pi s)) - X'_s(\cos(2^k\pi t) - \cos(2^k\pi s))\big] \Big| \\
&= \Big| 2\sum_{k=1}^{\infty} a_k \big[\sin(2^{k-1}\pi(s-t))\cos(2^{k-1}\pi(s+t)) \\
&\qquad\qquad + X'_s \sin(2^{k-1}\pi(s+t))\sin(2^{k-1}\pi(s-t))\big] \Big| \\
&= \Big| 2\sum_{k=1}^{\infty} a_k \sin(2^{k-1}\pi(s-t))\sqrt{1+(X'_s)^2}\,\sin(2^{k-1}\pi(s+t) + \arctan((X'_s)^{-1})) \Big|.
\end{aligned}
$$

Let us now investigate Hölder regularity at $s = 0$. First, assume $X'_0 > 0$, and take $t = 2^{-n}$ to obtain

$$
\begin{aligned}
&\frac{|X^1_{0,2^n} - X'_0 X^2_{0,2^n}|}{2^{-\beta n}} \\
&\quad = 2^{\beta n} \Big| 2\sum_{k=1}^{n} a_k \sin(2^{k-1-n}\pi)\sqrt{1+(X'_0)^2}\,\sin(2^{k-1-n}\pi + \arctan((X'_0)^{-1})) \Big| \\
&\quad \geq 2^{(\beta-\alpha)n} \sin\big(\frac{\pi}{2} + \arctan((X'_0)^{-1})\big).
\end{aligned}
$$

For $X'_0 < 0$ the same estimates work for $t_n = -2^{-n}$ instead. Therefore, the Hölder regularity at 0 cannot be better than α and in particular X cannot be self-controlled for $\frac{1}{\alpha} > 2$.

4.2. Föllmer integration

In his seminal paper Föllmer [Fö81] considered one dimensional pathwise integrals. He was able to give a pathwise meaning to the limit

$$
\int_0^T \mathrm{D}F(X_t)\,\mathrm{d}^{\pi_n} X_t := \lim_{n\to\infty} \sum_{[s,t]\in\pi_n} \langle \mathrm{D}F(X_s), X_{s,t}\rangle,
$$

provided $F \in C^2(\mathbb{R}^d, \mathbb{R})$. A translation of Föllmer's work, today named *Föllmer integration*, can be found in the appendix of [Son06]. His starting point was the hypothesis that quadratic variation of $X \in C([0,T], \mathbb{R}^d)$ exists along a sequence of partitions $(\pi_n)_{n\in\mathbb{N}}$ whose mesh tends to zero. Here $\langle \cdot, \cdot \rangle$ denotes the usual inner product on $\mathbb{R}^d$. As indicated and discussed below, this construction of an integral depends strongly on the chosen sequence of partitions $(\pi_n)_{n\in\mathbb{N}}$.

Before coming back to an approach to Föllmer's integral, we shall construct a Stratonovich type integral, thereby discussing the problem of dependence on a chosen sequence of partitions. As in the previous section, our approach is based on the notion of controlled paths. This will also lead us on a route which does not require the existence of iterated integrals as in the classical rough path approach. We fix a $\gamma \in [0,1]$, to discuss Stratonovich limits for Riemann sums where integrands are taken as convex combinations $\gamma Y_s + (1-\gamma) Y_t$ of the values of Y at the extremes of a partition interval $[s,t]$. We start by decomposing these sums into symmetric and antisymmetric parts. For $p, q \in [1, \infty)$, $X \in \mathcal{V}^p([0,T], \mathbb{R}^d)$ and $Y \in \mathscr{C}_X^q$ we have

$$
\begin{aligned}
\gamma\text{-}\int_0^T Y_t \,\mathrm{d}X_t &:= \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \langle Y_s + \gamma Y_{s,t}, X_{s,t}\rangle \\
&= \frac{1}{2}\Big(\gamma\text{-}\int_0^T Y_t \,\mathrm{d}X_t + \gamma\text{-}\int_0^T X_t \,\mathrm{d}Y_t\Big) + \frac{1}{2}\Big(\gamma\text{-}\int_0^T Y_t \,\mathrm{d}X_t - \gamma\text{-}\int_0^T X_t \,\mathrm{d}Y_t\Big) \\
&=: \frac{1}{2}\mathbb{S}_\gamma\langle X, Y\rangle + \frac{1}{2}\mathbb{A}_\gamma\langle X, Y\rangle.
\end{aligned} \tag{4.4}
$$

Note that $\gamma = 0$ corresponds to the classical Itô integral and $\gamma = \frac{1}{2}$ to the classical Stratonovich integral.

If the variation orders of X and Y fulfill $1/p + 1/q > 1$, we are in the framework of Young's integration theory. Below 1, either the existence of the rough path or control is needed. To illustrate this, we go back to Example 4.1.8.

Example 4.2.1. Let $X = (X^1, X^2)$ be given according to Example 4.1.8. In this case, we have seen that X^1 and X^2 are of finite $\frac{1}{\alpha}$-variation. With decomposition (4.4) we see that

$$
\begin{aligned}
\frac{1}{2}\text{-}\int_0^1 X_t^2 \,\mathrm{d}X_t^1 &:= \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \langle X_s^2 + \frac{1}{2}X_{s,t}^2, X_{s,t}^1\rangle = \frac{1}{2}\mathbb{S}_{\frac{1}{2}}\langle X^1, X^2\rangle + \frac{1}{2}\mathbb{A}_{\frac{1}{2}}\langle X^1, X^2\rangle \\
&= \frac{1}{2}\lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \frac{1}{2}\big(\langle X_s^2 + X_t^2, X_t^1 - X_s^1\rangle + \langle X_s^1 + X_t^1, X_t^2 - X_s^2\rangle\big) + \frac{1}{2}\mathbb{L}^{1,2}(X) \\
&= \frac{1}{2}\lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \langle X^1, X^2\rangle_{s,t} + \frac{1}{2}\mathbb{L}^{1,2}(X) = \frac{1}{2}(X_1^1 X_1^2 - X_0^1 X_0^2) + \frac{1}{2}\mathbb{L}^{1,2}(X),
\end{aligned}
$$

provided all terms are well-defined. Therefore, the integral exists if and only if Lévy's area exists, which is not the case for instance if $\alpha = \frac{1}{2}$. So beyond Young's theory, the existence of the $\frac{1}{2}$-Stratonovich integral is closely linked to the existence of Lévy's area.

Using a suitable control concept, we will next construct the Stratonovich integral described above, but not just with restriction to a particular sequence of partitions.

This time, the symmetry of the Gubinelli derivative of a controlled path plays an essential role. However, this symmetry assumption can be avoided if the involved paths control each other.

Definition 4.2.2. Let $X, Y \in \mathcal{V}^p([0,T],\mathbb{R}^d)$. We say that X and Y are *similar* if there exist $X', Y' \in \mathcal{V}^q([0,T],\mathbb{R}^{d\times d})$ such that $X \in \mathscr{C}_Y^q$ with Gubinelli derivative X', $Y \in \mathscr{C}_X^q$ with Gubinelli derivative Y', and $((X_t')^*)^{-1} = Y_t'$ for all $t \in [0,T]$. In this case we write $Y \in \mathscr{S}_X^q$.

Let us give a very simple example of two paths $X, Y \in \mathcal{V}^p([0,T],\mathbb{R}^d)$ such that $Y \in \mathscr{S}_X^q$ but neither $Y \in \mathscr{C}_X^q$ with Y' symmetric nor $X \in \mathscr{C}_Y^q$ with X' symmetric.

Example 4.2.3. For $p \in [2,3)$ take $X^1 \in \mathcal{V}^p([0,T],\mathbb{R})$ and $X^2, X^3 \in \mathcal{V}^{\frac{p}{2}}([0,T],\mathbb{R})$. If we set

$$X := (X^1, X^2, X^3) \quad \text{and} \quad Y := (X^1, 0, 0),$$

we obviously have $X, Y \in \mathcal{V}^p([0,T],\mathbb{R}^3)$. In this case we could choose X' and Y' identical to (z_1, z_2, z_3), where $z_1^* := (1,0,0)$, $z_2^* := (0,0,1)$, and $z_3^* := (0,-1,0)$. We see that $Y \in \mathscr{S}_X^p$, but X' and Y' are not symmetric matrices.

Under both assumptions we prove the existence of the Stratonovich integral described above. This time, thanks to the additional requirements of the Gubinelli derivative, the usual concept of controlled paths is sufficient, and Lévy's area is not needed.

Theorem 4.2.4. *Let $\gamma \in [0,1]$, $X \in \mathcal{V}^p([0,T],\mathbb{R}^d)$. If $Y \in \mathscr{C}_X^q$ and Y_t' is a symmetric matrix for all $t \in [0,T]$, then the antisymmetric part*

$$\mathbb{A}_\gamma\langle X, Y\rangle := \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \big(\langle Y_s + \gamma Y_{s,t}, X_{s,t}\rangle - \langle X_s + \gamma X_{s,t}, Y_{s,t}\rangle\big), \tag{4.5}$$

exists and satisfies

$$\mathbb{A}_\gamma\langle X, Y\rangle = \mathbb{A}\langle X, Y\rangle := \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \big(\langle Y_{s'}, X_{s,t}\rangle - \langle X_{s'}, Y_{s,t}\rangle\big)$$

for every choice of points $s' \in [s,t] \in \pi$. The same result holds if $Y \in \mathscr{S}_X^q$.

Proof. It is easy to verify that by definition the antisymmetric part, if it exists as a limit of the Riemann sums considered, has to satisfy the second formula of the claim at least with the choice $s' = s$, for all intervals $[s,t]$ belonging to a partition. To prove that this limit exists, we use Lemma 4.1.3. For this purpose, we set $\Xi_{s,t} := \langle Y_s, X_{s,t}\rangle - \langle X_s, Y_{s,t}\rangle$ for $(s,t) \in \Delta_T$. Since Y is controlled by X, we obtain

$$\begin{aligned}
\Xi_{s,t} - \Xi_{s,u} - \Xi_{u,t} &= \langle Y_u' X_{u,t} + R_{u,t}^Y, X_{s,u}\rangle - \langle X_{u,t}, Y_s' X_{s,u} + R_{s,u}^Y\rangle \\
&= \langle R_{u,t}^Y, X_{s,u}\rangle - \langle X_{u,t}, R_{s,u}^Y\rangle + \langle Y_u' X_{u,t}, X_{s,u}\rangle - \langle X_{u,t}, Y_s' X_{s,u}\rangle \\
&= \langle R_{u,t}^Y, X_{s,u}\rangle - \langle X_{u,t}, R_{s,u}^Y\rangle + \langle X_{u,t}, Y_u'' X_{s,u} - Y_s' X_{s,u}\rangle
\end{aligned}$$

for $0 \le s < u < t \le T$, where we used $\langle Y'_u X_{u,t}, X_{s,u}\rangle = \langle X_{u,t}, Y'_u X_{s,u}\rangle$ in the last line thanks to symmetry. With the same control ω for all functions involved as above, this gives

$$|\Xi_{s,t} - \Xi_{s,u} - \Xi_{u,t}| \le \omega(s,t)^{\frac{1}{p}+\frac{1}{r}} + \omega(s,t)^{\frac{1}{p}+\frac{1}{r}} + \omega(s,t)^{\frac{2}{p}+\frac{1}{q}} \le 3\omega(s,t)^{\vartheta}$$

with $\vartheta := \frac{2}{p} + \frac{1}{q} > 1$. So from Lemma 4.1.3 we conclude that the left-point Riemann sums converge. It remains to show that

$$\sum_{[s,t]\in\pi_n} \left(\langle Y_{s'}, X_{s,t}\rangle - \langle X_{s'}, Y_{s,t}\rangle\right) - \sum_{[s,t]\in\pi_n} \left(\langle Y_s, X_{s,t}\rangle - \langle X_s, Y_{s,t}\rangle\right) \tag{4.6}$$

tends to zero along every sequence of partitions (π_n) such that the mesh $|\pi_n|$ converges to zero. Applying the symmetry of Y', we get

$$\begin{aligned}\langle Y_{s'}, X_{s,t}\rangle - \langle X_{s'}, Y_{s,t}\rangle &- \left(\langle Y_s, X_{s,t}\rangle - \langle X_s, Y_{s,t}\rangle\right)\\ &= \langle Y'_s X_{s,s'} + R^Y_{s,s'}, X_{s,t}\rangle - \langle X_{s,s'}, Y'_s X_{s,t} + R^Y_{s,t}\rangle\\ &= \langle R^Y_{s,s'}, X_{s,t}\rangle - \langle X_{s,s'}, R^Y_{s,t}\rangle,\end{aligned}$$

and thus

$$\left|\langle Y_s, X_{s,t}\rangle - \langle X_s, Y_{s,t}\rangle - \left(\langle Y_{s'}, X_{s,t}\rangle - \langle X_{s'}, Y_{s,t}\rangle\right)\right| \le \omega(s,t)^{\vartheta}$$

with $\vartheta := \frac{1}{p} + \frac{1}{r} > 1$, where we choose the same control function ω for X and R^Y. Therefore, the properties of ω imply

$$\left|\sum_{[s,t]\in\pi_n} \left(\langle Y_{s,s'}, X_{s,t}\rangle - \langle X_{s,s'}, Y_{s,t}\rangle\right)\right| \le \sum_{[s,t]\in\pi_n} \omega(s,t)^{\vartheta} \le \max_{[s,t]\in\pi_n} \omega(s,t)^{\vartheta-1}\omega(0,T),$$

which means that (4.6) tends to zero as $|\pi_n|$ tends to zero.

If we instead assume, that X and Y are similar, we obtain

$$\begin{aligned}&\Xi_{s,t} - \Xi_{s,u} - \Xi_{u,t} = \langle X'_s Y_{s,u} + R^X_{s,u}, Y'_u X_{u,t} + R^Y_{u,t}\rangle - \langle Y_{s,u}, X_{u,t}\rangle\\ &= \langle X'_s Y_{s,u}, R^Y_{u,t}\rangle + \langle R^X_{s,u}, Y'_u X_{u,t}\rangle + \langle R^X_{s,u}, R^Y_{u,t}\rangle + \langle X'_s Y_{s,u}, Y'_u X_{u,t}\rangle - \langle Y_{s,u}, X_{u,t}\rangle\end{aligned}$$

for $0 \le s \le u \le t \le T$. The last two terms in the preceding formula can be rewritten as

$$\begin{aligned}\langle X'_s Y_{s,u}, Y'_u X_{u,t}\rangle - \langle Y_{s,u}, X_{u,t}\rangle &= \langle Y_{s,u}, (X'_t)^* Y'_u X_{u,t} - X_{u,t}\rangle\\ &= \langle Y_{s,u}, (X'_t)^*(Y'_u - Y'_t) X_{u,t}\rangle.\end{aligned}$$

Here we applied $((X'_t)^*)^{-1} = Y'_t$. Since the finite sum of control functions is again a control function, we can choose the same control function ω for X, X', R^X and Y, Y', R^Y, and obtain

$$\begin{aligned}&|\Xi_{s,t} - \Xi_{s,u} - \Xi_{u,t}|\\ &\quad\le ||X'||_\infty \omega(s,t)^{\frac{1}{p}+\frac{1}{r}} + ||Y'||_\infty \omega(s,t)^{\frac{1}{r}+\frac{1}{p}} + \omega(s,t)^{\frac{1}{r}+\frac{1}{r}} + ||X'||_\infty \omega(s,t)^{\frac{1}{q}+\frac{2}{p}}\\ &\quad\le (2||X'||_\infty + ||Y'||_\infty + \omega(0,T)^{\frac{1}{q}})\omega(s,t)^{\vartheta},\end{aligned}$$

where $||\cdot||_\infty$ denotes the supremum norm and $\vartheta := 2/p + 1/q > 1$. We therefore have shown that the left-point Riemann sums converge. It remains to prove that (4.6) goes to zero along every sequence of partitions (π_n) such that the mesh $|\pi_n|$ tends to zero. Since X and Y are similar, we observe that for $(s,t) \in \Delta_T$, and $s' \in [s,t]$

$$\begin{aligned}\langle Y_{s'}, X_{s,t}\rangle - \langle X_{s'}, Y_{s,t}\rangle - (\langle Y_s, X_{s,t}\rangle - \langle X_s, Y_{s,t}\rangle)& \\ = \langle Y'_s X_{s,s'}, X'_s Y_{s,t}\rangle + \langle Y'_s X_{s,s'}, R^X_{s,t}\rangle + \langle R^Y_{s,s'}, X'_s Y_{s,t}\rangle& \\ + \langle R^Y_{s,s'}, R^X_{s,t}\rangle - \langle Y_{s,t}, X_{s,s'}\rangle& \\ = \langle Y'_s X_{s,s'}, R^X_{s,t}\rangle + \langle R^Y_{s,s'}, X'_s Y_{s,t}\rangle + \langle R^Y_{s,s'}, R^X_{s,t}\rangle&.\end{aligned}$$

To obtain the last line, we once again use $((X'_s)^*)^{-1} = Y'_s$. Taking again the same control function ω for X, X', R^X and Y, Y', R^Y, we estimate

$$\left|\langle Y_s, X_{s,t}\rangle - \langle X_s, Y_{s,t}\rangle - (\langle Y_{s'}, X_{s,t}\rangle - \langle X_{s'}, Y_{s,t}\rangle)\right| \le C\omega(s,t)^{\vartheta},$$

where $C := ||X'||_\infty + ||Y'||_\infty + \omega(0,T)^{1/q}$ with $\vartheta := \frac{2}{p} + \frac{1}{p} > 1$. Superadditivity of ω finally gives

$$\left|\sum_{[s,t]\in\pi_n} (\langle Y_{s,s'}, X_{s,t}\rangle - \langle X_{s,s'}, Y_{s,t}\rangle)\right| \le C\omega(0,T) \max_{[s,t]\in\pi_n} \omega(s,t)^{\vartheta-1},$$

which means that (4.6) tends to zero as $|\pi_n|$ tends to zero. $\square$

Remark 4.2.5. *The proof of Theorem 4.2.4 works analogously under the assumption that X is controlled by Y and X'_t is a symmetric matrix for all $t \in [0,T]$. Moreover, if Y is controlled by X and Y'_t is an antisymmetric matrix for all $t \in [0,T]$, then an analogous result to Theorem 4.2.4 holds true for the symmetric part $\mathbb{S}_\gamma\langle X, Y\rangle$.*

In case $\gamma = \frac{1}{2}$ as in the example above, the symmetric part simplifies considerably, and therefore the preceding theorem will already imply the existence of the $\frac{1}{2}$-Stratonovich integral.

Corollary 4.2.6. *Let $X \in \mathcal{V}^p([0,T],\mathbb{R}^d)$, $Y \in \mathscr{C}^q_X$ and suppose Y'_t is a symmetric matrix for all $t \in [0,T]$ or $Y \in \mathscr{S}^q_X$. Then, the Stratonovich integral*

$$\int_0^T Y_t \circ \mathrm{d}X_t := \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \langle Y_s + \frac{1}{2} Y_{s,t}, X_{s,t}\rangle \tag{4.7}$$

exists and satisfies

$$\frac{1}{2}\text{-}\int_0^T Y_t \,\mathrm{d}X_t = \int_0^T Y_t \circ \mathrm{d}X_t = \frac{1}{2}(\langle Y_T, X_T\rangle - \langle Y_0, X_0\rangle) + \frac{1}{2}\mathbb{A}\langle X, Y\rangle.$$

Proof. By equation (4.4) we may separately treat the symmetric part $\mathbb{S}_{\frac{1}{2}}\langle X, Y\rangle$ and the antisymmetric part $\mathbb{A}_{\frac{1}{2}}\langle X, Y\rangle$ of the integral $\frac{1}{2}$-$\int_0^T Y_t \,\mathrm{d}X_t$. The existence of the antisymmetric part $\mathbb{A}_{\frac{1}{2}}\langle X, Y\rangle$ follows from Theorem 4.2.4. For the symmetric part, note that as in Example 4.1.8

$$\langle Y_s + \frac{1}{2} Y_{s,t}, X_{s,t}\rangle + \langle X_s + \frac{1}{2} X_{s,t}, Y_{s,t}\rangle = \langle Y, X\rangle_{s,t}, \quad (s,t) \in \Delta_T.$$

Therefore, $\mathbb{S}_{\frac{1}{2}}\langle X, Y\rangle$ is given by

$$\begin{aligned}\mathbb{S}_{\frac{1}{2}}\langle X, Y\rangle &= \lim_{|\pi|\to 0} \sum_{[s,t]\in\pi} \left(\langle Y_s + \frac{1}{2}Y_{s,t}, X_{s,t}\rangle + \langle X_s + \frac{1}{2}X_{s,t}, Y_{s,t}\rangle\right) \\ &= \langle Y_T, X_T\rangle - \langle X_0, Y_0\rangle. \end{aligned} \tag{4.8}$$

The proof works analogously for $Y \in \mathscr{S}^q_X$. □

The discussion of γ-Stratonovich integrals above has shown that the corresponding antisymmetric component can be treated by means of the concept of path control. In the case $\gamma \neq \frac{1}{2}$, a symmetric term is left to consider. This does not seem to be possible by means of the ideas used for the antisymmetric component. And this brings us back to Föllmer's approach. Our treatment of the symmetric part reflects the role played by *quadratic variation* in Föllmer's approach, and will therefore be strongly dependent on partition sequences. For this purpose we define the quadratic variation in the sense of Föllmer (cf. [Fö81]), and call a sequence of partitions (π_n) *increasing* if for all $[s,t] \in \pi_n$ there exist $[t_i, t_{i+1}] \in \pi_{n+1}$, $i = 1, \dots, N$, such that $[s,t] = \bigcup_{i=1}^N [t_i, t_{i+1}]$.

Definition 4.2.7. Let (π_n) be an increasing sequence of partitions such that $\lim_{n\to\infty} |\pi_n| = 0$. A continuous function $f\colon [0,T] \to \mathbb{R}$ has *quadratic variation* along (π_n) if the sequence of discrete measures on $([0,T], \mathcal{B}([0,T]))$ given by

$$\mu_n := \sum_{[s,t]\in\pi_n} |f_{s,t}|^2 \delta_s \tag{4.9}$$

converges weakly to a measure μ, where δ_s denotes the Dirac measure at $s \in [0,T]$. We write $[f]_t$ for the "distribution function" of the interval measure associated with μ. A continuous path $X = (X^1, \dots, X^d)$ has *quadratic variation* along (π_n) if (4.9) holds for all X^i and $X^i + X^j$, $1 \le i, j \le d$. In this case, we set

$$[X^i, X^j]_t := \frac{1}{2}\big([X^i + X^j]_t - [X^i]_t - [X^j]_t\big), \quad t \in [0,T].$$

Remark 4.2.8. *Since in our situation the limiting distribution function is continuous, weak convergence is equivalent to uniform convergence to the distribution function. Hence, $X = (X^1, \dots, X^d) \in C([0,T], \mathbb{R}^d)$ has quadratic variation in the sense of Föllmer if and only if*

$$[X^i, X^j]^n_t := \sum_{[u,v]\in\pi_n} X^i_{u\wedge t, v\wedge t} X^j_{u\wedge t, v\wedge t}$$

converges uniformly to $[X^i, X^j]$ in $C([0,T], \mathbb{R})$ for all $1 \le i, j \le d$, where $u \wedge t :=$ $\min\{u, t\}$*. See Lemma 2.3.22.*

Remark 4.2.9. *Let us emphasize here that* quadratic variation *should not be confused with the notion of* 2-variation*: quadratic variation depends on the choice of a partition sequence (π_n), 2-variation does not. In fact, for every continuous function $f \in C([0,T], \mathbb{R})$ there exits a sequence of partitions (π_n) with* $\lim_{n\to\infty} |\pi_n| = 0$ *such that $[f,f]_t = 0$ for all $t \in [0,T]$. See for instance Proposition 70 in [Fre83].*

The existence of quadratic variation guaranteed, Föllmer was able to prove a pathwise version of Itô's formula. In his case, the construction of the integral is closely linked to the partition sequence chosen for the quadratic variation. We will now aim at combining the techniques of controlled paths with the quadratic variation hypothesis, and derive a pathwise version of Itô's formula for paths with finite quadratic variation, in which the quadratic variation term may depend on a partition sequence, but the integral does not. As a first step, we derive the existence of γ-Stratonovich integrals for any $\gamma \in [0,1]$. To do so, we will need the following technical lemma, the easy proof of which is left to the reader.

Lemma 4.2.10. *Let $p \geq 1$, (π_n) be an increasing sequence of partitions such that $\lim_{n\to\infty} |\pi_n| = 0$, $X \in \mathcal{V}^p([0,T],\mathbb{R}^d)$ with quadratic variation along (π_n) and $Y \in \mathscr{C}^q_X$. In this case the* quadratic covariation *of X and Y exists and is given by*

$$[Y,X]_T := \lim_{n\to\infty} \sum_{[s,t]\in\pi_n} \langle X_{s,t}, Y_{s,t}\rangle = \sum_{1\leq i,j\leq d} \int_0^T Y'_t(i,j)\,\mathrm{d}^{\pi_n}[X^i,X^j]_t,$$

where $Y'_t = (Y'_t(i,j))_{1\leq i,j\leq d}$, for $0 \leq t \leq T$.

Theorem 4.2.11. *Let $X \in \mathcal{V}^p([0,T],\mathbb{R}^d)$, $Y \in \mathscr{C}^q_X$ and suppose Y'_t is a symmetric matrix for all $t \in [0,T]$ or $Y \in \mathscr{S}^q_X$. Let (π_n) be an increasing sequence of partitions such that $\lim_{n\to\infty}|\pi_n| = 0$ and X has quadratic variation along (π_n). Then for all $\gamma \in [0,1]$ the γ-$\int Y_t\,\mathrm{d}^{\pi_n}X_t$ integral exists and is given by*

$$\gamma\text{-}\int_0^T Y_t\,\mathrm{d}^{\pi_n}X_t = \int_0^T Y_t \circ \mathrm{d}X_t + \frac{1}{2}(2\gamma-1)\sum_{1\leq i,j\leq d}\int_0^T Y'_t(i,j)\,\mathrm{d}^{\pi_n}[X^i,X^j]_t,$$

where $Y'_t = (Y'_t(i,j))_{1\leq i,j\leq d}$.

Proof. Fix $\gamma \in [0,1]$. As before we split the sum as in (4.4) into its symmetric and antisymmetric part:

$$\begin{aligned}\sum_{[s,t]\in\pi_n} \langle Y_s + \gamma Y_{s,t}, X_{s,t}\rangle =& \frac{1}{2}\sum_{[s,t]\in\pi_n}\big(\langle Y_s + \gamma Y_{s,t}, X_{s,t}\rangle + \langle X_s + \gamma X_{s,t}, Y_{s,t}\rangle\big)\\ &+ \frac{1}{2}\sum_{[s,t]\in\pi_n}\big(\langle Y_s + \gamma Y_{s,t}, X_{s,t}\rangle - \langle X_s + \gamma X_{s,t}, Y_{s,t}\rangle\big).\end{aligned}$$

The second sum converges for every sequence of partitions (π_n) with $\lim_{n\to\infty}|\pi_n| = 0$ and is independent of γ thanks to Theorem 4.2.4. Taking $\gamma = 1/2$ we can apply Corollary 4.2.6 to see that

$$\frac{1}{2}\mathbb{A}\langle X,Y\rangle = \int_0^T Y_t \circ \mathrm{d}X_t - \frac{1}{2}\big(\langle X_T, Y_T\rangle - \langle X_0, Y_0\rangle\big). \tag{4.10}$$

For the symmetric part, we note for $(s,t) \in \Delta_T$

$$\begin{aligned}\langle Y_s + \gamma Y_{s,t}, X_{s,t}\rangle + \langle X_s + \gamma X_{s,t}, Y_{s,t}\rangle =& (1-\gamma)\big(\langle Y_t, X_t\rangle - \langle Y_s, X_s\rangle - \langle X_{s,t}, Y_{s,t}\rangle\big)\\ &+ \gamma\big(\langle Y_t, X_t\rangle - \langle Y_s, X_s\rangle + \langle X_{s,t}, Y_{s,t}\rangle\big)\\ =& \langle Y_t, X_t\rangle - \langle Y_s, X_s\rangle + (2\gamma-1)\langle X_{s,t}, Y_{s,t}\rangle.\end{aligned}$$

Thus the first sum reduces to

$$\frac{1}{2}\sum_{[s,t]\in\pi_n}\left(\langle Y_s+\gamma Y_{s,t}, X_{s,t}\rangle+\langle X_s+\gamma X_{s,t}, Y_{s,t}\rangle\right)$$
$$=\frac{1}{2}\left(\langle Y_T, X_T\rangle-\langle Y_0, X_0\rangle\right)+\frac{2\gamma-1}{2}\sum_{[s,t]\in\pi_n}\langle X_{s,t}, Y_{s,t}\rangle.$$

Therefore, the symmetric part converges along (π_n), and the assertion follows by (4.10) and Lemma 4.2.10.

The statement for $Y\in\mathscr{S}^q_X$ can be proven analogously. □

An application of Theorem 4.2.11 to the particular case $Y=\mathrm{D}F(X)$ for a smooth enough function F provides the classical Stratonovich formula.

Lemma 4.2.12. *Let $1\le p<3$, $X\in\mathcal{V}^p([0,T],\mathbb{R}^d)$ and $F\in C^2(\mathbb{R}^d,\mathbb{R})$. Suppose that the second derivative D^2F is α-Hölder continuous of order $\alpha>\max\{p-2,0\}$. Then the Stratonovich integral $\int\mathrm{D}F(X_t)\circ\mathrm{d}X_t$ exists and is given by*

$$\int_0^T\mathrm{D}F(X_t)\circ\mathrm{d}X_t=F(X_T)-F(X_0).$$

Proof. Let $X=(X^1,\dots,X^d)^*\in\mathcal{V}^p([0,T],\mathbb{R}^d)$ for $1\le p<3$. Then, with $r=\frac{p}{2}$ in the definition of controlled paths we easily see that $\mathrm{D}F(X)\in\mathscr{C}^p_X$. Thus by Corollary 4.2.6 the ($\frac{1}{2}$-)Stratonovich integral is well-defined and independent of the chosen sequence of partitions (π_n) along which the limit is taken. Now choose an increasing sequence of partitions (π_n) such that $\lim_{n\to\infty}|\pi_n|=0$ and $[X]_t=0$ along (π_n) for $t\in[0,T]$ (cf. Proposition 70 in [Fre83]). Applying Taylor's theorem to F, we observe that

$$\begin{aligned}F(X_T)-F(X_0)&=\frac{1}{2}\sum_{[s,t]\in\pi_n}\left((F(X_t)-F(X_s))-(F(X_s)-F(X_t))\right)\\&=\sum_{[s,t]\in\pi_n}\langle\frac{1}{2}\mathrm{D}F(X_s)+\frac{1}{2}\mathrm{D}F(X_t), X_{s,t}\rangle+\sum_{[s,t]\in\pi_n}(R(X_s,X_t)+\tilde{R}(X_s,X_t))\\&\quad+\frac{1}{4}\sum_{[s,t]\in\pi_n}\sum_{1\le i,j\le d}(\mathrm{D}^2_{i,j}(X_s)-\mathrm{D}^2_{i,j}(X_t))X^i_{s,t}X^j_{s,t},\end{aligned}$$

where $|R(x,y)|+|\tilde{R}(x,y)|\le\varphi(|x-y|)|x-y|^2$, for some increasing function $\varphi\colon[0,\infty)\to\mathbb{R}$ such that $\varphi(c)\to0$ as $c\to0$. Since X is continuous and has zero quadratic variation along (π_n), the last two terms converge to 0 as $n\to\infty$, and we obtain

$$\begin{aligned}\int_0^T\mathrm{D}F(X_t)\circ\mathrm{d}X_t&=\lim_{n\to\infty}\sum_{[s,t]\in\pi_n}\langle\mathrm{D}F(X_s)+\frac{1}{2}(\mathrm{D}F(X_t)-\mathrm{D}F(X_s)), X_{s,t}\rangle\\&=F(X_T)-F(X_0).\end{aligned}$$

□

We can now present the announced version of the pathwise formula by Föllmer (cf. [Fö81]), for which the proof reduces to combining the previous results of Theorem 4.2.11 and Lemma 4.2.12.

Corollary 4.2.13. *Let $1 \leq p < 3$, $\gamma \in [0,1]$ and (π_n) be an increasing sequence of partitions such that $\lim_{n\to\infty} |\pi_n| = 0$. Assume $F \in C^2(\mathbb{R}^d, \mathbb{R})$ with α-Hölder continuous second derivative $\mathrm{D}^2 F$ for some $\alpha > \max\{p-2, 0\}$. If $X \in \mathcal{V}^p([0,T], \mathbb{R}^d)$ has quadratic variation along (π_n), then the formula*

$$F(X_T) = F(X_0) + \gamma\text{-}\int_0^T \mathrm{D}F(X_t)\,\mathrm{d}^{\pi_n} X_t \\ - \frac{1}{2}(2\gamma - 1) \sum_{1 \leq i,j \leq d} \int_0^T \mathrm{D}^2_{i,j} F(X_s)\,\mathrm{d}^{\pi_n}[X^i, X^j]_s$$

holds.

The assumptions, that X is of finite p-variation for some $1 \leq p < 3$ and that the second derivative $\mathrm{D}^2 F$ is α-Hölder continuous for some $\alpha > \max\{p-2, 0\}$ can be considered as the price we have to pay for obtaining an integral of which the antisymmetric part does not depend on the chosen partition sequence. Föllmer [Fö81] does not need these hypotheses and especially not that the integrand is controlled by the integrator. This leads to a much bigger class of admissible integrands as we will see in the next subsection.

4.2.1. Functional Itô formula

In recent years, functional Itô calculus which extends classical calculus to functionals depending on the whole path of a stochastic process and not only on its current value, has received much attention. Based on the notion of derivatives due to Dupire [Dup09], in a series of papers Cont and Fournié [CF10a, CF10b, CF13] developed a functional Itô formula. One drawback of their approach is that the involved functional has to be defined on the space of càdlàg functions, or at least on a subspace strictly larger than $C([0,T], \mathbb{R}^d)$ (see [CR14]), and not only on $C([0,T], \mathbb{R}^d)$. In the spirit of Föllmer the paper [CF10b] provides a non-probabilistic version of a probabilistic Itô formula shown in [CF10a, CF13].

The present subsection takes reference to this program. We generalize Föllmer's pathwise Itô formula (cf. [Fö81] or Corollary 4.2.13) to twice Fréchet differentiable functionals defined on the space of continuous functions. Our functional Itô formula might be seen as the pathwise analogue to formulas stated in [Ahn97].

First we have to fix some further notation. Let (π_n) be an increasing sequence of partitions such that $\lim_{n\to\infty} |\pi_n| = 0$ and $X \in C([0,T], \mathbb{R}^d)$. We denote by X^n the piecewise linear approximation of X along (π_n), i.e.

$$X^n_t := \frac{X_{t^n_{j+1}} - X_{t^n_j}}{t^n_{j+1} - t^n_j}(t - t^n_j) + X_{t^n_j}, \quad t \in [t^n_j, t^n_{j+1}), \quad \text{for } [t^n_j, t^n_{j+1}] \in \pi_n. \tag{4.11}$$

In the following $\mathcal{C}$ stands for $C([0,T], \mathbb{R}^d)$ and $\mathcal{C}^*$ for the dual space of $\mathcal{C}$. For $X \in \mathcal{C}$ we define $X^t_s := X_s \mathbf{1}_{[0,t)}(s) + X_t \mathbf{1}_{[t,T]}(s)$ and $X^{n,t}_s := X^n_s \mathbf{1}_{[0,t)}(s) + X^n_t \mathbf{1}_{[t,T]}(s)$

for $s \in [0,T]$, where $\mathbf{1}_{[t,T]}$ is the indicator function of the interval $[t,T]$. Assume $F\colon \mathcal{C} \to \mathbb{R}$ is twice continuously (Fréchet) differentiable. That is, $\mathrm{D}F\colon \mathcal{C} \to \mathcal{C}^*$ and $\mathrm{D}^2F\colon \mathcal{C} \to \mathcal{L}(\mathcal{C},\mathcal{C}^*)$ are continuous with respect to the corresponding norms. It is well-known that $\mathcal{L}(\mathcal{C},\mathcal{C}^*)$ is isomorphic to $\mathcal{C}\otimes\mathcal{C}$. For each $t \in [0,T]$ we can understand $\mathbf{1}_{[t,T]}$ as an element of $\mathcal{C}^{**}$, the bidual of $\mathcal{C}$, and $\mathbf{1}_{[t,T]}\otimes\mathbf{1}_{[t,T]}$ as an element in $(\mathcal{C}\otimes\mathcal{C})^{**}$, respectively. Hence, $\langle \mathrm{D}F(X^s), \mathbf{1}_{[s,T]}\rangle$ and $\langle \mathrm{D}^2F(X^s), \mathbf{1}_{[s,T]} \otimes \mathbf{1}_{[s,T]}\rangle$ are well-defined as dual pairs.

Theorem 4.2.14. *Let (π_n) be an increasing sequence of partitions such that the mesh satisfies $\lim_{n\to\infty} |\pi_n| = 0$, and $X \in \mathcal{C}$ with quadratic variation along (π_n). Suppose $F\colon [0,T]\times\mathcal{C} \to \mathbb{R}$ is continuously differentiable with respect to the first argument and twice continuously differentiable with respect to the second. Furthermore, assume that $\partial_t F$ and D^2F are bounded and uniformly continuous. Then, for all $t \in [0,T]$ we have*

$$\begin{aligned} F(t,X^t) =& F(0,X^0) + \int_0^t \partial_t F(s,X^s)\,\mathrm{d}s + \sum_{i=1}^d \int_0^t \langle \mathrm{D}_i F(s,X^s), \mathbf{1}_{[s,T]}\rangle\,\mathrm{d}^{\pi_n} X^i_s \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \langle \mathrm{D}^2_{i,j}F(s,X^s), \mathbf{1}_{[s,T]}\otimes\mathbf{1}_{[s,T]}\rangle\,\mathrm{d}[X^i,X^j]_s, \end{aligned} \tag{4.12}$$

where the integral is given by

$$\begin{aligned} &\sum_{i=1}^d \int_0^t \langle \mathrm{D}_i F(s,X^s), \mathbf{1}_{[s,T]}\rangle\,\mathrm{d}^{\pi_n} X^i_s \\ &\qquad := \lim_{n\to\infty} \sum_{i=1}^d \sum_{[t^n_k,t^n_{k+1}]\in\pi_n(t)} \langle \mathrm{D}_i F(t^n_k, X^{n,t^n_k}), \eta^n_{t^n_j}\rangle X^i_{t^n_k,t^n_{k+1}}, \end{aligned}$$

where $\pi_n(t) := \{[u, v\wedge t] \,:\, [u,v]\in\pi_n,\ u<t\}$ and $\eta^n_{t^n_j}$ for $[t^n_j,t^n_{j+1}]\in\pi_n(t)$ by

$$\eta^n_{t^n_j}(s) := \frac{(s\vee t^n_{k+1}) - t^n_j}{t^n_{k+1}-t^n_k}\mathbf{1}_{[t^n_k,T]}(s), \quad s\in[0,T].$$

Proof. To increase the readability of the proof, we assume $d=1$. The general result follows analogously. Let $t\in[0,T]$ and (π_n) a sequence of partitions fulfilling the assumption of Theorem 4.2.14. We easily see that

$$\begin{aligned} F(t,X^{n,t}) - &F(0,X^{n,0}) \\ &= \sum_{[t^n_k,t^n_{k+1}]\in\pi_n(t)} \big(F(t^n_{k+1}, X^{n,t^n_{k+1}}) - F(t^n_k, X^{n,t^n_{k+1}}) \\ &\qquad\qquad + F(t^n_k, X^{n,t^n_{k+1}}) - F(t^n_k, X^{n,t^n_k})\big) \end{aligned} \tag{4.13}$$

and note that the right hand side converges uniformly to $F(t,X^t) - F(0,X^0)$ as $n\to\infty$. Applying a Taylor expansion, we obtain

$$F(t^n_{k+1}, X^{n,t^n_{k+1}}) - F(t^n_k, X^{n,t^n_{k+1}}) = \partial_t F(t^n_k, X^{n,t^n_{k+1}})(t^n_{k+1}-t^n_k) + R(t^n_k,t^n_{k+1}),$$

where one has $|R(t_k^n, t_{k+1}^n)| \le \varphi_1(|t_{k+1}^n - t_k^n|)|t_{k+1}^n - t_k^n|$, for some continuous function $\varphi_1 \colon [0,\infty) \to \mathbb{R}$ such that $\varphi_1(c) \to 0$ as $c \to 0$. With this observation and the continuity of $\partial_t F(s, X^s)$, we conclude by dominated convergence that

$$\lim_{n\to\infty} \sum_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} \left(F(t_{k+1}^n, X^{n,t_{k+1}^n}) - F(t_k^n, X^{n,t_{k+1}^n})\right) = \int_0^t \partial_t F(s, X^s)\,\mathrm{d}s.$$

For the second difference of equation (4.13), we use a second order Taylor expansion to get

$$\begin{aligned}
&\sum_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} F(t_k^n, X^{n,t_{k+1}^n}) - F(t_k^n, X^{n,t_k^n}) \\
&= \sum_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} \langle \mathrm{D}F(t_k^n, X^{n,t_k^n}), X^{n,t_{k+1}^n} - X^{n,t_k^n} \rangle \\
&\quad + \sum_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} \frac{1}{2} \langle \mathrm{D}^2 F(t_k^n, X^{n,t_k^n}), (X^{n,t_{k+1}^n} - X^{n,t_k^n}) \otimes (X^{n,t_{k+1}^n} - X^{n,t_k^n}) \rangle \\
&\quad + \sum_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} \tilde{R}(X^{n,t_k^n}, X^{n,t_{k+1}^n}) =: S_n^1(t) + S_n^2(t) + S_n^3(t),
\end{aligned}$$

where $|\tilde{R}(X^{n,t_k^n}, X^{n,t_{k+1}^n})| \le \varphi_2(\|X^{n,t_{k+1}^n} - X^{n,t_k^n}\|_\infty)\|X^{n,t_{k+1}^n} - X^{n,t_k^n}\|_\infty^2$, for some continuous function $\varphi_2 \colon \mathbb{R} \to \mathbb{R}$ such that $\varphi_2(c) \to 0$ as $c \to 0$. Since $X^{n,t_{k+1}^n} - X^{n,t_k^n} = \eta_{t_j^n}^n X_{t_k^n, t_{k+1}^n}$ and $[\cdot,\cdot]$ is bilinear, S_n^1 and S_n^2 can be rewritten by

$$\begin{aligned}
S_n^1(t) &= \sum_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} \langle \mathrm{D}F(t_k^n, X^{n,t_k^n}), \eta_{t_j^n}^n \rangle X_{t_k^n, t_{k+1}^n}, \\
S_n^2(t) &= \sum_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} \langle \mathrm{D}^2 F(t_k^n, X^{n,t_k^n}), \eta_{t_j^n}^n \otimes \eta_{t_j^n}^n \rangle X_{t_k^n, t_{k+1}^n}^2,
\end{aligned}$$

and S_n^3 estimated by

$$\sup_{t\in[0,T]} |S_n^3(t)| \le \max_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} \varphi_2(|X_{t_k^n, t_{k+1}^n}|) \sum_{[t_k^n, t_{k+1}^n] \in \pi_n(t)} X_{t_k^n, t_{k+1}^n}^2.$$

Because X has quadratic variation along (π_n) and $\varphi_2(|X_{t_k^n, t_{k+1}^n}|) \to 0$ as $n \to \infty$, $S_n^3(\cdot)$ tends uniformly to zero. To see the convergence of $S_n^2(t)$, we set $\lambda_n(s) := \max\{t_j^n : [t_j^n, t_{j+1}^n] \in \pi_n,\ t_j^n \le s\}$ and define

$$\begin{aligned}
f_n(s) &:= \langle \mathrm{D}^2 F(\lambda_{n(s)}, X^{n,\lambda_n(s)}), \eta_{\lambda_n(s)}^n \otimes \eta_{\lambda_n(s)}^n \rangle, \quad \text{and} \\
f(s) &:= \langle \mathrm{D}^2 F(s, X^s), \mathbf{1}_{[s,T]} \otimes \mathbf{1}_{[s,T]} \rangle, \quad s \in [0,T].
\end{aligned}$$

Note that (f_n) is a sequence of left-continuous functions which are uniformly bounded in n. Additionally, $\lim_{n\to\infty} f_n(s) = f(s)$ for each $s \in [0,T]$ as

$$\begin{aligned}
\lim_{n\to\infty} |f_n(s) - f(s)| \le{}& \lim_{n\to\infty} \left|\langle \mathrm{D}^2 F(\lambda_{n(s)}, X^{n,\lambda_n(s)}), \eta_{\lambda_n(s)}^n \otimes \eta_{\lambda_n(s)}^n - \mathbf{1}_{[s,T]} \otimes \mathbf{1}_{[s,T]} \rangle\right| \\
&+ \lim_{n\to\infty} \left|\langle \mathrm{D}^2 F(\lambda_{n(s)}, X^{n,\lambda_n(s)}) - \mathrm{D}^2 F(s, X^s), \mathbf{1}_{[s,T]} \otimes \mathbf{1}_{[s,T]} \rangle\right| = 0.
\end{aligned}$$

The first summand tends to zero by weak convergence of $\eta^n_{\lambda_n(s)} \otimes \eta^n_{\lambda_n(s)}$ to $\mathbf{1}_{[s,T]} \otimes \mathbf{1}_{[s,T]}$, and the second one by Lemma 3.2 in [Ahn97]. By Proposition 2.1 in [Ahn97] f is also left-continuous and so Lemma 12 in [CF10b] implies

$$\lim_{n\to\infty} S_n^2(t) = \int_0^t \langle \mathrm{D}^2 F(s, X^s), \mathbf{1}_{[s,T]} \otimes \mathbf{1}_{[s,T]} \rangle \,\mathrm{d}[X]_s.$$

In summary, we derived equation (4.12) and implicitly the convergence of $S_n^1(t)$. $\square$

It is fairly easy to see that $\langle \mathrm{D}F(t, X^t), \mathbf{1}_{[t,T]} \rangle$ is in general not controlled by a path increment of X, which we briefly illustrate by revisiting Example 2.3 in [Ahn97]. Especially, this explains why we cannot just rely on Theorem 4.2.11 to prove Theorem 4.2.14.

Example 4.2.15. Let μ be a finite signed Borel measure and let $F\colon C([0,T],\mathbb{R}) \to \mathbb{R}$ be given by

$$F(X) := \int_0^T g(s, X_s)\,\mu(\mathrm{d}s),$$

where $g(t,\cdot) \in C^2(\mathbb{R},\mathbb{R})$ for each $t \in [0,T]$ with bounded second partial derivatives $\mathrm{D}^2_{x,x} g$ and $g(\cdot, x)\colon [0,T] \to \mathbb{R}$ μ-measurable. In this case $\langle \mathbf{1}_{[t,T]}, \mathrm{D}F(X^t)\rangle$ is of course in general not controlled by a path increment of X as we see from the explicit calculation

$$\langle \mathrm{D}F(X^t), \mathbf{1}_{[t,T]} \rangle - \langle \mathrm{D}F(X^s), \mathbf{1}_{[s,T]} \rangle = -\int_s^t \mathrm{D}_x g(u, X_u)\,\mu(\mathrm{d}u), \quad 0 \le s \le t \le T.$$

5. Rough differential equations on Besov spaces

Differential equations belong to the most fundamental objects in numerous areas of mathematics gaining extra interest from their various fields of applications. A very important sub-class of classical ordinary differential equations (ODEs) are controlled ODEs, whose dynamics are given by

$$\mathrm{d}u(t) = F(u(t))\xi(t), \quad u(0) = u_0, \quad t \in \mathbb{R}, \tag{5.1}$$

where $u_0 \in \mathbb{R}^m$ is the initial condition, $u\colon \mathbb{R} \to \mathbb{R}^m$ is a continuous function, d denotes the differential operator and $F\colon \mathbb{R}^m \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a family of vector fields on $\mathbb{R}^m$. In such a dynamic $\xi\colon \mathbb{R} \to \mathbb{R}^n$ typically models the input signal and u the output.

If the signal ξ is very irregular, for instance if ξ has the regularity of white noise, equation (5.1) is called rough differential equation (RDE). Over the last two decades Lyons [Lyo98] and many other authors have developed the theory of rough paths to solve and analyze rough differential equations. A significant insight due to Lyons [Lyo98] was that the driving signal ξ must be enhanced to a "rough path" in some sense, in order to solve the RDE (5.1) and to restore the continuity of the Itô map defined by $\xi \mapsto u$ in a p-variation topology, cf. [LQ02, LCL07, FV10b]. In particular, the rough path framework allows for treating important examples as stochastic differential equations in a non-probabilistic setting. Parallel to the p-variation results, rough differential equations have been analyzed in the Hölder topology with similar tools, cf. [Fri05, FH14].

One core goal of this chapter is to unify the approach via the p-variation and the one via the Hölder topology in a common framework. To this end, we deal with rough differential equations on the very large and flexible class of Besov spaces $B^{\alpha}_{p,q}$, noting that, loosely speaking, the space of α-Hölder regular functions is given by the Besov space $B^{\alpha}_{\infty,\infty}$ and that the p-variation scale corresponds to $B^{1/p}_{p,q}$ (see [BLS06]). The results by Zähle [Zä98, Zä01, Zä05], who set up integration for functions in Sobolev–Slobodeckij spaces via fractional calculus, are covered by our results as well. In fact, Besov spaces unify numerous function spaces, including also Sobolev spaces and Bessel-potential spaces. For a comprehensive monograph we refer to Triebel [Tri10].

Due to this generality, studying solutions to the RDE (5.1) on Besov spaces is a highly interesting, but challenging problem. In a first step, provided the driving signal ξ is in $B^{\alpha-1}_{p,q}$ for $\alpha > 1/2$, $p \geq 2$, $q \geq 1$, the existence and uniqueness of a solution u to the RDE (5.1) is proven, see Theorem 5.2.1, and further it is shown that the corresponding Itô map is locally Lipschitz continuous with respect to the

Besov topology, see Theorem 5.2.2. In particular, with these results we recover the classical Young integration [You36] on Besov spaces.

In order to handle a more irregular driving signal ξ in $B^{\alpha-1}_{p,q}$ for $\alpha > 1/3$, $p \geq 3$, $q \geq 1$, the path itself has to be enhanced with an additional information, say $\pi(\vartheta, \xi)$, which always exists for a smooth path ξ and corresponds to the first iterated integral in rough path theory. In the spirit of the usual notion of geometric rough path, this leads naturally to the new definition of the space of geometric Besov rough paths $\mathcal{B}^{0;\alpha}_{p,q}$, cf. Definition 5.4.1. Starting with a smooth path ξ, it is shown that the Itô map associated to the RDE (5.1) extends continuously to the space of geometric Besov rough path, cf. Theorem 5.4.8. As a consequence there exists a unique pathwise solution to the RDE (5.1) driven by a geometric Besov rough path. Note that due to $\alpha > 1/p$ our results are restricted to continuous solutions, which seems to appear rather naturally, see Remark 5.4.10 for a discussion. Especially, for signals which are not self-similar like Brownian motion but whose regularity is determined by rare singularities, we can profit from measuring regularity in general Besov norms.

The immediate and highly non-trivial problem appearing in equation (5.1) is that the product $F(u)\xi$ is not well-defined for very irregular signals. While classical approaches as rough path theory formally integrate equation (5.1) and then give the appearing integral a meaning, the first step of our analysis is to give a direct meaning to the product in (5.1). Our analysis relies on the notion of paracontrolled distributions, very recently introduced by Gubinelli et al. [GIP12] on the Hölder spaces $B^{\alpha}_{\infty,\infty}$. Their key insight is that by applying Bony's decomposition to $F(u)\xi$ the appearing resonant term can be reduced to the resonant term $\pi(\vartheta, \xi)$ of ξ and its antiderivative ϑ, using a controlled ansatz to the solution u. The resonant term $\pi(\vartheta, \xi)$ turns out to be the necessary additional information to show the existence of a pathwise solution and corresponds to the first iterated integral in rough path theory as already mentioned above.

Generalizing the approach from [GIP12] to Besov spaces poses severe additional problems, which are solved by using the Besov space characterizations via Littlewood-Paley blocks as well as the one via the modulus of continuity. Besov spaces are a Banach algebra if and only if $p = q = \infty$ such that in general our results can only rely on pointwise multiplier theorems, Bony's decomposition and Besov embeddings. We thus need to generalize certain results in [BCD11] and [GIP12], including the commutator lemma, see Lemma 5.3.4. A second difficulty is that $u \in B^{\alpha}_{p,q}$ imposes an L^p-integrability condition on u. To overcome this problem, we localize the signal and consider a weighted Itô(-Lyons) map, both done in a way that does not change the dynamic of the RDE on a compact interval around the origin.

The paracontrolled distribution approach [GIP12] offers an extension of rough path theory to a multiparameter setting as also done by the innovative theory of regularity structures developed by Hairer [Hai14]. While Hairer's theory presumably has a much wider range of applicability, both successfully give a meaning to many stochastic partial differential equations (PDEs) like the KPZ equation [Hai13, GP15] and the dynamical Φ^4_3 equation [Hai14, CC13] just to name two. Even if the approach of Gubinelli et al. [GIP12] may not be a systematic theory as regularity structures, it comprises some advantages. The approach works with already well-studied tools like Bony's paraproduct and Littlewood-Paley theory, which leads to globally defined objects rather than the locally operating "jets" appearing in the theory of regularity structures. Since for stochastic PDEs the question about the "most suitable" function spaces seems not to be settled yet, it might be quite promising on its own to extend [GIP12] to a more general foundation as we do by working with general Besov spaces. For instance, let us refer to the very recent work of Hairer and Labbé [HL15], where the theory of regularity structures is adapted to a setting of weighted Besov spaces.

In probability theory the prototypical example of the differential equation (5.1) is a stochastic differential equation driven by a fractional Brownian motion B^H with Hurst index $H > 0$. It is well-known that the Besov regularity of such a fractional Brownian motion is $B^H_{p,\infty}$ for $p \in [1, \infty)$ and thus the results of the present chapter are applicable. For our Besov setting, an even more interesting example coming from stochastic analysis, recalling for example the Karhunen-Loève theorem, are Gaussian processes and stochastic processes given by a basis expansion with random coefficients. The Besov regularity of such random functions can be determined sharply and they are well-studied for instance when investigating the regularity of solutions for certain stochastic PDEs [CDD+12] or in non-parametric Bayesian statistics [ASS98, Boc13]. In order to make our results about RDEs accessible for these examples, we prove all the required sample path properties in Section 5.5. Especially the existence of the resonant term is provided.

This chapter is organized as follows. Section 5.1 introduces the functional analytic framework and gives some preliminary results. In Section 5.2 we recover Young integration on Besov spaces and deal with differential equations driven by paths with regularity $\alpha > 1/2$. The analytic foundation of the paracontrolled distribution approach on general Besov spaces is presented in Section 5.3. The application of the paracontrolled ansatz to rough differential equations is developed in Section 5.4 and in Section 5.5 it is used to solve certain stochastic differential equations. In Appendix A.6 some known results about Besov spaces are recalled and the proof for the local Lipschitz continuity of the Itô map is given.

5.1. Functional analytic preliminaries

For our analysis we need to recall the definition of Besov spaces, some elements of the Littlewood-Paley theory and Bony's paraproduct. For the properties of Besov spaces we refer to Triebel [Tri10]. The calculus of Bony's paraproduct is comprehensively studied by Bahouri et al. [BCD11], from which we also borrow most of our notation.

For the sake of clarification let us mention that $L^p(\mathbb{R}^d, \mathbb{R}^m)$ denotes the space of Lebesgue p-integrable functions for $p \in (0, \infty)$ and $L^\infty(\mathbb{R}^d, \mathbb{R}^m)$ denotes the space of bounded functions with the (quasi-)norms $\|\cdot\|_{L^p}$, $p \in (0, \infty]$. The space of α-Hölder continuous functions $f\colon \mathbb{R}^d \to \mathbb{R}^m$ is denoted by C^α equipped with the Hölder norm

$$\|f\|_\alpha := \sum_{|k|<\lfloor\alpha\rfloor} \|f^{(k)}\|_{L^\infty} + \sum_{|k|=\lfloor\alpha\rfloor} \sup_{x\neq y} \frac{|f^{(k)}(x) - f^{(k)}(y)|}{|x-y|^{\alpha-\lfloor\alpha\rfloor}},$$

where k denotes multi-indices with usual conventions and where $\lfloor\alpha\rfloor$ denotes the integer part of $\alpha > 0$. For operator valued functions $F\colon \mathbb{R}^m \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ we write $F \in C^n_b$, $n \in \mathbb{N}$, if F is bounded, continuous and n-times differentiable with bounded and continuous derivatives, and we use the abbreviation $C_b := C_b^0$. The first and second derivative are denoted by F' and F'', respectively, and higher derivatives by $F^{(n)}$. On the space C_b^n we introduce the norm

$$\|F\|_\infty := \sup_{x\in\mathbb{R}^m} \|F(x)\| \quad \text{and} \quad \|F\|_{C_b^n} := \|F\|_\infty + \sum_{j=1}^{n} \|F^{(n)}\|_\infty,$$

for $n \geq 1$, where $\|\cdot\|$ denotes the corresponding operator norms.

The presumably most fundamental way to define *Besov spaces* is given via the modulus of continuity of a function $f \in L^p(\mathbb{R}^d, \mathbb{R}^m)$

$$\omega_p(f,\delta) := \sup_{0<|h|<\delta} \|f(\cdot) - f(\cdot - h)\|_{L^p} \quad \text{for} \quad p, \delta > 0. \tag{5.2}$$

For $p, q \in [1,\infty]$ and $\alpha \in (0,1)$ Besov spaces are defined as

$$\begin{aligned} B^\alpha_{p,q}(\mathbb{R}^d) := B^\alpha_{p,q}(\mathbb{R}^d,\mathbb{R}^m) := \{f \in L^p(\mathbb{R}^d,\mathbb{R}^m) : \|f\|_{\omega:\alpha,p,q} < \infty\} \quad \text{with} \\ \|f\|_{\omega:\alpha,p,q} := \|f\|_{L^p} + \Big(\int_{\mathbb{R}^d} |h|^{-\alpha q}\omega_p(f,|h|)^q \frac{\mathrm{d}h}{|h|^d}\Big)^{1/q} \end{aligned} \tag{5.3}$$

and the usual modification if $q = \infty$. If $d = 1$ (and no confusion arises from the dimension m) we subsequently abbreviate $L^p := L^p(\mathbb{R},\mathbb{R}^m)$ and $B^\alpha_{p,q} := B^\alpha_{p,q}(\mathbb{R},\mathbb{R}^m)$. In $B^\alpha_{p,q}(\mathbb{R}^d)$ the regularity α is measured in the L^p-norm while q is basically a fine tuning parameter in view of the embedding $B^\alpha_{p,q_1}(\mathbb{R}^d) \subset B^\beta_{p,q_2}(\mathbb{R}^d)$ for $\beta < \alpha$ and any $q_1, q_2 \geq 1$. The classical Hölder spaces and Sobolev spaces are recovered as the special cases $B^\alpha_{\infty,\infty}(\mathbb{R}^d)$ (for non-integer α) and $B^\alpha_{2,2}(\mathbb{R}^d)$, respectively. Alternatively, Besov spaces can be characterized in terms of a Littlewood-Paley decomposition. Since our analysis mainly relies on this latter characterization, we describe it subsequently.

We write $\mathcal{S}(\mathbb{R}^d) := \mathcal{S}(\mathbb{R}^d,\mathbb{R}^m)$ for the space of Schwartz functions on $\mathbb{R}^d$ and denote its dual by $\mathcal{S}'(\mathbb{R}^d)$, which is the space of tempered distributions. For a function $f \in L^1$ the Fourier transform is defined by

$$\mathcal{F}f(z) := \int_{\mathbb{R}^d} e^{-i\langle z,x\rangle} f(x)\,\mathrm{d}x$$

and so the inverse Fourier transform is given by $\mathcal{F}^{-1}f(z) := (2\pi)^{-d}\mathcal{F}f(-z)$. If $f \in \mathcal{S}'(\mathbb{R}^d)$, then the usual generalization of the Fourier transform is considered. The Littlewood-Paley theory is based on localization in the frequency domain. Let χ and ρ be non-negative infinitely differentiable radial functions on $\mathbb{R}^d$ such that

(i) there is a ball $\mathcal{B} \subset \mathbb{R}^d$ and an annulus $\mathcal{A} \subset \mathbb{R}^d$ satisfying $\operatorname{supp}\chi \subset \mathcal{B}$ and $\operatorname{supp}\rho \subset \mathcal{A}$,

(ii) $\chi(z) + \sum_{j\geq 0} \rho(2^{-j}z) = 1$ for all $z \in \mathbb{R}^d$,

(iii) $\operatorname{supp}(\chi) \cap \operatorname{supp}(\rho(2^{-j}\cdot)) = \emptyset$ for $j \geq 1$ and $\operatorname{supp}(\rho(2^{-i}\cdot)) \cap \operatorname{supp}(\rho(2^{-j}\cdot)) = \emptyset$ for $|i-j| > 1$.

We say a pair (χ, ρ) with these properties is a *dyadic partition of unity* and throughout we use the notation

$$\rho_{-1} := \chi \quad \text{and} \quad \rho_j := \rho(2^{-j}\cdot) \quad \text{for } j \geq 0.$$

For the existence of such a partition we refer to [BCD11, Prop. 2.10]. Taking a dyadic partition of unity (χ, ρ), the *Littlewood-Paley blocks* are defined as

$$\Delta_{-1} f := \mathcal{F}^{-1}(\rho_{-1}\mathcal{F}f) \quad \text{and} \quad \Delta_j f := \mathcal{F}^{-1}(\rho_j \mathcal{F}f) \quad \text{for } j \geq 0. \tag{5.4}$$

Note that $\Delta_j f$ is a smooth function for every $j \geq -1$ and for every $f \in \mathcal{S}'(\mathbb{R}^d)$ we have

$$f = \sum_{j\geq -1} \Delta_j f := \lim_{j\to\infty} S_j f \quad \text{with} \quad S_j f := \sum_{i \leq j-1} \Delta_i f.$$

For $\alpha \in \mathbb{R}$ and $p, q \in (0, \infty]$ the Besov space can be characterized in full generality as

$$B^\alpha_{p,q}(\mathbb{R}^d, \mathbb{R}^m) = \Big\{ f \in \mathcal{S}'(\mathbb{R}^d, \mathbb{R}^m) : \|f\|_{\alpha,p,q} < \infty \Big\} \tag{5.5}$$
$$\text{with} \quad \|f\|_{\alpha,p,q} := \Big\| \big(2^{j\alpha} \|\Delta_j f\|_{L^p}\big)_{j\geq -1} \Big\|_{\ell^q}.$$

According to [Tri10, Thm. 2.5.12], the norms $\|\cdot\|_{\omega:\alpha,p,q}$ and $\|\cdot\|_{\alpha,p,q}$ are equivalent for $p, q \in (0,\infty]$ and $\alpha \in (\frac{d}{\min\{p,1\}} - d, 1)$. $B^\alpha_{p,q}(\mathbb{R}^d)$ is a quasi-Banach space and if $p, q \geq 1$, it is Banach space, cf. [Tri10, Thm. 2.3.3]. Although the (quasi-)norm $\|\cdot\|_{\alpha,p,q}$ depends on the dyadic partition (χ, ρ), different dyadic partitions of unity lead to equivalent norms.

We will frequently use the notation $A_\vartheta \lesssim B_\vartheta$, for a generic parameter ϑ, meaning that $A_\vartheta \leq C B_\vartheta$ for some constant $C > 0$ independent of ϑ. We write $A_\vartheta \sim B_\vartheta$ if $A_\vartheta \lesssim B_\vartheta$ and $B_\vartheta \lesssim A_\vartheta$. For integers $j_\vartheta, k_\vartheta \in \mathbb{Z}$ we write $j_\vartheta \lesssim k_\vartheta$ if there is some $N \in \mathbb{N}$ such that $j_\vartheta \leq k_\vartheta + N$, and $j_\vartheta \sim k_\vartheta$ if $j_\vartheta \lesssim k_\vartheta$ and $k_\vartheta \lesssim j_\vartheta$.

In view of the RDE (5.1) we need to study the product of two distributions. The standard estimate, cf. [Tri10, (24) on p. 143],

$$\|fg\|_{\alpha,p,q} \lesssim \|f\|_{\alpha,\infty,q} \|g\|_{\alpha,p,q} \tag{5.6}$$

applies only for $\alpha > 0$ and $p, q \geq 1$. However, in the context of RDEs the regularity α of the involved product will typically be negative. Given $f \in B^\alpha_{p_1,q_1}(\mathbb{R}^d)$ and $g \in B^\beta_{p_2,q_2}(\mathbb{R}^d)$, at least formally we can decompose the product fg in terms of Littlewood-Paley blocks as

$$fg = \sum_{j\geq -1} \sum_{i \geq -1} \Delta_i f \Delta_j g = T_f g + T_g f + \pi(f, g),$$

where

$$T_f g := \sum_{j \geq -1} S_{j-1} f \Delta_j g, \quad \text{and} \quad \pi(f, g) := \sum_{|i-j|\leq 1} \Delta_i f \Delta_j g. \tag{5.7}$$

We call $\pi(f, g)$ the *resonant term*. This decomposition was introduced by Bony [Bon81] and it comes with the following estimates:

Lemma 5.1.1. *Let $\alpha,\beta\in\mathbb{R}$ and $p_1,p_2,q_1,q_2\in[1,\infty]$ and suppose that*

$$\frac{1}{p}:=\frac{1}{p_1}+\frac{1}{p_2}\le 1\quad\text{and}\quad\frac{1}{q}:=\frac{1}{q_1}+\frac{1}{q_2}\le 1.$$

(i) For any $f\in L^{p_1}(\mathbb{R}^d)$ and $g\in B^{\beta}_{p_2,q}(\mathbb{R}^d)$ we have

$$\|T_fg\|_{\beta,p,q}\lesssim\|f\|_{L^{p_1}}\|g\|_{\beta,p_2,q}.$$

(ii) If $\alpha<0$, then for any $(f,g)\in B^{\alpha}_{p_1,q_1}(\mathbb{R}^d)\times B^{\beta}_{p_2,q_2}(\mathbb{R}^d)$ we have

$$\|T_fg\|_{\alpha+\beta,p,q}\quad\lesssim\quad\|f\|_{\alpha,p_1,q_1}\|g\|_{\beta,p_2,q_2}.$$

(iii) If $\alpha+\beta>0$, then for any $(f,g)\in B^{\alpha}_{p_1,q_1}(\mathbb{R}^d)\times B^{\beta}_{p_2,q_2}(\mathbb{R}^d)$ we have

$$\|\pi(f,g)\|_{\alpha+\beta,p,q}\lesssim\|f\|_{\alpha,p_1,q_1}\|g\|_{\beta,p_2,q_2}.$$

Proof. The last claim is Theorem 2.85 in [BCD11]. For the first claim and the second one we slightly generalize their Theorem 2.82. Since ρ_j is supported on 2^j times an annulus and the Fourier transform of $S_{k-1}f\Delta_kg$ is supported on 2^k times another annulus, it holds $\Delta_jT_fg=\Delta_j\sum_{j\sim k}S_{k-1}f\Delta_kg$. Using that Δ_j is a convolution with $\mathcal{F}^{-1}\rho_j=2^{jd}\mathcal{F}^{-1}\rho(2^j\cdot)$, $j\ge 0$, Young's inequality yields for any function $h\in L^p(\mathbb{R}^d)$ that $\|\Delta_jh\|_{L^p}\lesssim\|\mathcal{F}^{-1}\rho\|_{L^1}\|h\|_{L^p}$. Together with Hölder's inequality we obtain for any $j\ge -1$

$$\Big\|\Delta_j\Big(\sum_{k\ge -1}S_{k-1}f\Delta_kg\Big)\Big\|_{L^p}\lesssim\sum_{k\sim j}\|S_{k-1}f\Delta_kg\|_{L^p}\le\sum_{k\sim j}\|S_{k-1}f\|_{L^{p_1}}\|\Delta_kg\|_{L^{p_2}}.$$

Since $\lim_{k\to\infty}\|S_{k-1}f\|_{L^{p_1}}=\|f\|_{L^{p_1}}$, assertion (i) follows from

$$\begin{aligned}\|T_fg\|_{\beta,p,q}&\lesssim\Big\|2^{j\beta}\sum_{j\sim k}\|S_{k-1}f\|_{L^{p_1}}\|\Delta_kg\|_{L^{p_2}}\Big\|_{\ell^q}\\&\lesssim\|f\|_{L^{p_1}}\big\|2^{j\beta}\|\Delta_jg\|_{L^{p_2}}\big\|_{\ell^q}=\|f\|_{L^{p_1}}\|g\|_{\beta,p_2,q}.\end{aligned}$$

For (ii) another application of Hölder's inequality yields

$$\begin{aligned}\|T_fg\|_{\alpha+\beta,p,q}\lesssim&\Big\|2^{j(\alpha+\beta)}\sum_{j\sim k}\|S_{k-1}f\|_{\ell^{p_1}}\|\Delta_kg\|_{L^{p_2}}\Big\|_{\ell^q}\\\lesssim&\big\|2^{j\alpha}\|S_{j-1}f\|_{L^{p_1}}\big\|_{\ell^{q_1}}\big\|2^{j\beta}\|\Delta_jg\|_{L^{p_2}}\big\|_{\ell^{q_2}}\le\big\|2^{j\alpha}\|S_{j-1}f\|_{L^{p_1}}\big\|_{\ell^{q_1}}\|g\|_{\beta,p_2,q_2}.\end{aligned}$$

Finally, we apply Lemma A.4.3 to conclude that $(2^{j\alpha}\|S_{j-1}f\|_{L^{p_1}})_j\in\ell^{q_1}$ and that

$$\|(2^{j\alpha}\|S_{j-1}f\|_{L^{p_1}})_j\|_{\ell^{q_1}}\lesssim\|f\|_{\alpha,p_1,q_1}.\qquad\square$$

We finish this section with two elementary lemmas, which seem to be non-standard (cf. Lemma A.4 and A.10 in [GIP12] for the Hölder case). To control the norm of an antiderivative with respect to the function itself will play an import role, naturally restricted to the case $d=1$. The following lemma provides the counterpart to the well-known estimate $\|F'\|_{\alpha-1,p,q}\lesssim\|F\|_{\alpha,p,q}$ for any $F\in B^{\alpha}_{p,q}$, cf. [Tri10, Thm. 2.3.8]. For $p<\infty$ the antiderivative will in general have no finite L^p-norm such that we have to apply a weighting function to ensure integrability.

Lemma 5.1.2. *Let $p \in (1, \infty]$ and $\alpha \in (1/p, 1)$. For every $f \in B^{\alpha-1}_{p,q}(\mathbb{R})$ there exits a unique function $F\colon \mathbb{R} \to \mathbb{R}^m$ such that $F' = f$ and $F(0) = 0$. Moreover, it holds for any fixed $\psi \in C^1_b$ satisfying $C_\psi := \|\psi\|_{C^1_b} + \sum_{j,k\in\{0,1\}} \|t^j \psi^{(k)}(t)\|_{L^p} < \infty$ that*

$$\|\psi F\|_{\alpha,p,q} \lesssim C_\psi \|f\|_{\alpha-1,p,q}.$$

In particular, for any smooth ψ with $\operatorname{supp}\psi \subset [-\mathcal{T}, \mathcal{T}]$ *for some $\mathcal{T} > 0$ one has*

$$\|\psi F\|_{\alpha,p,q} \lesssim (1 \vee \mathcal{T}^2)\|\psi\|_{C^1_b} \|f\|_{\alpha-1,p,q}.$$

Proof. Since differentiating in spatial domain corresponds to multiplication in Fourier domain, we set

$$G(t) := \sum_{j\geq 0} \mathcal{F}^{-1}\Big[\frac{1}{iu}\rho_j(u)\mathcal{F}f(u)\Big](t) \quad \text{and} \quad H(t) := \int_0^t \Delta_{-1} f(s)\,\mathrm{d}s, \quad t \in \mathbb{R}.$$

Provided

$$\|\psi G\|_{\alpha,p,q} \lesssim \|\psi\|_{C^1_b} \|G\|_{\alpha,p,q} \lesssim \|\psi\|_{C^1_b} \|f\|_{\alpha-1,p,q}, \quad \|\psi H\|_{\alpha,p,q} \leq C_\psi \|f\|_{\alpha-1,p,q} \tag{5.8}$$

and noting that $B^\alpha_{p,q} \subset C(\mathbb{R})$ for $\alpha > 1/p$, the function $F := G + H - G(0)$ satisfies $F' = f$ and the asserted norm estimate. Uniqueness follows because any distribution with zero derivative is constant.

It remains to verify (5.8). Concerning G, we obtain for each Littlewood-Paley block, using $\operatorname{supp}(\rho_j) \cap \operatorname{supp}(\rho_k) = \emptyset$ for all $j, k \geq -1$ with $|k - j| > 1$,

$$\Delta_k G = \sum_{j=(k-1)\vee 0}^{k+1} \mathcal{F}^{-1}\Big[\frac{1}{iu}\rho_k(u)\rho_j(u)\mathcal{F}f(u)\Big] = \Big(\sum_{j=(k-1)\vee 0}^{k+1} \mathcal{F}^{-1}\Big[\frac{1}{iu}\rho_j(u)\Big]\Big) * \Delta_k f.$$

Using twice a substitution, we have for $j \geq 0$

$$\Big\|\mathcal{F}^{-1}\Big[\frac{\rho_j(u)}{iu}\Big]\Big\|_{L^1} = \Big\|\mathcal{F}^{-1}\Big[\frac{\rho(u)}{iu}\Big](2^j \cdot)\Big\|_{L^1} = 2^{-j}\Big\|\mathcal{F}^{-1}\Big[\frac{\rho(u)}{iu}\Big]\Big\|_{L^1}.$$

Hence, Young's inequality yields

$$\begin{aligned}\|G\|_{\alpha,p,q} &= \Big\|\big(2^{\alpha k}\|\Delta_k G\|_{L^p}\big)_k\Big\|_{\ell^q} \lesssim \Big\|\big(2^{(\alpha-1)k}\|\mathcal{F}[\rho(u)/(iu)]\|_{L^1}\|\Delta_k f\|_{L^p}\big)\Big\|_{\ell^q} \\ &\lesssim \|f\|_{\alpha-1,p,q}.\end{aligned}$$

To show the second part of (5.8), we use $\|\psi H\|_{\alpha,p,q} \lesssim \|\psi H\|_{1,p,\infty} \lesssim \|\psi H\|_{L^p} + \|(\psi H)'\|_{L^p}$ due to $\alpha < 1$. Hölder's inequality yields for $\bar{p} := \frac{p}{p-1}$ with the usual modification for $p = \infty$ that

$$\|\psi H\|_{L^p} \leq \|\Delta_{-1} f\|_{L^p} \big\|\psi(t) t^{1/\bar{p}}\big\|_{L^p} \lesssim \|(1 \vee t)\psi(t)\|_{L^p} \|f\|_{\alpha-1,p,q}$$

and similarly

$$\begin{aligned}\|(\psi H)'\|_{L^p} \leq \|\psi' H\|_{L^p} + \|\psi \Delta_{-1} f\|_{L^p} &\lesssim \|\Delta_{-1} f\|_{L^p}\big(\big\|\psi'(t) t^{1/\bar{p}}\big\|_{L^p} + \|\psi\|_\infty\big) \\ &\lesssim \big(\|\psi\|_\infty + \|(1 \vee t)\psi'(t)\|_{L^p}\big)\|f\|_{\alpha-1,p,q}. \qquad \square\end{aligned}$$

For later reference we finally investigate the *scaling operator* Λ_λ, given by $\Lambda_\lambda f(\cdot) := f(\lambda\cdot)$ for any $\lambda > 0$ and any function f, on Besov spaces.

Lemma 5.1.3. *For $\alpha \neq 0$, $p, q \geq 1$ and all $f \in B^\alpha_{p,q}(\mathbb{R}^d)$ we have*

$$\|\Lambda_\lambda f\|_{\alpha,p,q} \lesssim (1 + \lambda^\alpha |\log \lambda|)\lambda^{-d/p}\|f\|_{\alpha,p,q}.$$

Proof. Using $\Lambda_\kappa(\mathcal{F}f) = \kappa^{-d}\mathcal{F}[\Lambda_{\kappa^{-1}}f]$ for $\kappa > 0$, $f \in B^\alpha_{p,q}(\mathbb{R}^d)$, we first deduce

$$\begin{aligned}\Delta_j(\Lambda_\lambda f) &= \lambda^{-d}\mathcal{F}^{-1}[\rho_j\Lambda_{\lambda^{-1}}(\mathcal{F}f)] = \mathcal{F}^{-1}[\rho_j(\lambda\cdot)\mathcal{F}f](\lambda\cdot) \quad \text{and}\\ \Lambda_\lambda(\Delta_j f) &= \lambda^{-d}\mathcal{F}^{-1}[\rho_j(\lambda^{-1}\cdot)(\mathcal{F}f)(\lambda^{-1}\cdot)] = \mathcal{F}^{-1}[\rho_j(\lambda^{-1}\cdot)\mathcal{F}[\Lambda_\lambda f]]\end{aligned}$$

for all $\lambda > 0$. For $j \geq 0$ the Fourier transform of $\Lambda_\lambda(\Delta_j f)$ is consequently supported in $\lambda 2^j\mathcal{A}$, where $\mathcal{A}$ is the annulus containing the support of ρ, and we have $\Delta_k(\Lambda_\lambda\Delta_j f) \neq 0$ only if $2^k \sim \lambda 2^j$. Together with $\|\Delta_k f\|_{L^p} \leq \|\mathcal{F}^{-1}\rho_k\|_{L^1}\|f\|_{L^p} \lesssim \|f\|_{L^p}$ by Young's inequality we obtain

$$\|\Delta_k\Lambda_\lambda f\|_{L^p} \leq \sum_{j:2^k\sim\lambda 2^j}\|\Delta_k\Lambda_\lambda(\Delta_j f)\|_{L^p} \lesssim \lambda^{-d/p}\sum_{j:2^k\sim\lambda 2^j}\|\Delta_j f\|_{L^p} \quad \text{for } k \geq 0.$$

Applying again Young's inequality to the sequences $a := (\mathbf{1}_{[-|\log\lambda|,|\log\lambda|]}(k))_k$ and $(2^{j\alpha}\|\Delta_j f\|_{L^p})_j$, we infer

$$\begin{aligned}\Big\|\big(2^{k\alpha}\|\Delta_k\Lambda_\lambda f\|_{L^p}\big)_{k\geq 0}\Big\|_{\ell^q} &\lesssim \lambda^{-d/p}\Big\|\Big(\sum_{j:2^k\sim\lambda 2^j}\lambda^\alpha 2^{j\alpha}\|\Delta_j f\|_{L^p}\Big)_{k\geq 0}\Big\|_{\ell^q}\\ &\lesssim \lambda^{\alpha-d/p}\|a\|_{\ell^1}\|f\|_{\alpha,p,q} \lesssim |\log\lambda|\lambda^{\alpha-d/p}\|f\|_{\alpha,p,q}.\end{aligned}$$

Finally, we obtain analogously for $k = -1$ that

$$\begin{aligned}\|\Delta_{-1}\Lambda_\lambda f\|_{L^p} &\lesssim \lambda^{-d/p}\sum_{j:\lambda 2^j\lesssim 1}\|\Delta_j f\|_{L^p}\\ &\lesssim \lambda^{-d/p}\|f\|_{\alpha,p,q}\sum_{j:\lambda 2^j\lesssim 1}2^{-\alpha j} \lesssim (1+\lambda^\alpha)\lambda^{-d/p}\|f\|_{\alpha,p,q}.\end{aligned}$$ □

5.2. Young integration revisited

In the present section we start to consider the differential equation (5.1), which was given by

$$\mathrm{d}u(t) = F(u(t))\xi(t), \quad u(0) = u_0, \quad t \in \mathbb{R},$$

where $u_0 \in \mathbb{R}^m$, $u\colon \mathbb{R} \to \mathbb{R}^m$ is a continuous function and $F\colon \mathbb{R}^m \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Assuming our driving signal $\xi\colon \mathbb{R} \to \mathbb{R}^n$ is smooth enough, the differential equation (5.1) is well-defined and can be equivalently written in its integral form

$$u(t) = u_0 + \int_0^t F(u(s))\xi(s)\,\mathrm{d}s, \quad t \in [0,\infty), \tag{5.9}$$

and analogously for $t \in (-\infty, 0)$. According to Young [You36], the involved integral can be defined as limit of Riemann sums as long as the driving signal ξ is the derivative of a path ϑ which is of finite p-variation for $p < 2$. Then, equation (5.9) admits a unique solution on every bounded interval $[-\mathcal{T}, \mathcal{T}] \subset \mathbb{R}$ if $F \in C_b^2$ (see modern books as [LCL07, Theorem 1.28] or [Lej09, Theorem 1]). This result was first proven by Lyons [Lyo94] using a Picard iteration. The case of a $1/p$-Hölder continuous driving path ϑ was treated by Ruzmaikina [Ruz00]. Since then it is still of great interest to find new approaches to (5.9): Gubinelli [Gub04] has introduced the notion of controlled paths, Davie [Dav07] has shown the convergence of an Euler scheme, Hu [HN07] have used techniques from fractional calculus and Lejay [Lej10] has developed a simple approach similar to [Ruz00].

In this section we recover the analogous results on Besov spaces with a special focus on the situation when F is a linear functional. For a discussion of the importance of linear RDEs we refer to Coutin and Lejay [CL14] and references therein.

We first note that the function $F(u)$ inherits its regularity from the regularity of u. More precisely, [BCD11, Thm. 2.87] shows for $u \in B^{\alpha}_{p,q}$ satisfying $\|u\|_\infty < \infty$ and a family of sufficient regular vector fields F with $F(0) = 0$ (or $p = \infty$) that

$$\|F(u)\|_{\alpha,p,q} \lesssim \Big(\sum_{k=1}^{\lceil \alpha \rceil} \sup_{|x| \le \|u\|_\infty} \|F^{(k)}(x)\| \Big) \|u\|_{\alpha,p,q} \lesssim \|F\|_{C_b^{\lceil \alpha \rceil}} \|u\|_{\alpha,p,q}, \tag{5.10}$$

denoting the smallest integer larger or equal than $\alpha > 0$ by $\lceil \alpha \rceil$ and provided the norms on the right-hand side are finite. If the product $F(u)\xi$ is regular enough, we can understand the differential equation (5.1) in its integral form (5.9) where the integral is given by the antiderivative of the product, i.e.

$$\mathrm{d}\Big(\int_0^t F(u(s))\xi(s)\,\mathrm{d}s \Big) = F(u(t))\xi(t) \quad \text{and} \quad \int_0^0 F(u(s))\xi(s)\,\mathrm{d}s = 0.$$

In view of Lemma 5.1.2 the solution u of (5.1) cannot be expected to be contained in $B^{\alpha}_{p,q}$. Therefore, we consider instead a localized version of the differential equation. Alternatively, the solution of the RDE (5.1) could be studied in homogenous or weighted Besov spaces, which can only lead to very similar results. In order to provide our results in the most commonly used notion of Besov spaces, we focus on localized equations. We impose the following standing assumption:

Assumption 1. Let $\varphi\colon \mathbb{R} \to \mathbb{R}_+$ be fixed smooth function with support $[-2, 2]$ and equal to 1 on $[-1, 1]$. Denote $\varphi_{\mathcal{T}}(x) := \varphi(x/\mathcal{T})$ for $\mathcal{T} > 0$.

Theorem 5.2.1. *Let $\mathcal{T} > 0$, $\alpha \in (1/2, 1]$ and assume that $\xi \in B^{\alpha-1}_{p,q}$ for $p \in [2, \infty]$ and $q \in [1, \infty]$. If $F\colon \mathbb{R}^m \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a linear mapping, then for every $u_0 \in \mathbb{R}^d$ there exists a unique global solution $u \in B^{\alpha}_{p,q}$ to the Cauchy problem*

$$u(t) = \varphi_{\mathcal{T}}(t) u_0 + \varphi_{\mathcal{T}}(t) \int_0^t F(u(s))\xi(s)\,\mathrm{d}s, \quad t \in \mathbb{R}, \tag{5.11}$$

with the usual convention for $t < 0$. This result extends to nonlinear $F \in C_b^2$ if $p = \infty$.

Proof. Step 1: First we establish a contraction principle under the assumption that $\|F'\|_{C_b^1}$ is sufficiently small. Without loss of generality we may assume $u_0 = 0$. Following a fixed point argumentation, we consider the solution map

$$\Phi\colon B_{p,q}^{\alpha} \to B_{p,q}^{\alpha}, \quad u \mapsto \tilde{u} := \varphi_{\mathcal{T}} \int_0^{\cdot} F(u(s))\xi(s)\,\mathrm{d}s, \quad t \in \mathbb{R}.$$

In order to verify that Φ is indeed well-defined, we use Lemma 5.1.2 to observe

$$\|\varphi_{\mathcal{T}} F\|_{\alpha,p,q} \lesssim (1 \vee \mathcal{T}^2)(1 \vee \mathcal{T}^{-1})\|\varphi\|_{C_b^1}\|f\|_{\alpha-1,p,q} \lesssim C_{\mathcal{T},\varphi}\|f\|_{\alpha-1,p,q},$$

where $C_{\mathcal{T},\varphi} := (\mathcal{T}^{-1} \vee \mathcal{T}^2)\|\varphi\|_{C_b^1}$, for any given $f \in B_{p,q}^{\alpha-1}$ with $\mathrm{d}F = f$ and $F(0) = 0$. We thus have

$$\|\Phi(u)\|_{\alpha,p,q} = \Big\|\varphi_{\mathcal{T}}\Big(\int_0^{\cdot} F(u(s))\xi(s)\,\mathrm{d}s\Big)\Big\|_{\alpha,p,q} \lesssim C_{\mathcal{T},\varphi}\|F(u)\xi\|_{\alpha-1,p,q}.$$

Applying Bony's decomposition, the Besov embedding $B_{p/2,q}^{2\alpha-1} \subset B_{p,q}^{\alpha-1}$ (cf. [Tri10, Thm. 2.7.1]) for $p > 1/\alpha$ and Lemma 5.1.1, we obtain

$$\begin{aligned}
\|\Phi(u)\|_{\alpha,p,q} &\lesssim C_{\mathcal{T},\varphi}(\|T_{F(u)}\xi\|_{\alpha-1,p,q} + \|\pi(F(u),\xi)\|_{2\alpha-1,p/2,q} + \|T_{\xi}(F(u))\|_{\alpha-1,p,q}) \\
&\lesssim C_{\mathcal{T},\varphi}(\|F(u)\|_{\infty}\|\xi\|_{\alpha-1,p,q} \\
&\qquad + \|F(u)\|_{\alpha,p,2q}\|\xi\|_{\alpha-1,p,2q} + \|\xi\|_{\alpha-1,p,q}\|F(u)\|_{0,\infty,\infty}\big).
\end{aligned}$$

Using the embeddings $B_{p,q}^{\alpha} \subset L^{\infty}$ and $B_{p,q}^{\alpha} \subset B_{\infty,\infty}^{0}$ for $\alpha > 1/p$ and (5.10), we deduce that

$$\|\Phi(u)\|_{\alpha,p,q} \lesssim C_{\mathcal{T},\varphi}\|F'\|_{\infty}\|\xi\|_{\alpha-1,p,q}\|u\|_{\alpha,p,q}. \tag{5.12}$$

To apply Banach's fixed point theorem, it remains to show that Φ is a contraction. For $u, \tilde{u} \in B_{p,q}^{\alpha}$ Lemma 5.1.2 again yields

$$\begin{aligned}
\|\Phi(u) - \Phi(\tilde{u})\|_{\alpha,p,q} &\lesssim C_{\mathcal{T},\varphi}\|\big(F(u) - F(\tilde{u})\big)\xi\|_{\alpha-1,p,q} \\
&\lesssim C_{\mathcal{T},\varphi}\int_0^1 \|F'(u + t(u-\tilde{u}))(u-\tilde{u})\xi\|_{\alpha-1,p,q}\,\mathrm{d}t.
\end{aligned}$$

Denoting by $v_t := F'(u + t(u-\tilde{u}))(u-\tilde{u})$, we conclude as above

$$\begin{aligned}
&\|\Phi(u) - \Phi(\tilde{u})\|_{\alpha,p,q} \\
&\quad\lesssim C_{\mathcal{T},\varphi}\int_0^1 \big(\|T_{v_t}\xi\|_{\alpha-1,p,q} + \|\pi(v_t,\xi)\|_{2\alpha-1,p/2,q/2} + \|T_{\xi}v_t\|_{\alpha-1,p,q}\big)\,\mathrm{d}t \\
&\quad\lesssim C_{\mathcal{T},\varphi}\int_0^1 \big(\|v_t\|_{\alpha,p,q}\|\xi\|_{\alpha-1,p,q}\big)\,\mathrm{d}t.
\end{aligned}$$

By the standard estimate (5.6), we obtain

$$\|\Phi(u) - \Phi(\tilde{u})\|_{\alpha,p,q} \lesssim C_{\mathcal{T},\varphi}\Big(\int_0^1 \|F'(u + t(\tilde{u}-u))\|_{\alpha,\infty,q}\,\mathrm{d}t\Big)\|\xi\|_{\alpha-1,p,q}\|u-\tilde{u}\|_{\alpha,p,q}. \tag{5.13}$$

Hence, if F is linear and $\|F'\|_\infty$ is small enough, Φ is a contraction. Provided $p=\infty$ and $F\in C_b^2$, it suffices if $\|F'\|_{C_b^1}$ is sufficiently small:

$$\|\Phi(u)-\Phi(\tilde u)\|_{\alpha,p,q}\lesssim C_{\mathcal{T},\varphi}\|F'\|_{C_b^1}(\|u\|_{\alpha,\infty,q}+\|\tilde u\|_{\alpha,\infty,q})\|\xi\|_{\alpha-1,\infty,q}\|u-\tilde u\|_{\alpha,\infty,q}. \quad (5.14)$$

Step 2: In order to ensure that $\|F'\|_{C_b^1}$ is small enough, we scale ξ as follows: For some fixed $\epsilon\in(0,\alpha-1/p)$ and for some $\lambda\in(0,1)$ to be chosen later we set

$$\xi^\lambda := \lambda^{1-\alpha+1/p+\epsilon}\Lambda_\lambda\xi, \quad (5.15)$$

where we recall the scaling operator $\Lambda_\lambda f=f(\lambda\cdot)$ for $f\in\mathcal{S}'$. Lemma 5.1.3 yields

$$\|\xi^\lambda\|_{\alpha-1,p,q}=\lambda^{1-\alpha+1/p+\epsilon}\|\Lambda_\lambda\xi\|_{\alpha-1,p,q}\lesssim(\lambda^\epsilon|\log\lambda|+\lambda^{1-\alpha+\epsilon})\|\xi\|_{\alpha-1,p,q}\le\|\xi\|_{\alpha-1,p,q}.$$

For $\lambda>0$ sufficiently small Step 1 provides a unique global solution $u^\lambda\in B^\alpha_{p,q}$ to the (localized) differential equation

$$u^\lambda(t)=\varphi_{\mathcal{T}}(t)u_0+\varphi_{\mathcal{T}}(t)\int_0^t\lambda^{\alpha-1/p-\epsilon}F(u^\lambda(s))\xi^\lambda(s)\,\mathrm{d}s, \quad (5.16)$$

for all $u_0\in\mathbb{R}$. Setting now $u:=\Lambda_{\lambda^{-1}}u^\lambda$, we have constructed a unique solution to

$$u(t)=\Lambda_{\lambda^{-1}}u^\lambda(t)=\varphi_{\lambda\mathcal{T}}(t)u_0+\varphi_{\lambda\mathcal{T}}(t)\int_0^tF(u(s))\xi(s)\,\mathrm{d}s,$$

which coincides with (5.11) on $[-\lambda\mathcal{T},\lambda\mathcal{T}]$.

Step 3: Since the choice of λ does not depend on u_0, we can iteratively apply Step 2 on intervals of length $2\lambda\mathcal{T}$ to construct a unique global solution $u\in B^\alpha_{p,q}$ to equation (5.11). □

In this simple setting it turns out that the Itô map S defined by

$$S\colon\mathbb{R}^d\times B^{\alpha-1}_{p,q}\to B^\alpha_{p,q}\quad\text{via}\quad(u_0,\xi)\mapsto u, \quad (5.17)$$

where u denotes the solution of the (localized) Cauchy problem (5.11), is a locally Lipschitz continuous map with respect to the Besov norm.

Theorem 5.2.2. *Let $\alpha\in(1/2,1]$, $q\in[1,\infty]$ and $F\colon\mathbb{R}^m\to\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$. If either F is a linear mapping and $p\in[2,\infty]$ or $F\in C_b^2$ and $p=\infty$, then the Itô map S given by (5.17) is locally Lipschitz continuous.*

Proof. Let $u_0^i\in\mathbb{R}^d$, $\xi^i\in B^{\alpha-1}_{p,q}$ be such that $\|\xi^i\|_{\alpha,p,q}\le R$ and $|u_0^i|\le R$ for some $R>0$ and denote by u^i the unique solution to corresponding Cauchy problems (5.11) for $i=1,2$, which exists thanks to Theorem 5.2.1. In order to avoid repetition, we just consider a linear mapping F. The non-linear case works analogously.

Step 1: Suppose that $\|F'\|_\infty$ is sufficiently small. Recalling $C_{\mathcal{T},\varphi}=(\mathcal{T}^{-1}\vee\mathcal{T}^2)\|\varphi\|_{C_b^1}$, we deduce similarly to (5.12) that

$$\|u^i\|_{\alpha,p,q}\lesssim\|\varphi_{\mathcal{T}}\|_{\alpha,p,q}|u_0^i|+C_{\mathcal{T},\varphi}\|F'\|_\infty\|\xi^i\|_{\alpha-1,p,q}\|u^i\|_{\alpha,p,q},$$

which, provided $\|F'\|_\infty$ is small enough, depending only on R, φ and $\mathcal{T}$, leads to

$$\|u^i\|_{\alpha,p,q} \lesssim \|\varphi_{\mathcal{T}}\|_{\alpha,p,q} R, \quad \text{for} \quad i = 1,2.$$

For the difference $u^1 - u^2$ we have

$$\begin{aligned}
&\|u^1 - u^2\|_{\alpha,p,q} \\
&\quad \le \|\varphi_{\mathcal{T}}(u_0^1 - u_0^2)\|_{\alpha,p,q} + \Big\|\varphi_{\mathcal{T}}\Big(\int_0^{\cdot} F(u^1(s))\xi^1(s)\,\mathrm{d}s - \int_0^{\cdot} F(u^2(s))\xi^2(s)\,\mathrm{d}s\Big)\Big\|_{\alpha,p,q} \\
&\quad \lesssim \|\varphi_{\mathcal{T}}\|_{\alpha,p,q}|u_0^1 - u_0^2| + \Big\|\varphi_{\mathcal{T}}\int_0^{\cdot} (F(u^1(s)) - F(u^2(s)))\xi^1(s)\,\mathrm{d}s\Big\|_{\alpha,p,q} \\
&\qquad + C_{\mathcal{T},\varphi}\|F(u^2)(\xi^1 - \xi^2)\|_{\alpha-1,p,q}.
\end{aligned}$$

The second term can be estimated as in (5.13) and for the last one Bony's decomposition, Lemma 5.1.1 and (5.10) yield

$$\|F(u^2)(\xi^1-\xi^2)\|_{\alpha-1,p,q} \lesssim \|F(u^2)\|_{\alpha,p,q}\|\xi^1-\xi^2\|_{\alpha-1,p,q} \le \|F'\|_\infty\|u^2\|_{\alpha,p,q}\|\xi^1-\xi^2\|_{\alpha-1,p,q}.$$

Therefore, we can combine the above estimates to

$$\begin{aligned}
\|u^1 - u^2\|_{\alpha,p,q} \lesssim & C_{\mathcal{T},\varphi}\Big(|u_0^1 - u_0^2| + \|\varphi_{\mathcal{T}}\|_{\alpha,p,q}\|F'\|_\infty R\|\xi^1 - \xi^2\|_{\alpha-1,p,q} \\
&+ \Big(\int_0^1 \|(F'(u^1 + t(u^2 - u^1))\|_{\alpha-1,\infty,q}\,\mathrm{d}t\Big)R\|u^1 - u^2\|_{\alpha,p,q}\Big).
\end{aligned}$$

If F is linear with sufficiently small $\|F\|_{C_b^1}$, we obtain the desired estimate by rearranging:

$$\|u^1 - u^2\|_{\alpha,p,q} \lesssim C_{\mathcal{T},\varphi}(|u_0^1 - u_0^2| + \|\varphi_{\mathcal{T}}\|_{\alpha,p,q}\|F\|_{C_b^1} R\|\xi^1 - \xi^2\|_{\alpha-1,p,q}).$$

Step 2: The assumption on $\|F'\|_\infty$ can be translated to an assumption on $\mathcal{T}$ using the same scaling argument as in Step 2 in the proof of Theorem 5.2.1. More precisely, we define $\xi^{\lambda,1}$ and $\xi^{\lambda,2}$ for $\lambda > 0$ as in (5.15) and note $\|\xi^{\lambda,i}\|_{\alpha,p,q} \lesssim R$ for $i = 1,2$. Therefore, for sufficiently small λ there exists a unique solution $u^{\lambda,i}$ to (5.16) for $i = 1,2$. Setting again $u^i := \Lambda_{\lambda^{-1}} u^\lambda$ and applying twice Lemma 5.1.3 together with Step 1 gives

$$\begin{aligned}
&\|u^1 - u^2\|_{\alpha,p,q} \\
&\quad \lesssim (1 + \lambda^{-\alpha}|\log\lambda^{-1}|)\lambda^{1/p}\|u^{\lambda,1} - u^{\lambda,2}\|_{\alpha,p,q} \\
&\quad \lesssim C_{\mathcal{T},\varphi}(1 + \lambda^{-\alpha}|\log\lambda^{-1}|)\lambda^{1/p}(|u_0^1 - u_0^2| + \|\varphi_{\mathcal{T}}\|_{\alpha,p,q}\|F'\|_\infty R\|\xi^{\lambda,1} - \xi^{\lambda,2}\|_{\alpha-1,p,q}) \\
&\quad \lesssim C_{\mathcal{T},\varphi}(1 + \lambda^{-\alpha}|\log\lambda^{-1}|)\lambda^{1/p}(|u_0^1 - u_0^2| + \|\varphi_{\mathcal{T}}\|_{\alpha,p,q}\|F'\|_\infty R\|\xi^1 - \xi^2\|_{\alpha-1,p,q}).
\end{aligned}$$

In conclusion, the Itô map is locally Lipschitz continuous given $\mathcal{T} > 0$ is sufficiently small because u^i is a solution to

$$u^i(t) = \varphi_{\lambda\mathcal{T}}(t)u_0^i + \varphi_{\lambda\mathcal{T}}(t)\int_0^t F(u^i(s))\xi^i(s)\,\mathrm{d}s, \quad i = 1,2.$$

Step 3: The local Lipschitz continuity for arbitrary $\mathcal{T}$ follows by a pasting argument. For this purpose choose a partition of unity $(\mu_j)_{j\in\mathbb{Z}} \subset C_b^\infty$ satisfying

$\mu_j(t_j+\epsilon) = 1$, $\epsilon \in [-\frac{1}{2}\lambda\mathcal{T}, \frac{1}{2}\lambda\mathcal{T}]$, for anchor points $t_j \in \mathbb{R}$ with $t_0 = 0$ and $|t_j - t_{j-1}| \le \lambda\mathcal{T}/2$ and fulfilling

$$|\mathrm{supp}\mu_j| := \sup\{|x-y| \;:\; x, y \in \mathrm{supp}\mu_j\} \le \lambda\mathcal{T} \quad \text{and} \quad \sum_{j\in\mathbb{Z}} \mu_j(x) = 1 \quad \text{for all } x \in \mathbb{R}.$$

Since the u^i for $i = 1,2$ have compact support, there is some $N \in \mathbb{N}$ such that one has, using (5.6),

$$\|u^1 - u^2\|_{\alpha,p,q} \le \sum_{j=-N}^{N} \|\mu_j(u^1-u^2)\|_{\alpha,p,q} \lesssim \sum_{j=-N}^{N} \|\mu_j\|_{C_b^1}\|u_j^1 - u_j^2\|_{\alpha,p,q},$$

where u_j^i is the unique solution to

$$u_j^i(t) = \varphi_{\lambda\mathcal{T}}(t-t_j)u_{t_j}^i + \varphi_{\lambda\mathcal{T}}(t-t_j)\int_{t_j}^{t} F(u_j^i(s))\xi^i(s)\,\mathrm{d}s$$

with initial condition $u_{t_j}^i := u^i(t_j)$ for $i = 1,2$. Noting that $|u_{t_j}^1 - u_{t_j}^2| \lesssim \|u_{j-1}^1 - u_{j-1}^2\|_{\alpha,p,q}$ for $j \ge 1$ and similarly for negative j, Step 2 yields

$$\|u^1 - u^2\|_{\alpha,p,q} \lesssim C_{\mathcal{T},\varphi}(|u_0^1 - u_0^2| + \|\varphi_{\mathcal{T}}\|_{\alpha,p,q}\|F'\|_\infty R\|\xi^1 - \xi^2\|_{\alpha-1,p,q}). \qquad \square$$

To extend these results to nonlinear functions F for $p < \infty$ and to less regular driving signals ξ, more precisely $\xi \in B_{p,q}^{\alpha-1}$ for $\alpha \in (1/3, 1/2)$, is the aim of the following two sections.

5.3. Linearization and commutator estimate

In order to deal with more irregular driving signals $\xi \in B_{p,q}^{\alpha-1}$, we shall apply Bony's decomposition to rigorously define the product $F(u)\xi$, which appears in the RDE (5.1). Let us first formally decompose $F(u)\xi$ and analyze the Besov regularity of the different terms as follows

$$F(u)\xi = \underbrace{T_{F(u)}\xi}_{\in B_{p,q}^{\alpha-1}} + \underbrace{\pi(F(u),\xi)}_{\in B_{p/2,q/2}^{2\alpha-1} \text{ if } 2\alpha-1>0} + \underbrace{T_\xi(F(u))}_{\in B_{p/2,q/2}^{2\alpha-1}}. \tag{5.18}$$

The first term $T_{F(u)}\xi$ is in $B_{p,q}^{\alpha-1}$ due to Lemma 5.1.1 and the boundedness of F. The regularity of the third term $T_\xi F(u) \in B_{p/2,q/2}^{2\alpha-1}$ for $\alpha < 1$ can also be deduced from Lemma 5.1.1 since naturally the solution u has regularity $B_{p,q}^\alpha$ and thus $F(u) \in B_{p,q}^\alpha$ by (5.10). The regularity estimate of the resonant term can be applied only if $2\alpha - 1 > 0$. This is the main reason, why it was possible for $\alpha \in (1/2, 1]$ to show the existence of a solution to the (localized) RDE (5.1) in Section 5.2 without taking any additional information about ξ into account. However, this high Besov regularity assumption on ξ is violated in most of the basic examples from stochastic analysis as for instance for stochastic differential equations driven by Brownian motion or martingales. The aim of this section is to reduce the resonant term $\pi(F(u),\xi)$ to $\pi(u,\xi)$:

Proposition 5.3.1. *Let $\alpha \in (\frac{1}{3}, \frac{1}{2})$, $p \in [3, \infty]$ and $F \in C_b^{2+\gamma}(\mathbb{R})$ for some $\gamma \in (0,1]$ satisfying $F(0) = 0$. Then there is a map $\Pi_F \colon B^{\alpha}_{p,\infty}(\mathbb{R}) \times B^{\alpha-1}_{p,\infty}(\mathbb{R}) \to B^{3\alpha-1}_{p/3,\infty}(\mathbb{R})$ such that for any $u \in B^{\alpha}_{p,\infty}(\mathbb{R})$ and $\xi \in B^{\alpha-1}_{p,\infty}(\mathbb{R})$ we have*

$$\pi(F(u), \xi) = F'(u)\pi(u, \xi) + \Pi_F(u, \xi) \tag{5.19}$$

with

$$\|\Pi_F(u, \xi)\|_{3\alpha-1,p/3,\infty} \lesssim \|F\|_{C_b^2} \|u\|^2_{\alpha,p,\infty} \|\xi\|_{\alpha-1,p,\infty}. \tag{5.20}$$

Moreover, Π_F is locally Hölder continuous satisfying for any $u^1, u^2 \in B^{\alpha}_{p,q}(\mathbb{R})$ and $\xi^1, \xi^2 \in B^{\alpha-1}_{p,q}(\mathbb{R})$

$$\begin{aligned} &\|\Pi_F(u^1, \xi^1) - \Pi_F(u^2, \xi^2)\|_{3\alpha-1,p/3,\infty} \\ &\quad \lesssim \|F\|_{C_b^{2+\gamma}} C(u^1, u^2, \xi^1, \xi^2)\Big(\|u^1 - u^2\|^{\gamma}_{\infty} + \|u^1 - u^2\|_{\alpha,p,\infty} + \|\xi^1 - \xi^2\|_{\alpha-1,p,\infty}\Big) \end{aligned}$$

where

$$\begin{aligned} C(u^1, u^2, \xi^1, \xi^2) :=& \|u^1\|^2_{\alpha,p,\infty} \wedge \|u^2\|^2_{\alpha,p,\infty} + (\|u^1\|_{\alpha,p,\infty} + \|u^2\|_{\alpha,p,\infty}) \\ &\times (1 + \|\xi^1\|_{\alpha-1,p,\infty} \wedge \|\xi^2\|_{\alpha-1,p,\infty}). \end{aligned}$$

As we will see in the next section, it suffices to consider only $q = \infty$ in Proposition 5.3.1. Taking into account the embedding $B^{\alpha}_{p,q} \subset B^{\alpha}_{p,\infty}$ for any $q \in [1, \infty]$, this case corresponds to the weakest Besov norm for fixed α and p.

In order to prove this proposition, we need the subsequent lemmas. As the first step, we show the following paralinearization result, which is a slight generalization of Theorem 2.92 in [BCD11]. Our proof is inspired by [GIP12, Lem. 2.6] and relies on the characterization of Besov spaces via the modulus of continuity. We obtain that the composition $F(u)$ can be written as a paraproduct of $F'(u)$ and u up to some more regular remainder.

Lemma 5.3.2. *Let $0 < \beta \leq \alpha < 1$ and $F \in C_b^{1+\beta/\alpha}(\mathbb{R}^m)$. Let $p \geq \beta/\alpha + 1$ and define $p' := \alpha p/(\alpha + \beta)$. Then for any $g \in B^{\alpha}_{p,\infty}(\mathbb{R}^d) \cap L^{\infty}(\mathbb{R}^d)$ there is some $R_F(g) \in B^{\alpha+\beta}_{p',\infty}(\mathbb{R}^d)$ satisfying*

$$F(g) - F(0) = T_{F'(g)} g + R_F(g) \quad \textit{and} \quad \|R_F(g)\|_{\alpha+\beta,p',\infty} \lesssim \|F\|^{2-\beta/\alpha}_{C_b^{1+\beta/\alpha}} \|g\|^{1+\beta/\alpha}_{\alpha,p,\infty}.$$

Moreover, if $F \in C_b^{2+\gamma}$ for some $\gamma \in (0,1]$ and if $p > 2 \vee 1/\alpha$ then the map

$$R_F \colon B^{\alpha}_{p,\infty}(\mathbb{R}^d) \cap L^{\infty}(\mathbb{R}^d) \to B^{2\alpha}_{p/2,\infty}(\mathbb{R}^d)$$

is locally Hölder continuous with

$$\begin{aligned} &\|R_F(g) - R_F(h)\|_{2\alpha,p/2,\infty} \\ &\lesssim \|F\|_{C_b^{2+\gamma}} \Big(\|g\|^2_{\alpha,p,\infty} \wedge \|h\|^2_{\alpha,p,\infty} + \|g\|_{\alpha,p,\infty} + \|h\|_{\alpha,p,\infty}\Big)(\|g - h\|^{\gamma}_{\infty} + \|g - h\|_{\alpha,p,\infty}). \end{aligned}$$

Proof. The remainder $R_F(g)$ is given by

$$R_F(g) = F(g) - F(0) - T_{F'(g)}g = \sum_{j\geq -1} F_j$$
$$\text{with} \quad F_j := \Delta_j(F(g) - F(0)) - S_{j-1}(F'(g))\Delta_j g.$$

For $j \leq 0$ Young's inequality and the Lipschitz continuity of F yield

$$\|F_j\|_{L^p} = \|\Delta_j(F(g) - F(0))\|_{L^p} \leq \|\mathcal{F}^{-1}\rho_j\|_{L^1}\|F(g) - F(0)\|_{L^p} \lesssim \|F\|_{C_b^1}\|g\|_{L^p}$$

and we have $\|F_j\|_{L^{p'}} \leq \|F_j\|_{L^p}\|F_j\|_{L^{\alpha p/\beta}} \lesssim \|F_j\|_\infty^{1-\beta/\alpha}\|F_j\|_{L^p}^{1+\beta/\alpha}$.

For $j > 0$ we have $\Delta_j F(0) = 0$ and the Fourier transform of F_j is supported in 2^j times some annulus. Defining the kernel functions $K_j := \mathcal{F}^{-1}\rho_j$ and $K_{<j-1} := \sum_{k<j-1} K_k$ and using that $\int K_j(x)\,\mathrm{d}x = \rho_j(0) = 0$, the blocks F_j can be written as convolution

$$\begin{aligned}
F_j(x) &= \int_{\mathbb{R}^2} K_j(x-y)K_{<j-1}(x-z)\big(F(g(y)) - F'(g(z))g(y)\big)\,\mathrm{d}y\,\mathrm{d}z \\
&= \int_{\mathbb{R}^2} K_j(x-y)K_{<j-1}(x-z)\Big(F(g(y)) - F(g(z)) - F'(g(z))\big(g(y)-g(z)\big)\Big)\,\mathrm{d}y\,\mathrm{d}z \\
&= \int_{\mathbb{R}^2} K_j(x-y)K_{<j-1}(x-z) \\
&\qquad \times \Big(\big(F'(g(z) + \xi_{yz}(g(y)-g(z))) - F'(g(z))\big)\big(g(y)-g(z)\big)\Big)\,\mathrm{d}y\,\mathrm{d}z, \quad (5.21)
\end{aligned}$$

where we used in the in last equality the mean value theorem for intermediate points $\xi_{yz} \in [0,1]$. By the Hölder continuity of F' the above display can be estimated by

$$\begin{aligned}
|F_j(x)| &\leq \|F\|_{C_b^{1+\beta/\alpha}} \int_{\mathbb{R}^2} |K_j(x-y)K_{<j-1}(x-z)|\xi_{yz}^{\beta/\alpha}|g(y)-g(z)|^{\beta/\alpha+1}\,\mathrm{d}y\,\mathrm{d}z \\
&\leq \|F\|_{C_b^{1+\beta/\alpha}} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(z)||g(x-y)-g(x-z)|^{\beta/\alpha+1}\,\mathrm{d}y\,\mathrm{d}z.
\end{aligned}$$

Now we can estimate the $L^{p'}$-norm of the integral by the integral of the $L^{p'}$-norm, which yields

$$\begin{aligned}
\|F_j\|_{L^{p'}} &\leq \|F\|_{C_b^{1+\beta/\alpha}} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(z)|\,\big\||g(\cdot-(y-z)) - g(\cdot)|^{\beta/\alpha+1}\big\|_{L^{p'}}\,\mathrm{d}y\,\mathrm{d}z \\
&\leq \|F\|_{C_b^{1+\beta/\alpha}} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(z)| \sup_{|h|\leq|y-z|} \big\||g(\cdot)-g(\cdot-h)|^{\beta/\alpha+1}\big\|_{L^{p'}}\,\mathrm{d}y\,\mathrm{d}z \\
&= \|F\|_{C_b^{1+\beta/\alpha}} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(z)| \sup_{|h|\leq|y-z|} \|g(\cdot)-g(\cdot-h)\|_{L^p}^{1+\beta/\alpha}\,\mathrm{d}y\,\mathrm{d}z.
\end{aligned}$$

Recalling the modulus of continuity from (5.2) and the corresponding representation of the Besov norm, we obtain with Hölder's inequality for any $q \in [1,\infty]$ with $q^* = q/(q-1)$

$$\begin{aligned}
\|F_j\|_{L^{p'}} &\leq \|F\|_{C_b^{1+\beta/\alpha}} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(y-h)|\omega_p(g,|h|)^{1+\beta/\alpha}\,\mathrm{d}y\,\mathrm{d}h \\
&\lesssim \|F\|_{C_b^{1+\beta/\alpha}}\|g\|_{\alpha,p,(1+\beta/\alpha)q}^{1+\beta/\alpha}\Big(\int \Big(|h|^{\alpha+\beta+d/q}\int |K_j(y)K_{<j-1}(y-h)|\,\mathrm{d}y\Big)^{q^*}\,\mathrm{d}h\Big)^{1/q^*}
\end{aligned} \quad (5.22)$$

(with $d/q := 0$ for $q = \infty$ and the usual modification for $q^* = \infty$). Abbreviating $\delta := \alpha + \beta + d/q$, the last integral can be written as

$$\begin{aligned}
&\big\||h|^\delta(|K_j| * |K_{<j-1}(-\cdot)|)(h)\big\|_{L^{q^*}} \\
&\quad\le \big\|(|h|^\delta|K_j|(h)) * |K_{<j-1}(-\cdot)|\big\|_{L^{q^*}} + \big\||K_j| * (|h|^\delta|K_{<j-1}(-h)|)\big\|_{L^{q^*}} \\
&\quad\le \||h|^\delta|K_j|(h)\|_{L^{q^*}}\|K_{<j-1}\|_{L^1} + \|K_j\|_{L^1}\||h|^\delta|K_{<j-1}(-h)|\|_{L^{q^*}},
\end{aligned}$$

where we apply Young's inequality in the last estimate. Due to $K_j = \mathcal{F}^{-1}\rho_j = (2\pi)^{-d}2^{jd}\mathcal{F}\rho(2^j\cdot)$, we see easily that $\||h|^\delta|K_j|(h)\|_{L^{q^*}} \lesssim 2^{-j(\alpha+\beta)}$ and $\|K_j\|_{L^1} \lesssim 1$. To bound similarly the norms of $K_{<j-1}$ note that $\mathcal{F}K_{<j-1}$ is uniformly bounded and supported on a ball with radius of order 2^j. We conclude

$$\|F_j\|_{L^{p'}} \lesssim 2^{-j(\alpha+\beta)}\|F\|_{C_b^{1+\beta/\alpha}}\|g\|_{\alpha,p,(1+\beta/\alpha)q}^{1+\beta/\alpha}.$$

The claimed bound $\|R_F(g)\|_{\alpha+\beta,p,\infty}$ thus follows from Lemma A.4.1 and choosing $q = \infty$.

To show the Hölder continuity, we will apply similar arguments. For convenience we define $\Delta f(y,z) := f(y) - f(z)$ for any function f. Using the additional regularity of F, we obtain from (5.21) that

$$\begin{aligned}
F_j(x) &= \int_{\mathbb{R}^2} K_j(x-y)K_{<j-1}(x-z)\int_0^1 \Big(F'(g(z) + s\Delta g(y,z)) - F'(g(z))\Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times \Delta g(y,z)\,\mathrm{d}s\,\mathrm{d}y\,\mathrm{d}z \\
&= \int_{\mathbb{R}^2} K_j(x-y)K_{<j-1}(x-z)\int_0^1\int_0^1 sF''(g(z) + rs\Delta g(y,z))\Delta g(y,z)^2\,\mathrm{d}r\,\mathrm{d}s\,\mathrm{d}y\,\mathrm{d}z.
\end{aligned}$$

Hence, we can write

$$R_F(g) - R_F(h) = \sum_{j\ge -1} G_j$$

with

$$\begin{aligned}
G_j(x) &= \int_{\mathbb{R}^2}\int_0^1\int_0^1 K_j(x-y)K_{<j-1}(x-z)s\Big(F''(g(z) + rs\Delta g(y,z))\Delta g(y,z)^2 \\
&\qquad\qquad - F''(h(z) + rs\Delta h(y,z))\Delta h(y,z)^2\Big)\,\mathrm{d}r\,\mathrm{d}s\,\mathrm{d}y\,\mathrm{d}z \\
&= \int_{\mathbb{R}^2}\int_0^1\int_0^1 K_j(x-y)K_{<j-1}(x-z)s \\
&\qquad \times \Big(\Big(F''(g(z) + rs\Delta g(y,z)) - F''(h(z) + rs\Delta h(y,z))\Big) \\
&\qquad \times \Delta g(y,z)^2 + F''(h(z) + rs\Delta h(y,z))(\Delta g(y,z)^2 - \Delta h(y,z)^2)\Big)\,\mathrm{d}r\,\mathrm{d}s\,\mathrm{d}y\,\mathrm{d}z.
\end{aligned}$$

The Hölder continuity of F'' yields

$$\begin{aligned}|G_j(x)| \leq &\|F\|_{C_b^{2+\gamma}} \int_{\mathbb{R}^2} \int_0^1 \int_0^1 |K_j(x-y)K_{<j-1}(x-z)| \\ &\times \Big(\big|(g-h)(z) + rs\Delta(g-h)(y,z)\big|^\gamma \\ &\times |\Delta g(y,z)|^2 + |\Delta(g-h)(y,z)|\big(|\Delta g(y,z)| + |\Delta h(y,z)|\big)\Big) \,\mathrm{d}r \,\mathrm{d}s \,\mathrm{d}y \,\mathrm{d}z \\ \leq &\|F\|_{C_b^{2+\gamma}} \int_{\mathbb{R}^2} |K_j(x-y)K_{<j-1}(x-z)|\Big(\|g-h\|_\infty^\gamma |\Delta g(y,z)|^2 \\ &+ |\Delta(g-h)(y,z)|\big(|\Delta g(y,z)| + |\Delta h(y,z)|\big)\Big) \,\mathrm{d}y \,\mathrm{d}z.\end{aligned}$$

Using the inequalities by Minkowski and Cauchy-Schwarz, we obtain analogously to (5.22)

$$\begin{aligned}\|G_j\|_{L^{p/2}} \leq &\|F\|_{C_b^{2+\gamma}} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(z)|\Big(\|g-h\|_\infty^\gamma \|\Delta g(x-y,x-z)\|_{L^p}^2 \\ &+ \|\Delta(g-h)(x-y,x-z)\|_{L^p} \\ &\times \big(\|\Delta g(x-y,x-z)\|_{L^p} + \|\Delta h(x-y,x-z)\|_{L^p}\big)\Big) \,\mathrm{d}y \,\mathrm{d}z \\ \leq &\|F\|_{C_b^{2+\gamma}} \int_{\mathbb{R}} \big(|K_j| * |K_{<j-1}(-\cdot)|\big)(z) \\ &\times \Big(\|g-h\|_\infty^\gamma \omega_p(g,|z|)^2 + \omega_p(g-h,|z|)\big(\omega_p(g,|z|) + \omega_p(h,|z|)\big)\Big) \,\mathrm{d}z \\ \leq &\|F\|_{C_b^{2+\gamma}} \Big(\|g-h\|_\infty^\gamma \|g\|_{\alpha,p,2q}^2 + \|g-h\|_{\alpha,p,2q}\big(\|g\|_{\alpha,p,2q} + \|h\|_{\alpha,p,2q}\big)\Big) 2^{-j2\alpha}.\end{aligned}$$

The claimed bound follows again from Lemma A.4.1 and the symmetry in g and h. □

In the situation of Proposition 5.3.1 we conclude

$$F(u) = T_{F'(u)}u + R_F(u) \quad \text{with} \quad \|R_F(u)\|_{2\alpha,p/2,\infty} \lesssim \|u\|_{\alpha,p,\infty}^2.$$

Due to this linearization it remains to study $\pi(T_{F'(u)}u,\xi)$. For Hölder continuous functions [GIP12, Lem. 2.4] have shown that the terms $\pi(T_{F'(u)}u,\xi)$ and $F'(u)\pi(u,\xi)$ only differ by a smoother remainder. To find an estimate of the regularity for the commutator

$$\Gamma(f,g,h) := \pi(T_f g, h) - f\pi(g,h) \tag{5.23}$$

in general Besov norms, we first prove the following auxiliary lemma, cf. [BCD11, Lem. 2.97].

Lemma 5.3.3. *Let $p, p_1, p_2 \geq 1$ such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} \leq 1$. Then for $\alpha \in (0,1)$ and for any $f \in B_{p_1,\infty}^\alpha(\mathbb{R}^d)$ and $g \in L^{p_2}(\mathbb{R}^d)$ the operator $[\Delta_j, f]g := \Delta_j(fg) - f\Delta_j g$ satisfies*

$$\|[\Delta_j, f]g\|_{L^p} \lesssim 2^{-j\alpha}\|f\|_{\alpha,p_1,\infty}\|g\|_{L^{p_2}}.$$

Proof. Since $\Delta_j f = (\mathcal{F}^{-1}\rho_j) * f$, we observe

$$[\Delta_j, f]g(x) = \mathcal{F}^{-1}\rho_j * (fg)(x) - f(\mathcal{F}^{-1}\rho_j * g)(x) \\ = \int_{\mathbb{R}} \mathcal{F}^{-1}\rho_j(y)\big(f(x-y) - f(x)\big)g(x-y)\,\mathrm{d}y, \quad x \in \mathbb{R}^d.$$

Minkowski's and Hölder's inequalities yield

$$\|[\Delta_j, f]g\|_{L^p} \leq \int_{\mathbb{R}} \big\|\mathcal{F}^{-1}\rho_j(y)(f(\cdot - y) - f)g(\cdot - y)\big\|_{L^p}\,\mathrm{d}y \\ \leq \|g\|_{L^{p_2}} \int_{\mathbb{R}} |\mathcal{F}^{-1}\rho_j(y)| \|f(\cdot - y) - f\|_{L^{p_1}}\,\mathrm{d}y.$$

With the modulus of continuity (5.2) and the corresponding Besov norm, we obtain

$$\|[\Delta_j, f]g\|_{L^p} \leq \|g\|_{L^{p_2}} \int_{\mathbb{R}} |\mathcal{F}^{-1}\rho_j(y)\omega_{p_1}(f, |y|)|\,\mathrm{d}y \\ \leq \|g\|_{L^{p_2}} \sup_{y \in \mathbb{R}^d} \{|y|^{-\alpha}\omega_{p_1}(f, |y|)\} \int_{\mathbb{R}} |y|^{\alpha}|\mathcal{F}^{-1}\rho_j(y)|\,\mathrm{d}y \\ \sim \|f\|_{\alpha,p_1,\infty}\|g\|_{L^{p_2}} \big\||y|^{\alpha}|\mathcal{F}^{-1}\rho_j(y)|\big\|_{L^1}.$$

For $j = -1$ the previous L^1-norm is finite because χ is smooth and compactly supported. For $j \geq 0$ we additionally note that $\mathcal{F}^{-1}\rho_j = 2^{jd}\mathcal{F}\rho(2^j \cdot)$ implies

$$\big\||y|^{\alpha}|\mathcal{F}^{-1}\rho_j(y)|\big\|_{L^1} = 2^{-j\alpha}\big\||y|^{\alpha}|\mathcal{F}^{-1}\rho(y)|\big\|_{L^1} \lesssim 2^{-j\alpha}.$$ □

Lemma 5.3.4. *Let $\alpha \in (0,1)$, $\beta, \gamma \in \mathbb{R}$ such that $\alpha + \beta + \gamma > 0$ and $\beta + \gamma < 0$. Moreover, let $p_1, p_2, p_3 \geq 1$ satisfy $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = \frac{1}{p} \leq 1$ and let $q \geq 1$. Then for $f, g, h \in \mathcal{S}(\mathbb{R}^d)$ the commutator operator from (5.23) satisfies*

$$\|\Gamma(f, g, h)\|_{\alpha+\beta+\gamma,p,q} \lesssim \|f\|_{\alpha,p_1,q}\|g\|_{\beta,p_2,q}\|h\|_{\gamma,p_3,q}.$$

Therefore, Γ can be uniquely extended to a bounded trilinear operator

$$\Gamma\colon B^{\alpha}_{p_1,q}(\mathbb{R}^d) \times B^{\beta}_{p_2,q}(\mathbb{R}^d) \times B^{\gamma}_{p_3,q}(\mathbb{R}^d) \to B^{\alpha+\beta+\gamma}_{p,q}(\mathbb{R}^d).$$

Proof. Let $f, g, h \in \mathcal{S}(\mathbb{R}^d)$. Using $T_f g = \sum_{k \geq -1}\sum_{l \geq k+2} \Delta_k f \Delta_l g = \sum_{k \geq -1} \Delta_k f (g - S_{k+2}g)$, we decompose

$$\begin{aligned}
\Gamma(f, g, h) &= \pi(T_f g, h) - f\pi(g, h) \\
&= \sum_{j \geq -1} \sum_{i:|i-j| \leq 1} (\Delta_i(T_f g)\Delta_j h - f\Delta_i g \Delta_j h) \\
&= \sum_{j,k \geq -1} \sum_{i:|i-j| \leq 1} \Big(\Delta_i((\Delta_k f)(g - S_{k+2}g)) - \Delta_k f \Delta_i g\Big)\Delta_j h \\
&= -\sum_{k \geq -1} \underbrace{\sum_{j \geq -1} \sum_{i:|i-j| \leq 1} \Delta_k f \Delta_i(S_{k+2}g)\Delta_j h}_{=:a_k} \\
&\quad + \sum_{j \geq -1} \underbrace{\sum_{k \geq -1} \sum_{i:|i-j| \leq 1} ([\Delta_i, \Delta_k f](g - S_{k+2}g))\Delta_j h}_{=:b_j}.
\end{aligned} \tag{5.24}$$

We will separately estimate both sums in the following.

For $k \geq -1$ we have $\Delta_i(S_{k+2}g) = 0$ for $i > k+2$ due to property (iii) of the dyadic partition of unity. Consequently,

$$a_k = \sum_{i=-1}^{k+2} \sum_{j:|i-j|\leq 1} \Delta_k f \Delta_i(S_{k+2}g)\Delta_j h$$

and its Fourier transform satisfies $\operatorname{supp}\mathcal{F}a_k \subset 2^k\mathcal{B}$ for some ball $\mathcal{B}$. Hölder's inequality yields

$$\|a_k\|_{L^p} \leq \|\Delta_k f\|_{L^{p_1}} \sum_{i=-1}^{k+2} \sum_{j:|i-j|\leq 1} \|\Delta_i(S_{k+2}g)\|_{L^{p_2}} \|\Delta_j h\|_{L^{p_3}}.$$

Owing to $\Delta_i(S_{k+2}g) = \Delta_i g$ for $i \leq k$ and $\|\Delta_i\Delta_k g\|_{L^{p_2}} \leq \|\mathcal{F}^{-1}\rho_i\|_{L^1}\|\Delta_k g\|_{L^{p_2}} \lesssim \|\Delta_k g\|_{L^{p_2}}$ by Young's inequality, we have

$$\begin{aligned}\|a_k\|_{L^p} &\lesssim \|\Delta_k f\|_{L^{p_1}} \sum_{i=-1}^{k+2} \sum_{j:|i-j|\leq 1} \|\Delta_i g\|_{L^{p_2}} \|\Delta_j h\|_{L^{p_3}} \\ &\lesssim \|\Delta_k f\|_{L^{p_1}} \|g\|_{\beta,p_2,\infty}\|h\|_{\gamma,p_3,\infty} \sum_{i=-1}^{k+2} 2^{-i(\beta+\gamma)} \\ &\lesssim 2^{-k(\beta+\gamma)}\|\Delta_k f\|_{L^{p_1}} \|g\|_{\beta,p_2,\infty}\|h\|_{\gamma,p_3,\infty},\end{aligned}$$

using $\beta+\gamma<0$ in the last estimate. Since $2^{k\alpha}\|\Delta_k f\|_{L^{p_1}} \in \ell^q$, Lemma A.4.2 yields

$$\Big\|\sum_{k\geq -1} a_k\Big\|_{\alpha+\beta+\gamma,p,q} \lesssim \|f\|_{\alpha,p_1,q}\|g\|_{\beta,p_2,\infty}\|h\|_{\gamma,p_3,\infty}.$$

Now, let us consider the second sum in (5.24). Note that

$$b_j = \sum_{i:|i-j|\leq 1} \sum_{k\geq -1} \sum_{l\geq k+2} ([\Delta_i, \Delta_k f]\Delta_l g)\Delta_j h = \sum_{i:|i-j|\leq 1} \sum_{l\geq -1} ([\Delta_i, S_{l-1}f]\Delta_l g)\Delta_j h.$$

Since the support of the Fourier transform of $S_{l-1}f\Delta_l g$ is of the form $2^l\mathcal{A}$ for some annulus $\mathcal{A}$, we have that

$$[\Delta_i, S_{l-1}f]\Delta_l g = \Delta_i(S_{l-1}f\Delta_l g) - (S_{l-1}f)(\Delta_i\Delta_l g)$$

vanishes if $|i-l| \geq N$ for some $N \in \mathbb{N}$. Therefore,

$$b_j = \sum_{i:|i-j|\leq 1} \sum_{l\sim i} ([\Delta_i, S_{l-1}f]\Delta_l g)\Delta_j h$$

has a Fourier transform supported on 2^j times some annulus. Using Hölder's inequality and Lemma 5.3.3, we estimate

$$\begin{aligned}\|b_j\|_{L^p} &\lesssim \sum_{i:|i-j|\leq 1} \sum_{l\sim i} 2^{-i\alpha}\|S_{k-1}f\|_{\alpha,p_1,\infty}\|\Delta_l g\|_{L^{p_2}}\|\Delta_j h\|_{L^{p_3}} \\ &\lesssim 2^{-j(\alpha+\beta+\gamma)}\|f\|_{\alpha,p_1,\infty}(2^{j\beta}\sum_{l\sim j}\|\Delta_l g\|_{L^{p_2}})2^{j\gamma}\|\Delta_j h\|_{L^{p_3}}.\end{aligned}$$

For any $q_2, q_3 \geq q$ satisfying $\frac{1}{q} = \frac{1}{q_2} + \frac{1}{q_3}$ Hölder's inequality and Lemma A.4.2 yield then

$$\Big\| \sum_{j\geq -1} b_j \Big\|_{\alpha+\beta+\gamma,p,q} \lesssim \|f\|_{\alpha,p_1,\infty} \|g\|_{\beta,p_2,q_2} \|h\|_{\gamma,p_3,q_3}.$$

To obtain the claimed norm bound, recall that $B^\alpha_{p,q}(\mathbb{R}^d)$ continuously embeds into $B^\alpha_{p,q'}(\mathbb{R}^d)$ for any $q \leq q'$.

For $p, q < \infty$ the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is dense $B^\alpha_{p,q}(\mathbb{R}^d)$ for any $\alpha \in \mathbb{R}$ such that there is a unique extension of C on $B^\alpha_{p_1,q}(\mathbb{R}^d) \times B^\beta_{p_2,q}(\mathbb{R}^d) \times B^\gamma_{p_3,q}(\mathbb{R}^d)$. For $p = \infty$ or $q = \infty$ a similar argument as in [GIP12, Lem. 2.4] applies. $\square$

Combining the previous results, we obtain the following corollary, cf. [GIP12, Lem. 2.7], which immediately implies Proposition 5.3.1 due to the embedding $B^\alpha_{p,q} \subset L^\infty$ for $\alpha > 1/p$ and $d = 1$.

Corollary 5.3.5. *Let $p_1, p_2 \in [1,\infty]$ satisfy $\frac{2}{p_1} + \frac{1}{p_2} =: \frac{1}{p} \leq 1$. Let $\alpha \in (0,1)$ and $\beta < 0$ such that $2\alpha + \beta > 0$ and $\alpha + \beta < 0$. Further, suppose $F \in C^{2+\gamma}_b(\mathbb{R}^m)$ for some $\gamma \in (0,1]$ satisfying $F(0) = 0$. Then there exists a map $\Pi_F \colon B^\alpha_{p_1,\infty}(\mathbb{R}^d) \times B^\beta_{p_2,\infty}(\mathbb{R}^d) \to B^{2\alpha+\beta}_{p,\infty}(\mathbb{R}^d)$ such that*

$$\pi(F(f), g) = F'(f)\pi(f,g) + \Pi_F(f,g)$$

and

$$\|\Pi_F(f,g)\|_{2\alpha+\beta,p,\infty} \lesssim \|F\|_{C^2_b} \|f\|^2_{\alpha,p_1,\infty} \|g\|_{\beta,p_2,\infty}.$$

For $f_1, f_2 \in B^\alpha_{p_1,\infty}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ and $g_1, g_2 \in B^\beta_{p_2,\infty}(\mathbb{R}^d)$ we have furthermore

$$\begin{aligned}
&\|\Pi_F(f_1,g_1) - \Pi_F(f_2,g_2)\|_{2\alpha+\beta,p,\infty} \\
&\lesssim \|F\|_{C^{2+\gamma}_b} \Big(\|f_1\|^2_{\alpha,p_1,q} \wedge \|f_2\|^2_{\alpha,p_1,\infty} + (\|f_1\|_{\alpha,p_1,\infty} + \|f_2\|_{\alpha,p_1,\infty}) \\
&\times (1 + \|g_1\|_{\beta,p_2,\infty} \wedge \|g_1\|_{\beta,p_2,\infty}) \Big) \Big(\|f_1 - f_2\|^\gamma_\infty + \|f_1 - f_2\|_{\alpha,p_1,\infty} + \|g_1 - g_2\|_{\beta,p_2,\infty} \Big).
\end{aligned}$$

Proof. Setting $\Pi_F(f,g) := \Gamma(F'(f), f, g) + \pi(R_F(f), g)$, we can write

$$\pi(F(f),g) = F'(f)\pi(f,g) + \Gamma(F'(f),f,g) + \pi(R_F(f),g) = F'(f)\pi(f,g) + \Pi_F(f,g).$$

Lemmas 5.1.1, 5.3.2 and 5.3.4 yield

$$\begin{aligned}
\|\Pi_F(f,g)\|_{2\alpha+\beta,p,\infty} \leq& \|\Gamma(F'(f),f,g)\|_{2\alpha+\beta,p,\infty} + \|\pi(R_F(f),g)\|_{2\alpha+\beta,p,\infty} \\
\lesssim& \|F'(f)\|_{\alpha,p_1,\infty} \|f\|_{\alpha,p_1,\infty} \|g\|_{\beta,p_2,\infty} + \|R_F(f)\|_{2\alpha,p_1/2,\infty} \|g\|_{\beta,p_2,\infty} \\
\lesssim& (\|F'(f)\|_{\alpha,p_1,\infty} + \|F\|_{C^2_b} \|f\|_{\alpha,p_1,\infty}) \|f\|_{\alpha,p_1,\infty} \|g\|_{\beta,p_2,\infty},
\end{aligned}$$

where we again used Besov embeddings. Finally, we apply (5.10).

The bound of $\|\Pi_F(f_1,g_1) - \Pi_F(f_2,g_2)\|_{2\alpha+\beta,p,\infty}$ follows from analogous estimates, using the argument-wise linearity of Γ and π, the Hölder continuity of R_F from

Lemma 5.3.2 and

$$\begin{aligned}\|F'(f_1)-F'(f_2)\|_{\alpha,p_1,q} =& \Big\| \int_0^1 F''(f_1+s(f_2-f_1))(f_1-f_2)\,\mathrm{d}s \Big\|_{\alpha,p_1,q}\\ \le& \int_0^1 \|F''(f_1+s(f_2-f_1))(f_1-f_2)\|_{\alpha,p_1,q}\,\mathrm{d}s\\ \le& \|F''\|_\infty \|f_1-f_2\|_{\alpha,p_1,q} \qquad (5.25)\end{aligned}$$

for any $q\in[1,\infty]$. □

5.4. The paracontrolled ansatz

Assuming that the driving signal ξ satisfies $\xi\in B^{\alpha}_{p,q}$ for $\alpha>1/3$, we come back to the RDE (5.1). Recall that it was given by

$$\mathrm{d}u(t)=F(u(t))\xi(t),\quad u(0)=u_0,\quad t\in\mathbb{R},$$

where $u_0\in\mathbb{R}^m$, $u\colon\mathbb{R}\to\mathbb{R}^m$ is a continuous function and $F\colon\mathbb{R}^m\to\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ is a family of vector fields on $\mathbb{R}^m$. In Section 5.2 we have already considered the case $\alpha>1/2$. The classical way to continuously extend Young's approach to more irregular driving signals is Lyons' rough path theory, which additionally to the signal ξ needs to handle the corresponding "iterated integral".

As an alternative, we use in the present section a new paracontrolled ansatz similar to Gubinelli et al. [GIP12]. We postulate that the solution u of the RDE (5.1) is of the form

$$u=T_{u^\vartheta}\vartheta+u^{\#}$$

with $\vartheta,u^\vartheta\in B^{\alpha}_{p,q}$ and a remainder $u^{\#}\in B^{2\alpha}_{p/2,q}$. Decomposing $F(u)\xi$ in terms of Littlewood-Paley blocks and linearizing F by Proposition 5.3.1, we have

$$F(u)\xi=T_{F(u)}\xi+\pi(F(u),\xi)+T_\xi(F(u))=T_{F(u)}\xi+F'(u)\pi(u,\xi)+\Pi_F(u,\xi)+T_\xi(F(u)).$$

The presumed controlled structure yields that understanding the (problematic) term $\pi(u,\xi)$ reduces further to the analysis of $\pi(\vartheta,\xi)$ owing to the commutator from (5.23):

$$\pi(u,\xi)=\pi(T_{u^\vartheta}\vartheta,\xi)+\pi(u^{\#},\xi)=u^\vartheta\pi(\vartheta,\xi)+\underbrace{\Gamma(u^\vartheta,\vartheta,\xi)}_{\in B^{3\alpha-1}_{p/3,q}}+\underbrace{\pi(u^{\#},\xi)}_{\in B^{3\alpha-1}_{p/3,q}}.$$

Plugging the paracontrolled ansatz into the RDE (5.1), the Leibniz rule and the above observation yield

$$T_{u^\vartheta}\,\mathrm{d}\vartheta+T_{\mathrm{d}u^\vartheta}\vartheta+\,\mathrm{d}u^{\#}=\,\mathrm{d}u=T_{F(u)}\xi+F'(u)\pi(u,\xi)+\Pi_F(u,\xi)+T_\xi(F(u)).$$

Comparing the least regular terms on the left-hand and on the right-hand side, we choose ϑ as the solution to $\mathrm{d}\vartheta=\xi$ with $\vartheta(0)=0$ and $u^\vartheta=F(u)$.

As already noted in Section 5.2, we cannot expect ϑ to be contained in any Besov space (cf. Lemma 5.1.2). This requirement would especially be violated in most interesting examples from probability theory, for instance, ϑ being Brownian motion or a martingale. In order to circumvent this issue, we use again the localizing function φ from Assumption 1. Still relying on $\mathrm{d}\vartheta = \xi$ and $\vartheta(0) = 0$, we introduce the local version of the signal

$$\vartheta_{\mathcal{T}} := \varphi_{\mathcal{T}}\vartheta \quad \text{and} \quad \xi_{\mathcal{T}} := \mathrm{d}\vartheta_{\mathcal{T}} = \varphi_{\mathcal{T}}\xi + \varphi_{\mathcal{T}}'\vartheta.$$

The corresponding localized RDE is then given by

$$\mathrm{d}u = F(u)\xi_{\mathcal{T}}, \quad u(0) = u_0. \tag{5.26}$$

This differential equation coincides with the original one on the interval $[-\mathcal{T}, \mathcal{T}]$ due to $\varphi(t) = 1$ and $\varphi'(t) = 0$ for $|t| \leq \mathcal{T}$.

Summarizing briefly the above discussion, we need two additional pieces of information about very irregular signals. Namely, $\xi_{\mathcal{T}}$ has to be the derivative of a path $\vartheta_{\mathcal{T}}$ with compact support and the resonant term $\pi(\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}})$ has to be well-defined. This precisely corresponds to the classical rough path theory, where a path ϑ defined on some compact interval is enhanced with the information of the iterated integral $\int \vartheta_s \,\mathrm{d}\vartheta_s$.

Analogously to the notion of geometric rough path (cf. for example Section 2.2. in [FH14]), we introduce now the notion of geometric Besov rough path:

Definition 5.4.1. Let $\mathcal{T} > 0$ and let $C_{\mathcal{T}}^{\infty}$ be the space of smooth functions $\vartheta_{\mathcal{T}} \colon \mathbb{R} \to \mathbb{R}^n$ with support $\operatorname{supp} \vartheta_{\mathcal{T}} \subset [-2\mathcal{T}, 2\mathcal{T}]$ and $\vartheta_{\mathcal{T}}(0) = 0$. The closure of the set $\{(\vartheta_{\mathcal{T}}, \pi(\vartheta_{\mathcal{T}}, \mathrm{d}\vartheta_{\mathcal{T}})) : \vartheta_{\mathcal{T}} \in C_{\mathcal{T}}^{\infty}\} \subset B_{p,q}^{\alpha} \times B_{p/2,q}^{2\alpha-1}$ with respect to the norm $\|\cdot\|_{\alpha,p,q} + \|\cdot\|_{2\alpha-1,p/2,q}$ is denoted by $\mathcal{B}_{p,q}^{0,\alpha}$ and $(\vartheta_{\mathcal{T}}, \eta_{\mathcal{T}}) \in \mathcal{B}_{p,q}^{0,\alpha}$ is called *geometric Besov rough path*.

Even with the driving signal $(\vartheta, \eta) \in \mathcal{B}_{p,q}^{0,\alpha}$ we unfortunately cannot expect in general that the solution u to the Cauchy problem (5.26) with $\xi_{\mathcal{T}} = \mathrm{d}\vartheta_{\mathcal{T}}$ lies in any Besov spaces $B_{p,q}^{\alpha}$ for finite p and q. On the other hand, Besov spaces on the compact domain $[-\mathcal{T}, \mathcal{T}]$ seem not be convenient for the paraproduct approach since Littlewood-Paley theory and Bony's paraproduct are from their very nature constructed on the whole real line. It appears to be natural to instead consider a weighted version of the Itô-Lyons $\hat{S}$ map given by

$$\hat{S} \colon \mathbb{R}^d \times \mathcal{B}_{p,q}^{0,\alpha} \to B_{p,q}^{\alpha} \quad \text{via} \quad (u_0, \vartheta_{\mathcal{T}}, \pi(\vartheta_{\mathcal{T}}, \mathrm{d}\vartheta_{\mathcal{T}})) \mapsto \psi u, \tag{5.27}$$

where u solves (5.26) with $\xi_{\mathcal{T}} = \mathrm{d}\vartheta_{\mathcal{T}}$ and $\psi \colon \mathbb{R} \to (0, \infty)$ is a regular weight function being constant one on $[-2\mathcal{T}, 2\mathcal{T}]$. Consequently, provided $\vartheta_{\mathcal{T}} \in C_{\mathcal{T}}^{\infty}$ with $\xi_{\mathcal{T}} = \mathrm{d}\vartheta_{\mathcal{T}}$ the weighted solution $\tilde{u} := \psi u$ possesses the dynamic

$$\mathrm{d}\tilde{u} = \psi \,\mathrm{d}u + \psi' u = F(\tilde{u})\xi_{\mathcal{T}} + \frac{\psi'}{\psi}\tilde{u}, \quad \tilde{u}(0) = u_0. \tag{5.28}$$

Let us emphasize that also this weighted differential equation still coincides with the original RDE (5.1) restricted to the interval $[-\mathcal{T}, \mathcal{T}]$.

The aim is now to continuously extend the weighted Itô-Lyons map $\hat{S}$ from smooth functions with support in $[-2\mathcal{T}, 2\mathcal{T}]$ to the geometric Besov rough paths or more precisely from the domain $\mathbb{R}^d \times \{(\vartheta_{\mathcal{T}}, \pi(\vartheta_{\mathcal{T}}, \mathrm{d}\vartheta_{\mathcal{T}})) : \vartheta_{\mathcal{T}} \in C_{\mathcal{T}}^{\infty}\}$ to $\mathbb{R}^d \times \mathcal{B}_{p,q}^{0,\alpha}$. For this purpose we specify our assumptions on the weight function ψ as follows:

Assumption 2. For any $\mathcal{T} > 0$ let $\psi = \psi_{\mathcal{T}} \in B^{\alpha}_{p,q} \cap C^1_b$ be a strictly positive function which is equal to one on $[-2\mathcal{T}, 2\mathcal{T}]$ and suppose that there exist two constants $C_\psi, c_\psi > 0$ such that $\|\psi'/\psi\|_\infty \lesssim C_\psi$ and $\max\{\psi(2\mathcal{T}+1), \psi(-2\mathcal{T}-1)\} > c_\psi$.

The conditions on ψ are quite weak and allow for a large variety of weight functions as illustrated by the following examples.

Example 5.4.2. Let $\alpha \in (0,1)$, $\mathcal{T} > 0$ and $\kappa \in (0,1)$.

(i) The function

$$\psi_{\mathcal{T}}(t) := \begin{cases} 1, & |t| \le 2\mathcal{T}, \\ \exp\Big(-\frac{\kappa(|t|-2\mathcal{T})^2}{1+|t|-2\mathcal{T}}\Big), & |t| > 2\mathcal{T}, \end{cases}$$

satisfies Assumption 2 for $C_\psi = \kappa$ and $c_\psi = e^{-1/2}$.

(ii) The function

$$\psi_{\mathcal{T}}(t) := \begin{cases} 1, & |t| \le 2\mathcal{T}, \\ \big(1+\kappa(|t|-2\mathcal{T})^2\big)^{-2}, & |t| > 2\mathcal{T}, \end{cases}$$

satisfies Assumption 2 for $C_\psi = \sqrt{\kappa}$ and $c_\psi = 1/4$.

For later reference let us remark a property which makes weight functions fulfilling Assumption 2 so suitable in our context.

Remark 5.4.3. *For any two weight functions ψ and $\tilde\psi$ satisfying Assumption 2, the resulting weighted Besov norms of the solution u are equivalent. More precisely, it is elementary to show*

$$\|\psi u\|_{\alpha,p,q} \lesssim (1 + c_{\tilde\psi}^{-1}\|\tilde\psi - \psi\|_{\alpha,p,q})\|\tilde\psi u\|_{\alpha,p,q}$$

for any $u \in B^{\alpha}_{p,q}$ which is constant on $(-\infty, -2\mathcal{T}]$ and on $[2\mathcal{T}, \infty)$.

In order to analyze the weighted RDE (5.28), we modify our ansatz to

$$\tilde u = T_{F(\tilde u)}\vartheta_{\mathcal{T}} + u^{\#}, \quad \text{where} \quad u^{\#} \in B^{2\alpha}_{p/2,q},\ \vartheta_{\mathcal{T}} \in C^\infty_{\mathcal{T}}.$$

Roughly speaking, in the terminology of [GIP12] the pair $(\tilde u, F(\tilde u)) \in (B^{\alpha}_{p,q})^2$ is said to be *paracontrolled* by $\vartheta_{\mathcal{T}} \in B^{\alpha}_{p,q}$. The dynamic of $u^{\#}$ is characterized in the next lemma.

Lemma 5.4.4. *Let $u_0 \in \mathbb{R}^m$, let $\vartheta_{\mathcal{T}} \in C^\infty_{\mathcal{T}}$ with derivative $\xi_{\mathcal{T}} = \mathrm{d}\vartheta_{\mathcal{T}}$ and suppose that ψ satisfies Assumption 2. Then the following conditions are equivalent:*

(i) u is the solution to the ODE (5.26),

(ii) u can be written as $u = \psi^{-1}\tilde u$ where $\tilde u$ solves the ODE (5.28),

(iii) $\tilde u$ can be written as $\tilde u = T_{F(\tilde u)}\vartheta_{\mathcal{T}} + u^{\#}$ where $u^{\#}$ solves

$$\mathrm{d}u^{\#} = F(\tilde u)\xi_{\mathcal{T}} - \mathrm{d}(T_{F(\tilde u)}\vartheta_{\mathcal{T}}) + \frac{\psi'}{\psi}\tilde u, \qquad u^{\#}(0) = u_0 - T_{F(\tilde u)}\vartheta_{\mathcal{T}}(0). \tag{5.29}$$

Proof. For the equivalence between (i) and (ii) note that $u = \psi^{-1}\tilde{u}$ is well-defined by Assumption 2 and that we have by the Leibniz rule

$$\mathrm{d}u = \mathrm{d}(\psi^{-1}\tilde{u}) = \psi^{-1}\,\mathrm{d}\tilde{u} - \frac{\psi'}{\psi^2}\tilde{u} = F(u)\xi_{\mathcal{T}}, \qquad u(0) = \psi^{-1}(0)\tilde{u}(0) = u_0.$$

The equivalence between (ii) and (iii) follows by combining $\tilde{u} = T_{F(\tilde{u})}\vartheta_{\mathcal{T}} + u^{\#}$ and (5.28), which yields

$$\mathrm{d}u^{\#} = \mathrm{d}\tilde{u} - \mathrm{d}(T_{F(\tilde{u})}\vartheta_{\mathcal{T}}) = F(\tilde{u})\xi_{\mathcal{T}} - \mathrm{d}(T_{F(\tilde{u})}\vartheta_{\mathcal{T}}) + \frac{\psi'}{\psi}\tilde{u}$$

and due to $\tilde{u}(0) = u(0) = u_0$ the initial condition satisfies $u^{\#}(0) = u_0 - T_{F(\tilde{u})}\vartheta_{\mathcal{T}}(0)$. □

As we have seen in the discussion at the beginning of the present section, we want to reduce the resonant term $\pi(F(\tilde{u}), \xi_{\mathcal{T}})$ to the resonant term $\pi(\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}})$. Indeed, this is possible as proven in the following proposition. The specific form of u allows to improve the quadratic estimate (5.20) in Proposition 5.3.1 to a linear one. Its proof is inspired by Lemma 5.2 by Gubinelli et al. [GIP12].

Proposition 5.4.5. *Let $\alpha \in (\frac{1}{3}, \frac{1}{2})$, $p \geq 3$, $q \geq 1$, and $F \in C_b^2$ with $F(0) = 0$. If $\vartheta_{\mathcal{T}} \in C_{\mathcal{T}}^{\infty}$ with derivative $\xi_{\mathcal{T}} = \mathrm{d}\vartheta_{\mathcal{T}}$, then for $\tilde{u} = T_{F(\tilde{u})}\vartheta_{\mathcal{T}} + u^{\#}$ with $\tilde{u} \in B_{p,q}^{\alpha}$ and $u^{\#} \in B_{p/2,q}^{2\alpha}$ one has*

$$\begin{aligned}\|\pi(F(\tilde{u}), \xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q} \lesssim{}& \big(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2\big)\big(\|\tilde{u}\|_{\alpha,p,q} + \|u^{\#}\|_{2\alpha,p/2,q}\big) \\ &\times \big(\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q} + \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}^2 + \|\pi(\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\big).\end{aligned}$$

Proof. Step 1: To avoid the quadratic estimate, we first need a modified version of Lemma 5.3.2. We will borrow some notation from the proof of this former lemma. For brevity we define $v_u := T_{F(\tilde{u})}\vartheta_{\mathcal{T}}$ and recall $\tilde{u} := \psi u$ such that $\tilde{u} = v_u + u^{\#}$. We write

$$F(\tilde{u}) - F(0) = T_{F'(\tilde{u})}\tilde{u} + R_F(\tilde{u})$$

with

$$R_F(\tilde{u}) = \sum_{j\geq -1} F_j \quad \text{with} \quad F_j := \Delta_j(F(\tilde{u}) - F(0)) - S_{j-1}(F'(\tilde{u}))\Delta_j\tilde{u}.$$

For $j \leq 0$, we saw in Lemma 5.3.2 that $\|F_j\|_{L^{p/2}} \lesssim \|F\|_{C_b^1}\|\tilde{u}\|_{L^{p/2}}$ which yields

$$\begin{aligned}\|F_j\|_{L^{p/2}} \lesssim{}& \|F\|_{C_b^1}(\|v_u\|_{L^{p/2}} + \|u^{\#}\|_{L^{p/2}}) \\ \leq{}& \|F\|_{C_b^1}(\|T_{F(\tilde{u})}\vartheta_{\mathcal{T}}\|_{L^{p/2}} + \|u^{\#}\|_{L^{p/2}}).\end{aligned}$$

For $j > 0$, we deduce from (5.21) and our ansatz that

$$\begin{aligned}|F_j| ={}& \Big|\int_{\mathbb{R}^2} K_j(x-y)K_{<j-1}(x-z)\Big(\big(F'(\tilde{u}(z) + \xi_{yz}(\tilde{u}(y) - \tilde{u}(z))) - F'(\tilde{u}(z))\big) \\ &\times \big(v_u(y) - v_u(z) + u^{\#}(y) - u^{\#}(z)\big)\Big)\,\mathrm{d}y\,\mathrm{d}z\Big|, \\ \leq{}& \|F\|_{C_b^2}\int_{\mathbb{R}^2} |K_j(x-y)K_{<j-1}(x-z)||\tilde{u}(y) - \tilde{u}(z)||v_u(y) - v_u(z)|\,\mathrm{d}y\,\mathrm{d}z \\ &+ 2\|F\|_{C_b^1}\int |K_j(x-y)K_{<j-1}(x-z)||u^{\#}(y) - u^{\#}(z)|\,\mathrm{d}y\,\mathrm{d}z.\end{aligned}$$

Proceeding as in proof of Lemma 5.3.2 and applying Hölder's inequality, we obtain for $q^* = q/(q-1)$

$$\begin{aligned}
&\|F_j\|_{L^{p/2}} \\
&\leq \|F\|_{C_b^2} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(z)| \big\|\tilde{u}(x-(y-z)) - \tilde{u}(x)\big\|_{L^p} \\
&\qquad\qquad \times \big\|(v_u(x-(y-z)) - v_u(x)\big\|_{L^p} \,\mathrm{d}y\,\mathrm{d}z \\
&\quad + 2\|F\|_{C_b^1} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(z)| \big\|u^{\#}(x-(y-z)) - u^{\#}(x)\big\|_{L^{p/2}} \,\mathrm{d}y\,\mathrm{d}z \\
&\leq \|F\|_{C_b^2} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(y-h)| \omega_p(\tilde{u}, |h|)\omega_p(v_u, |h|) \,\mathrm{d}y\,\mathrm{d}h \\
&\quad + 2\|F\|_{C_b^1} \int_{\mathbb{R}^2} |K_j(y)K_{<j-1}(y-h)| \omega_{p/2}(u^{\#}, |h|) \,\mathrm{d}y\,\mathrm{d}h \\
&\leq \|F\|_{C_b^2} \Big\| |h|^{2\alpha+1/q} (|K_j| * |K_{<j-1}(-\cdot)|)(h) \Big\|_{L^{q^*}} \\
&\quad \times \Big(\Big\| |h|^{-\alpha} \omega_p(v_u, |h|) \Big\|_{\infty} \Big\| (|h|^{-\alpha-1/q} \omega_p(\tilde{u}, |h|) \Big\|_{L^q} + 2 \Big\| |h|^{-2\alpha-1/q} \omega_{p/2}(u^{\#}, |h|) \Big\|_{L^q} \Big) \\
&\lesssim 2^{-j2\alpha} \|F\|_{C_b^2} (\|v_u\|_{\alpha,p,\infty} \|\tilde{u}\|_{\alpha,p,q} + \|u^{\#}\|_{2\alpha,p/2,q}).
\end{aligned}$$

Due to Lemma 5.1.1 one further has

$$\|v_u\|_{\alpha,p,\infty} = \|T_{F(\tilde{u})}\vartheta_{\mathcal{T}}\|_{\alpha,p,\infty} \lesssim \|T_{F(\tilde{u})}\vartheta_{\mathcal{T}}\|_{\alpha,p,q} \lesssim \|F\|_{\infty} \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}$$

and thus Lemma A.4.1 gives

$$\|R_F(\tilde{u})\|_{2\alpha,p/2,\infty} \lesssim \|F\|_{C_b^2} (1 + \|F\|_{\infty} \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}) (\|\tilde{u}\|_{\alpha,p,q} + \|u^{\#}\|_{2\alpha,p/2,q}). \tag{5.30}$$

Step 2: Plugging in the ansatz once again and keeping the definition of our commutator (5.23) in mind, we decompose

$$\begin{aligned}
\pi(F(\tilde{u}), \xi_{\mathcal{T}}) &= \pi(T_{F'(\tilde{u})}\tilde{u}, \xi_{\mathcal{T}}) + \pi(R_F(\tilde{u}), \xi_{\mathcal{T}}) \\
&= \pi(T_{F'(\tilde{u})} T_{F(\tilde{u})} \vartheta_{\mathcal{T}}, \xi_{\mathcal{T}}) + \pi(T_{F'(\tilde{u})} u^{\#}, \xi_{\mathcal{T}}) + \pi(R_F(\tilde{u}), \xi_{\mathcal{T}}) \\
&= F'(\tilde{u}) \pi(T_{F(\tilde{u})}\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}}) + \Gamma(F'(\tilde{u}), T_{F(\tilde{u})}\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}}) + \pi(T_{F'(\tilde{u})} u^{\#}, \xi_{\mathcal{T}}) + \pi(R_F(\tilde{u}), \xi_{\mathcal{T}}) \\
&= F'(\tilde{u}) F(\tilde{u}) \pi(\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}}) + F'(\tilde{u}) \Gamma(F(\tilde{u}), \vartheta_{\mathcal{T}}, \xi_{\mathcal{T}}) + \Gamma(F'(\tilde{u}), T_{F(\tilde{u})}\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}}) \\
&\qquad + \pi(T_{F'(\tilde{u})} u^{\#}, \xi_{\mathcal{T}}) + \pi(R_F(\tilde{u}), \xi_{\mathcal{T}}).
\end{aligned} \tag{5.31}$$

Therefore, we can bound $\|\pi(F(\tilde{u}), \xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q}$ by estimating these five terms separately. We will apply the following bound which holds owing to the Besov embedding $B^{3\alpha-1}_{p/3,q/2} \subset B^{2\alpha-1}_{p/2,q/2}$ due to $\alpha > 1/p$ and which uses Bony's estimates and $2\alpha - 1 < 0$: for $f \in L^{\infty} \cup B^{\alpha}_{p,\infty}$ and $g \in B^{2\alpha-1}_{p/2,q/2}$ it holds

$$\begin{aligned}
\|fg\|_{2\alpha-1,p/2,q/2} &\lesssim \|T_f g\|_{2\alpha-1,p/2,q/2} + \|\pi(f,g)\|_{3\alpha-1,p/3,q/2} + \|T_g f\|_{2\alpha-1,p/2,q/2} \\
&\lesssim \|f\|_{\infty} \|g\|_{2\alpha-1,p/2,q/2} + \big(\|f\|_{0,\infty,\infty} \|g\|_{3\alpha-1,p/3,q/2} \wedge \|f\|_{\alpha,p,\infty} \|g\|_{2\alpha-1,p/2,q/2} \big) \\
&\qquad + \|g\|_{2\alpha-1,p/2,q/2} \|f\|_{0,\infty,\infty} \\
&\lesssim (\|f\|_{\infty} \|g\|_{3\alpha-1,p/3,q/2}) \wedge (\|f\|_{\alpha,p,\infty} \|g\|_{2\alpha-1,p/2,q/2}).
\end{aligned} \tag{5.32}$$

Furthermore, note for the following estimates that $\|\xi_{\mathcal{T}}\|_{\alpha-1,p,q} \lesssim \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}$ thanks to the lifting property of Besov spaces, cf. [Tri10, Thm. 2.3.8].

Applying (5.32) and (5.10) to $\tilde{F} := F'F$, we obtain for the first summand

$$\begin{aligned}\|F'(\tilde{u})F(\tilde{u})\pi(\vartheta_{\mathcal{T}},\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q} \lesssim& \|\tilde{F}(\tilde{u})\|_{\alpha,p,\infty}\|\pi(\vartheta_{\mathcal{T}},\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ \lesssim& \|F\|_{C_b^1}\|F\|_{C_b^2}\|\tilde{u}\|_{\alpha,p,q}\|\pi(\vartheta_{\mathcal{T}},\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q}.\end{aligned}$$

For the second term the above estimate (5.32) and Lemma 5.3.4 yield

$$\begin{aligned}\|F'(\tilde{u})\Gamma(F(\tilde{u}),\vartheta_{\mathcal{T}},\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q} &\lesssim \|F'\|_\infty\|\Gamma(F(\tilde{u}),\vartheta_{\mathcal{T}},\xi_{\mathcal{T}})\|_{3\alpha-1,p/3,q}\\ &\lesssim \|F'\|_\infty\|F(\tilde{u})\|_{\alpha,p,q}\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}\|\xi_{\mathcal{T}}\|_{\alpha-1,p,q}\\ &\lesssim \|F\|_{C_b^1}^2\|\tilde{u}\|_{\alpha,p,q}\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}^2,\end{aligned}$$

where (5.10) is used in the last line. Lemmas 5.1.1 and 5.3.4 again together with (5.10) gives for the third term

$$\begin{aligned}\|\Gamma(F'(\tilde{u}),T_{F(\tilde{u})}\vartheta_{\mathcal{T}},\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q} &\lesssim \|F'(\tilde{u})\|_{\alpha,p,q}\|T_{F(\tilde{u})}\vartheta_{\mathcal{T}}\|_{\alpha,p,q}\|\xi_{\mathcal{T}}\|_{\alpha-1,p,q}\\ &\lesssim \|F\|_{C_b^1}^2\|\tilde{u}\|_{\alpha,p,q}\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}^2.\end{aligned}$$

The second last term in (5.31) can be estimated by

$$\begin{aligned}\|\pi(T_{F'(\tilde{u})}u^{\#},\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q} &\lesssim \|T_{F'(\tilde{u})}u^{\#}\|_{2\alpha,p/2,q}\|\xi_{\mathcal{T}}\|_{\alpha-1,p,q}\\ &\lesssim \|F'\|_\infty\|u^{\#}\|_{2\alpha,p/2,q}\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}\end{aligned}$$

where a Besov embedding, Lemma 5.1.1 and (5.10) are used. Finally, for the last term, note that there is some $\epsilon \in (0,\alpha-\frac{1}{p})$ such that $3\alpha-1-\epsilon>0$. Applying Lemma 5.1.1, Step 1 and Besov embeddings, we get

$$\begin{aligned}\|\pi(R_F(\tilde{u}),\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q} &\lesssim \|\pi(R_F(\tilde{u}),\xi_{\mathcal{T}})\|_{3\alpha-1-\epsilon,p/3,q}\\ &\lesssim \|R_F(\tilde{u})\|_{2\alpha-\epsilon,p/2,q}\|\xi_{\mathcal{T}}\|_{\alpha-1,p,q}\\ &\lesssim \|F\|_{C_b^2}(1+\|F\|_\infty\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q})(\|\tilde{u}\|_{\alpha,p,q}+\|u^{\#}\|_{2\alpha,p/2,q})\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}.\end{aligned}$$

These five estimates combined lead to the asserted bound. □

Having established a linear upper bound for the resonant term $\pi(F(\tilde{u}),\xi_{\mathcal{T}})$, we deduce the boundedness of the solution to the localized RDE (5.26) in the weighted Besov norm.

Corollary 5.4.6. *Let $\alpha \in (1/3,1/2)$, $p \geq 3$, $q \geq 1$ and $F \in C_b^2$ with $F(0)=0$. Let $\vartheta_{\mathcal{T}} \in C_{\mathcal{T}}^\infty$ with derivative $\xi_{\mathcal{T}} = \mathrm{d}\vartheta_{\mathcal{T}}$. If the bound*

$$\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2 < c(\mathcal{T}^3 \vee 1)\big(\|\vartheta_{\mathcal{T}}\|_{\alpha-1,p,q}+\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}^2+\|\pi(\vartheta_{\mathcal{T}},\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\big)^{-1}$$

holds for a universal constant $c>0$, independent of ϑ, F, u_0 and if ψ satisfies Assumption 2 for some sufficiently small C_ψ, then the solution u to (5.26) satisfies

$$\begin{aligned}\|\psi u\|_{\alpha,p,q} \lesssim& (\mathcal{T}^2 \vee 1)\big(|u(0)| + (\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^3)(\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}+1)\\ &\times(\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}+\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}^2+\|\pi(\vartheta_{\mathcal{T}},\xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q})\big).\end{aligned}$$

Proof. We recall the characterization of $\tilde{u} = \psi u$ from Lemma 5.4.4. In order to obtain the desired estimate of the norm, we apply Bony's decomposition and calculate

$$\begin{aligned}
\mathrm{d}u^{\#} &= F(\tilde{u})\xi_{\mathcal{T}} - \mathrm{d}(T_{F(\tilde{u})}\vartheta_{\mathcal{T}}) + \frac{\psi'}{\psi}\tilde{u} \\
&= T_{F(\tilde{u})}\xi_{\mathcal{T}} + \pi(F(\tilde{u}), \xi_{\mathcal{T}}) + T_{\xi_{\mathcal{T}}}(F(\tilde{u})) - \mathrm{d}(T_{F(\tilde{u})}\vartheta_{\mathcal{T}}) + \frac{\psi'}{\psi}\tilde{u} \\
&= \pi(F(\tilde{u}), \xi_{\mathcal{T}}) + T_{\xi_{\mathcal{T}}}(F(\tilde{u})) - T_{\mathrm{d}(F(\tilde{u}))}\vartheta_{\mathcal{T}} + \frac{\psi'}{\psi}\tilde{u}. \qquad (5.33)
\end{aligned}$$

We bound the $B^{2\alpha-1}_{p/2,q}$-norm of these four terms separately. The first term is bounded by Proposition 5.4.5. To estimate the second term in (5.33), Lemma 5.1.1, (5.10) and a Besov embedding yield

$$\begin{aligned}
\|T_{\xi_{\mathcal{T}}}(F(\tilde{u}))\|_{2\alpha-1,p/2,q} &\lesssim \|F\|_{C_b^1}\|\xi_{\mathcal{T}}\|_{\alpha-1,p,2q}\|\tilde{u}\|_{\alpha,p,2q} \\
&\lesssim \|F\|_{C_b^1}\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}\|\tilde{u}\|_{\alpha,p,q}.
\end{aligned}$$

The third term in (5.33) can be estimated with the lifting property of Besov spaces, Lemma 5.1.1, (5.10) and a Besov embedding

$$\begin{aligned}
\|T_{\mathrm{d}(F(\tilde{u}))}\vartheta_{\mathcal{T}}\|_{2\alpha-1,p/2,q} &\lesssim \|\,\mathrm{d}F(\tilde{u})\|_{\alpha-1,p,2q}\|\vartheta_{\mathcal{T}}\|_{\alpha,p,2q} \\
&\lesssim \|F(\tilde{u})\|_{\alpha,p,2q}\|\vartheta_{\mathcal{T}}\|_{\alpha,p,2q} \lesssim \|F\|_{C_b^1}\|\tilde{u}\|_{\alpha,p,q}\|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}.
\end{aligned}$$

For the last term in (5.33) we note the norm equivalence $\|\psi u\|_{L^{p/2}} \sim \|\tilde{\psi} u\|_{L^{p/2}}$ with for u being constant outside of $[-2\mathcal{T}, 2\mathcal{T}]$, where we set $\tilde{\psi} := \psi\psi_2$ for another weight function ψ_2 satisfying Assumption 2. Hence, $\|\tilde{u}\|_{L^{p/2}} \lesssim \|\psi_2\tilde{u}\|_{L^{p/2}} \le \|\psi_2\|_{L^p}\|\tilde{u}\|_{L^p}$ by Hölder's inequality. Since $2\alpha - 1 < 0$, a Besov embedding yields

$$\|\frac{\psi'}{\psi}\tilde{u}\|_{2\alpha-1,p/2,q} \lesssim \|\frac{\psi'}{\psi}\tilde{u}\|_{L^{p/2}} \lesssim \|\frac{\psi'}{\psi}\|_{\infty}\|\tilde{u}\|_{L^{p/2}} \lesssim (\mathcal{T} \vee 1)\|\frac{\psi'}{\psi}\|_{\infty}\|\tilde{u}\|_{L^p}.$$

Combining all the above estimates, we obtain

$$\begin{aligned}
\|\,\mathrm{d}u^{\#}\|_{2\alpha-1,p/2,q} \lesssim& C_{\xi,\vartheta}(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2)(\|\tilde{u}\|_{\alpha,p,q} + \|u^{\#}\|_{2\alpha,p/2,q}) \\
&+ (\mathcal{T} \vee 1)\|\frac{\psi'}{\psi}\|_{\infty}\|\tilde{u}\|_{\alpha,p,q}
\end{aligned}$$

with

$$C_{\xi,\vartheta} := \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q} + \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}^2 + \|\pi(\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}})\|_{2\alpha-1,p/2,q}.$$

Applying again the lifting property of Besov spaces [Tri10, Thm. 2.3.8] together with the definition of $u^{\#}$, $\|\tilde{u}\|_{L^{p/2}} \lesssim (\mathcal{T} \vee 1)\|\tilde{u}\|_{L^p}$ and the compact support of $\vartheta_{\mathcal{T}}$, we have

$$\begin{aligned}
\|u^{\#}\|_{2\alpha,p/2,q} \lesssim& \|u^{\#}\|_{L^{p/2}} + \|\,\mathrm{d}u^{\#}\|_{2\alpha-1,p/2,q} \\
\le& \|T_{F(\tilde{u})}\vartheta_{\mathcal{T}}\|_{L^{p/2}} + \|\tilde{u}\|_{L^{p/2}} + \|\,\mathrm{d}u^{\#}\|_{2\alpha-1,p/2,q} \\
\lesssim& (\mathcal{T} \vee 1)(\|F\|_{\infty}\|\vartheta_{\mathcal{T}}\|_{L^p} + \|\tilde{u}\|_{L^p}) + \|\,\mathrm{d}u^{\#}\|_{2\alpha-1,p/2,q}. \qquad (5.34)
\end{aligned}$$

Hence, combining the last two inequalities leads to

$$\begin{aligned}\|\,\mathrm{d}u^{\#}\|_{2\alpha-1,p/2,q} \lesssim & C_{\xi,\vartheta}(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2)\big(\|\tilde{u}\|_{\alpha,p,q} + \|\,\mathrm{d}u^{\#}\|_{2\alpha-1,p/2,q}\big) \\ & + (\mathcal{T} \vee 1)\big(C_{\xi,\vartheta}(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2)(\|F\|_\infty \|\vartheta_{\mathcal{T}}\|_{L^p} \\ & + \|\tilde{u}\|_{L^p}) + \|\frac{\psi'}{\psi}\|_\infty \|\tilde{u}\|_{\alpha,p,q}\big).\end{aligned}$$

If $C_{\xi,\vartheta}(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2)$ is sufficiently small, we thus obtain

$$\begin{aligned}&\|\,\mathrm{d}u^{\#}\|_{2\alpha-1,p/2,q} \\ &\quad \lesssim (\mathcal{T} \vee 1) C_{\xi,\vartheta}(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2)\big(\|\tilde{u}\|_{\alpha,p,q} + \|F\|_\infty \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}\big) \\ &\quad\quad + (\mathcal{T} \vee 1)\|\frac{\psi'}{\psi}\|_\infty \|\tilde{u}\|_{\alpha,p,q}.\end{aligned} \tag{5.35}$$

In combination with the ansatz and the bounds from above, Lemma 5.1.1 reveals

$$\begin{aligned}\|\,\mathrm{d}\tilde{u}\|_{\alpha-1,p,q} &\leq \|\,\mathrm{d}(T_{F(\tilde{u})}\vartheta_{\mathcal{T}})\|_{\alpha-1,p,q} + \|\,\mathrm{d}u^{\#}\|_{\alpha-1,p,q} \\ &\lesssim \|T_{\mathrm{d}F(\tilde{u})}\vartheta_{\mathcal{T}}\|_{2\alpha-1,p/2,q} + \|T_{F(\tilde{u})}\xi_{\mathcal{T}}\|_{\alpha-1,p,q} + \|\,\mathrm{d}u^{\#}\|_{2\alpha-1,p/2,q} \\ &\lesssim (\mathcal{T} \vee 1)\Big(C_{\xi,\vartheta}(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2)\big(\|\tilde{u}\|_{\alpha,p,q} + \|F\|_\infty \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q} + 1\big) + \|\frac{\psi'}{\psi}\|_\infty \|\tilde{u}\|_{\alpha,p,q}\Big).\end{aligned}$$

Due to Remark 5.4.3 applied to $\tilde{\psi} = \psi\psi_2$, we can apply Lemma 5.1.2 to obtain

$$\begin{aligned}\|\tilde{u}\|_{\alpha,p,q} &\lesssim \|\psi_2 \tilde{u}\|_{\alpha,p,q} \leq (\mathcal{T}^2 \vee 1)\big(|u(0)| + \|\,\mathrm{d}\tilde{u}\|_{\alpha-1,p,q}\big). \\ &\lesssim \underbrace{\Big((\mathcal{T}^3 \vee 1) C_{\xi,\vartheta}(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2) + \|\frac{\psi'}{\psi}\|_\infty\Big)}_{=:D} \|\tilde{u}\|_{\alpha,p,q} \\ &\quad + (\mathcal{T}^2 \vee 1)\big(|u(0)| + C_{\xi,\vartheta}(\|F\|_{C_b^2} \vee \|F\|_{C_b^2}^2)(\|F\|_\infty \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q} + 1)\big).\end{aligned}$$

For D smaller than some universal constant we conclude the assertion. $\square$

For any $F \in C_b^3$ and $\|F\|_{C_b^3}$ small enough, the following lemma reveals that the weighted Itô-Lyons map $\hat{S}$ as introduced in (5.27) is locally Lipschitz continuous with respect to the Besov norms on $\mathbb{R}^d \times B_{p,q}^{\alpha-1} \times B_{p/2,q}^{2\alpha-1}$ and thus it can be uniquely extended in a continuous way.

Lemma 5.4.7. *Let $\alpha \in (1/3, 1/2)$, $p \geq 3$, $q \geq 1$ and let $F \in C_b^3$ with $F(0) = 0$. Assume ψ is a weight function satisfying Assumption 2 and let $\vartheta_{\mathcal{T}} \in C_0^\infty$ with derivative $\xi_{\mathcal{T}} = \mathrm{d}\vartheta_{\mathcal{T}}$. Then there exits a polynomial on $\mathbb{R}^3$ such that, provided the bound*

$$\|F\|_{C_b^3} + \|F\|_{C_b^2}^3 \leq P(\mathcal{T} \vee 1, \|\vartheta_{\mathcal{T}}\|_{\alpha,p,q}, \|\pi(\xi_{\mathcal{T}}, \vartheta_{\mathcal{T}})\|_{2\alpha-1,p,q})^{-1},$$

holds and C_ψ is sufficiently small, there exists for every $u_0 \in \mathbb{R}^d$ a unique global solution $u \in \mathcal{S}'$ with $\psi u \in B_{p,q}^\alpha$ to the Cauchy problem (5.26). Furthermore, for fixed $\mathcal{T}$, ψ and F the weighted Itô-Lyons map $\hat{S}$ is local Lipschitz continuous on $\mathbb{R}^d \times C_{\mathcal{T}}^\infty$ around $(u_0, \vartheta_{\mathcal{T}}, \pi(\vartheta_{\mathcal{T}}, \xi_{\mathcal{T}}))$.

The local Lipschitz continuity is the key ingredient to extend the weighted Itô-Lyons map from smooth paths to irregular ones. The proof works similarly to the proofs of Proposition 5.4.5 and Corollary 5.4.6 with an additional application of the Lipschitz result in Proposition 5.3.1. Due to the necessary, but quite lengthy estimations, we postpone the proof to Appendix A.5 with the hope to increase the readability of the chapter.

The requirement $F(0) = 0$ seems to be a purely technical assumption. However, we decided not to get rid of this condition because it would only make all estimates even more involved without the need of conceptually new ideas.

Finally, we can state our main result: There exist a continuous extension of the weighted Itô-Lyons map $\hat{S}$ from $\mathbb{R}^d \times C^\infty_{\mathcal{T}}$ to the domain $\mathbb{R}^d \times \mathcal{B}^{0,\alpha}_{p,q}$. Similarly to Theorem 5.2.1 we use a dilation argument together with a localization procedure to circumvent the assumption that $\|F\|_{C^3_b}$ has to be small. Allowing for general Besov spaces, this theorem generalizes Lyons' celebrated Universal Limit Theorem [LQ02, Thm. 6.2.2] and in particular [GIP12, Thm. 3.3].

Theorem 5.4.8. *Let $\mathcal{T} > 0$, $\alpha \in (1/3, 1/2)$, $p \geq 3$, $q \geq 1$ and $F \in C^3_b$ with $F(0) = 0$. If the weight function ψ satisfies Assumption 2 with C_ψ sufficiently small, then the weighted Itô-Lyons map $\hat{S}$ as introduced in (5.27) can be continuously extended from $\mathbb{R}^d \times C^\infty_{\mathcal{T}}$ to* the domain *$\mathbb{R}^d \times \mathcal{B}^{0,\alpha}_{p,q}$. In particular, there exists a unique solution to (5.27) for any geometric Besov rough path $(\vartheta_{\mathcal{T}}, \pi(\vartheta_{\mathcal{T}}, \mathrm{d}\vartheta_{\mathcal{T}})) \in \mathcal{B}^{0,\alpha}_{p,q}$.*

An elementary formulation of Theorem 5.4.8 is presented in the next lemma. The proof of Theorem 5.4.8 is then an immediate consequence.

Lemma 5.4.9. *Assume the weight function ψ satisfies Assumption 2 with C_ψ sufficiently small. Let $\mathcal{T} > 0$, $\alpha \in (1/3, 1/2)$, $p \geq 3$, $q \geq 1$ and $F \in C^3_b$ with $F(0) = 0$. Let further $u_0 \in \mathbb{R}^m$ be an initial condition and $(\vartheta_{\mathcal{T}}, \eta_{\mathcal{T}}) \in \mathcal{B}^{0,\alpha}_{p,q}$ be a geometric Besov rough path. Let $(\vartheta^n_{\mathcal{T}}) \subset C^\infty_{\mathcal{T}}$ be a sequence of functions with corresponding derivatives $(\xi^n_{\mathcal{T}})$ and $(u^n_0) \subset \mathbb{R}^m$ be a sequence of initial conditions such that $(u^n_0, \vartheta^n_{\mathcal{T}}, \pi(\vartheta^n_{\mathcal{T}}, \xi^n_{\mathcal{T}}))$ converges to $(u_0, \vartheta_{\mathcal{T}}, \eta_{\mathcal{T}})$ in $\mathbb{R}^m \times B^{\alpha-1}_{p,q} \times B^{2\alpha-1}_{p/2,q}$. Denote by u^n the unique solution to the Cauchy problem (5.26) with u^n_0 and $\xi^n_{\mathcal{T}}$ for all $n \in \mathbb{N}$. Then there exists $u \in \mathcal{S}'$ such that $\psi u \in B^\alpha_{p,q}$ and $\psi u^n \to \psi u$ in $B^\alpha_{p,q}$. The limit u depends only on $(u_0, \vartheta_{\mathcal{T}}, \eta_{\mathcal{T}})$ and not on the approximating family $(u^n_0, \vartheta^n_{\mathcal{T}}, \pi(\vartheta^n_{\mathcal{T}}, \xi^n_{\mathcal{T}}))$.*

Proof. In order to apply Lemma 5.4.7, we first need to ensure that $\|F\|_{C^3_b}$ is small enough. Thus, as similarly done in Step 2 of the proof of Theorem 5.2.1, we scale $\vartheta^n_{\mathcal{T}}$: For some fixed $\epsilon \in (0, \alpha - 1/p)$ and for $\lambda \in (0,1)$ we set

$$\vartheta^{n,\lambda}_{\mathcal{T}} := \lambda^{-\alpha+1/p+\epsilon} \Lambda_\lambda \vartheta^n_{\mathcal{T}} \quad \text{and} \quad \xi^{n,\lambda}_{\mathcal{T}} := \lambda^{1-\alpha+1/p+\epsilon} \Lambda_\lambda \xi^n_{\mathcal{T}},$$

where we recall the scaling operator $\Lambda_\lambda f = f(\lambda \cdot)$ for $f \in \mathcal{S}'$. Given this scaling, still $\xi^{n,\lambda}_{\mathcal{T}} = \mathrm{d}\vartheta^{n,\lambda}_{\mathcal{T}}$ holds true and the corresponding norms of $\xi^{n,\lambda}_{\mathcal{T}}$ and $\vartheta^{n,\lambda}_{\mathcal{T}}$ can be controlled by the Lemmas 5.1.2 and 5.1.3, i.e.

$$\begin{aligned} &\|\xi^{n,\lambda}_{\mathcal{T}}\|_{\alpha-1,p,q} \lesssim \|\xi^n_{\mathcal{T}}\|_{\alpha-1,p,q} \quad \text{and} \\ &\|\vartheta^{n,\lambda}_{\mathcal{T}}\|_{\alpha,p,q} \lesssim (1 \vee \mathcal{T}^2) \|\xi^{n,\lambda}_{\mathcal{T}}\|_{\alpha-1,p,q} \lesssim (1 \vee \mathcal{T}^2) \|\xi^n_{\mathcal{T}}\|_{\alpha-1,p,q}. \end{aligned}$$

Moreover, again using Lemma 5.1.3 we can estimate

$$\begin{aligned}\|\pi(\vartheta_{\mathcal{T}}^{n,\lambda},\xi_{\mathcal{T}}^{n,\lambda})\|_{2\alpha-1,p/2,q} &= \lambda^{1-2\alpha+2/p+2\epsilon}\|\pi(\Lambda_\lambda\vartheta_{\mathcal{T}}^n,\Lambda_\lambda\xi_{\mathcal{T}}^n)\|_{2\alpha-1,p/2,q}\\ &\lesssim (\lambda^{2\epsilon}|\log\lambda|+\lambda^{1-2\alpha+2\epsilon})\|\pi(\vartheta_{\mathcal{T}}^n,\xi_{\mathcal{T}}^n)\|_{2\alpha-1,p/2,q}.\end{aligned}$$

Let us take once more the localization function φ from Assumption 1 and noticing that $\varphi_{2\mathcal{T}}\vartheta_{\mathcal{T}}^n=\vartheta_{\mathcal{T}}^n$ for all $n\in\mathbb{N}$. Therefore, Lemma 5.4.7 provides for $\lambda>0$ sufficiently small a unique global solution $u^{n,\lambda}\in B_{p,q}^\alpha$ to

$$\mathrm{d}u^{n,\lambda}=\lambda^{\alpha-1/p-\epsilon}F(u^{n,\lambda})\,\mathrm{d}(\varphi_{2\mathcal{T}}\vartheta_{\mathcal{T}}^{n,\lambda}),\quad u^{n,\lambda}(0)=u_0^n.$$

Setting now $u^n:=\Lambda_{\lambda^{-1}}u^{n,\lambda}$, we have constructed a unique global solution to

$$\mathrm{d}u^n=F(u^n)\,\mathrm{d}(\varphi_{2\lambda\mathcal{T}}\vartheta_{\mathcal{T}}^n),\quad u(0)=u_0^n.$$

Since $(u_0^n,\vartheta_{\mathcal{T}}^{n,\lambda},\pi(\vartheta_{\mathcal{T}}^{n,\lambda},\xi_{\mathcal{T}}^{n,\lambda}))$ converges to $(u_0,\vartheta_{\mathcal{T}}^\lambda,\pi(\vartheta_{\mathcal{T}}^\lambda,\xi_{\mathcal{T}}^\lambda))$ in $\mathbb{R}^d\times B_{p,q}^{\alpha-1}\times B_{p/2,q}^{2\alpha-1}$, the continuity of the Itô-Lyons map established in Lemma 5.4.7 implies that $u^{n,\lambda}$ converges to some u^λ in $B_{p,q}^\alpha$ weighted by ψ. Therefore, the solution u^n converges to $u:=\Lambda_{\lambda^{-1}}u^\lambda$ in $B_{p,q}^\alpha$ weighted by ψ, due to Lemma 5.1.3 and 5.1.3, which can be seen analogously to Step 2 of the proof of Theorem 5.2.2. We note that $u|_{[-\lambda\mathcal{T},\lambda\mathcal{T}]}$ does not depend on $\varphi_{\lambda\mathcal{T}}$.

Following the same argumentation as in Step 3 of the proof of Theorem 5.2.2, we can iterate this construction of u^n and u on intervals of the length $2\lambda T$. In this way we end up with a continuous function u such that $\psi u\in B_{p,q}^\alpha$ and ψu^n converges to ψu in $B_{p,q}^\alpha$. Note that u depends only on $(u_0,\vartheta_{\mathcal{T}},\pi(\vartheta_{\mathcal{T}},\xi_{\mathcal{T}}))$ but neither on approximating family $(u_0^n,\vartheta_{\mathcal{T}}^n,\pi(\vartheta_{\mathcal{T}}^n,\xi_{\mathcal{T}}^n))$ nor on $\varphi_{\lambda\mathcal{T}}$. □

While general Besov spaces contain functions with jumps, the paracontrolled distribution approach to rough differential equations as explored in the present section only studies continuous functions. Therefore, we think a discussion is in order why the paracontrolled distribution approach seems to be naturally restricted to continuous functions.

Remark 5.4.10. *The results in Section 5.3 apply only to Besov spaces $B_{p,q}^\alpha$ for $p\geq 1$. According to (5.20), our estimates result in a bound of the $B_{p/3,q}^{3\alpha-1}$-norm. Consequently, we require $p\geq 3$ and $\alpha>1/3$ in order to have positive regularity. In particular, our main theorem applies only to the case $\alpha>1/p$ which implies that $B_{p,q}^\alpha$ embeds into the space of continuous functions.*

If we want to extend our results to discontinuous functions, corresponding to $\alpha<1/p$, then we could hope that it helps to verify the previous results for $p<1$. Let us sketch some details on this idea, where we have to deal with the quasi-Banach space $B_{p,q}^\alpha$ for $p<1$. In that case the triangle inequality only holds true up to a multiplicative constant

$$\|f+g\|_{\alpha,p,q}\leq 2^{1/p-1}(\|f\|_{\alpha,p,q}+\|g\|_{\alpha,p,q})\quad \textit{for } f,g\in B_{p,q}^\alpha.$$

Following the lines of the proof of Lemma 2.84 (or Lemma 2.49 respectively) in Bahouri [BCD11], we obtain in the case $p \in (0,1)$*,* $q > 1$*,* $\alpha > 1/p - 1$*, for* $u := \sum_j u_j$ *with* $\operatorname{supp} u_j \subset 2^j \mathcal{B}$ *for some ball* $\mathcal{B}$ *that*

$$\|u\|_{s-(1/p-1),p,q} \lesssim \big\| \big(2^{js} \|u_j\|_{L^p}\big)_j \big\|_{\ell^q},$$

provided the right-hand side is finite. For the commutator lemma in the case $p \in (0,1)$ *we thus cannot hope for more than the following: Replacing the assumption* $p \geq 1$ *with* $\alpha + \beta + \gamma > (\frac{1}{p} - 1) \vee 0$ *in the situation of Lemma 5.3.4, we conjecture*

$$\|\Gamma(f,g,h)\|_{\alpha+\beta+\gamma-(\frac{1}{p}-1)\vee 0,p,q} \lesssim \|f\|_{\alpha,p_1,q} \|g\|_{\beta,p_2,q} \|h\|_{\gamma,p_3,q}.$$

Applying this bound to (5.20), we obtain for $p \in (0,1)$

$$\|\Pi_F(u,\xi)\|_{3\alpha-1-(3/p-1),p/3,q} < \big(\|F'(u)\|_{\alpha,p,q} + \|u\|_{\alpha,p,q}\big) \|u\|_{\alpha,p,q} \|\xi\|_{\alpha-1,p,q}.$$

However, $3\alpha - 1 - (3/p - 1) > 0$ *is equivalent to* $\alpha > 1/p$*, which is the same condition as we had before, excluding discontinuous functions.*

Alternatively, a higher order expansion in the linearization Lemma 5.3.2 could be studied (corresponding to more additional information). If such a second order expansion would succeed, we may have the condition $4\alpha - 1 > 0$*, but with the price of imposing* $p/4 \geq 1$*. Consequently, we would again obtain* $\alpha > 1/p$*.*

In conclusion, it appears natural that this approach is restricted to continuous functions.

5.5. Stochastic differential equations

The purely analytic results from the previous sections for rough differential equations allow for treating a large class of stochastic differential equations (SDEs) in a pathwise way. While we assumed so far that the driving signal ξ of the RDE (5.1) is given by a deterministic function with a certain Besov regularity, we suppose from now on that ξ is the distributional derivative of some continuous stochastic process X. Provided all involved stochastic objects live on a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and setting $\xi := \mathrm{d}X$, the RDE (5.1) becomes an SDE with the dynamic

$$\mathrm{d}u(t) = F(u(t))\,\mathrm{d}X_t, \quad u(0) = u_0, \quad t \in [0,1], \tag{5.36}$$

where u_0 is a random variable in $\mathbb{R}^m$ and X is some d-dimensional stochastic process for simplicity on the interval $[0,1]$.

Instead of relying on classical stochastic integration in order to give the SDE (5.36) a meaning, we shall demonstrate here that the results of Section 5.2 and 5.4 are feasible for a wide class of SDEs. For this propose the present section is devoted to show the required sample path properties of a couple of stochastic processes. This allows for solving SDEs which are beyond the scope of classical probability theory as well as for recovering well-known examples. Let us emphasize that we present here only a few exemplary stochastic processes to illustrate our results and do not aim for the most general class of stochastic processes.

Gaussian processes

A well-known but very common example for a stochastic driving signal X is the fractional Brownian motion, cf. [Cou07, Mis08]. A d-dimensional fractional Brownian motion $B^H = (B^1, \dots, B^d)$ with Hurst index $H \in (0,1)$ is a Gaussian process with zero mean, independent components, and covariance function given by

$$\mathbb{E}[B_s^i B_t^i] = \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H}), \quad s,t \in [0,1],$$

for $i = 1, \dots, d$. The Besov regularity of (fractional) Brownian motion is already know for a long time due to Roynette [Roy93] and Ciesielski [CKR93]: it holds $(B_t^H)_{t\in[0,1]} \in B^H_{p,\infty}([0,1], \mathbb{R}^d)$ almost surely for any $p \in [1,\infty]$ and $(B_t^H)_{t\in[0,1]} \notin B^H_{p,q}([0,1], \mathbb{R}^d)$ almost surely if $q < \infty$, see for instance [Ver09, Corrollary 5.3]. More recently, Veraar [Ver09] investigated the Besov regularity for more general Gaussian processes. The self-similar behavior of fractional Brownian motion implies that B^H has the same regularity H with respect to all p-scales of the Besov spaces. Therefore, it suffices to focus on $p = \infty$ for this example.

Even if one could still rely on results from rough path theory (Lyons [Lyo98] or Gubinelli et al. [GIP12]) in the case $H > 1/3$, the following lemma shows how to recover the results for SDEs with our machinery. It in particular covers the fractional Brownian motion.

Lemma 5.5.1 ([GIP12, Cor. 3.10])**.** *Let X be a centered d-dimensional Gaussian process with independent components whose covariance function fulfills for some $H \in (1/4, 1)$ the Coutin-Qian condition*

$$\begin{aligned} \mathbb{E}[|X_t - X_s|^2] &\lesssim |t-s|^{2H} \quad \text{and} \\ |\mathbb{E}[(X_{s+r} - X_s)(X_{t+r} - X_t)]| &\lesssim |t-s|^{2H-2} r^2, \end{aligned} \tag{5.37}$$

for all $s,t \in \mathbb{R}$ and all $r \in [0, |t-s|)$. For every $\alpha < H$ and any smooth function φ with compact support we have $\varphi X \in B^\alpha_{\infty,\infty}$. Moreover, there exists an $\eta \in B^{2\alpha-1}_{\infty,\infty}$ such that for every $\delta > 0$ and every $\psi \in \mathcal{S}$ with $\int \psi(t)\,\mathrm{d}t = 1$ it holds

$$\lim_{n\to\infty} \mathbb{P}\big(\|\psi^n * (\varphi X) - (\varphi X))\|_{\alpha,\infty,\infty} + \|\pi(\psi^n * (\varphi X),\, \mathrm{d}(\psi^n * (\varphi X)) - \eta)\|_{2\alpha-1,\infty,\infty} > \delta\big) = 0,$$

where we denote $\psi^n := n\psi(n\cdot)$.

In other words, every d-dimensional Gaussian process X satisfying the Coutin-Qian condition (5.37) for some $H \in (1/3, 1/2)$ can be enhanced to a geometric Besov rough path and especially Theorem 5.4.8 can be applied to solve the SDE (5.36), cf. Coutin and Qian [CQ02] or Friz and Victoir [FV10a].

Stochastic processes via Schauder expansions

Instead of approximating stochastic processes by processes with smooth sample paths, in probability theory it is often more convenient to construct a process via an expansion with respect to a basis of L^2. The presumably most famous construction of this type is the Karhunen-Loève expansion of Gaussian processes.

A classical construction of a Brownian motion on the interval $[0,1]$ is the Lévy-Ciesielski construction based on Schauder functions. More generally, Schauder functions are a very frequently applied tool in stochastic analysis. Notably, they are used to investigate the Besov regularity of stochastic processes, cf. for example Ciesielski et al. [CKR93] and Rosenbaum [Ros09], and very recently Gubinelli et al. [GIP14] constructed directly the rough path integral in terms of Schauder expansions.

The Schauder functions can be defined as the antiderivatives of the Haar functions. More explicitly they are given by

$$G_{j,k}(t) := 2^{-j/2}\psi(2^j t-(k-1)) \quad \text{with} \quad \psi(t) := t\mathbf{1}_{[0,1/2]}(t)-(t-\tfrac{1}{2})\mathbf{1}_{(1/2,1]}(t), \quad t\in\mathbb{R},$$

for $j\in\mathbb{N}$ and $1\le k\le 2^n$, and $G_{0,0}(0):=1$. The Haar functions form a basis of $L^2([0,1],\mathbb{R})$ and it is obvious that $G_{n,k}\in B^{\beta}_{p,q}$ for $0<\beta<1$ and $p,q\in[1,\infty]$ with $\beta>1/p$, cf. [Ros09, Prop. 9]. The next lemma explains why an approximation of stochastic processes in terms of Schauder expansions can also be used to show that a process can be enhanced to a geometric Besov rough path.

Lemma 5.5.2. *Let $\alpha\in(1/3,1/2)$, $\beta\in(1/2,1]$, $p\ge 2$ and $q\ge 1$. Suppose $(f^n)\subset B^{\beta}_{p,q}$ is a sequence of functions such that* $\operatorname{supp} f^n\subset[0,1]$ *for all $n\in\mathbb{N}$. If $(f^n,\pi(f^n,\,\mathrm{d}f^n))$ converges in $B^{\alpha}_{p,q}\times B^{2\alpha-1}_{p/2,q}$ to some $(f,\pi(f,\,\mathrm{d}f))\in B^{\alpha}_{p,q}\times B^{2\alpha-1}_{p/2,q}$, then $(f,\pi(f,\,\mathrm{d}f))\in\mathcal{B}^{0,\alpha}_{p,q}$.*

Proof. Let us recall that C^{∞}_1 is dense in $\{g\in B^{\beta}_{p,q} : \operatorname{supp} g\subset[0,1]\}$. Hence, for every $n\in\mathbb{N}$ there exists a sequence of smooth functions $(f^{n,m})_m\subset C^{\infty}_1$ such that $(f^{n,m},\,\mathrm{d}f^{n,m})$ converges to $(f^n,\,\mathrm{d}f^n)$ in $B^{\beta}_{p,q}\times B^{\beta-1}_{p,q}$ as m goes to infinity, where the convergence of the second component follows by the first one using the lifting property of Besov spaces. Since $\beta>1/2$, we also have by Lemma 5.1.1 that $\pi(f^{n,m},\,\mathrm{d}f^{n,m})$ converges to $\pi(f^n,\,\mathrm{d}f^n)$ as m goes to infinity. Therefore, taking a diagonal sequence there exists a sequence of smooth functions $(f^{n,m(n)})_n\subset C^{\infty}_1$ such that $(f,\pi(f,\,\mathrm{d}f)) = \lim_{n\to\infty}\pi(f^{n,m(n)},\,\mathrm{d}f^{n,m(n)})$ where the limit is taken in $B^{\alpha}_{p,q}\times B^{2\alpha-1}_{p/2,q}$. □

Based on Lemma 5.5.2 it is now an immediate consequence of Theorem 6.5 and 6.6. in [GIP14] that suitable hypercontractive processes and continuous martingales can be lifted to geometric Besov rough paths since the Lévy area term in [GIP14] corresponds to our resonant term. Especially, all examples from probability theory in [GIP14] are feasible with our results as well.

Random functions via wavelet expansions: a prototypical example

Motivated from the previous construction, we shall consider as a last example more general stochastic processes which can be constructed as series expansion with random coefficients and with respect to a wavelet basis. There are several applications of such models, for instance, in non-parametric Bayesian statistics to construct priors on function spaces. One advantage is that the sample path regularity of such processes can be determined precisely, cf. Abramovich et al. [ASS98], Chioica et al. [CDD+12] and Bochkina [Boc13].

Wavelets can be taken to be localized in the time domain as well as in the Fourier domain. The latter property is quite convenient when working with Littlewood-Paley theory as we demonstrate in the following. Let $\{\psi_{j,k} : j \in \mathbb{N}, k \in \mathbb{Z}\}$ be an orthonormal wavelet basis of $L^2(\mathbb{R})$, where $\psi_{j,k}(t) := 2^{j/2}\psi(2^j t - k)$ for $j \geq 1$, $k \in \mathbb{Z}$, $t \in \mathbb{R}$, and $\psi \in L^2(\mathbb{R})$. Then, any function $f \in L^2(\mathbb{R})$ can be written as

$$f(t) := \sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}}\langle f, \psi_{j,k}\rangle\psi_{j,k}(t), \quad t \in \mathbb{R}, \quad \text{with} \quad \langle f, \psi_{j,k}\rangle := \int_{\mathbb{R}} f(s)\psi_{j,k}(s)\,\mathrm{d}s.$$

Replacing the deterministic wavelet coefficients with real valued random variables $(Z_{j,k})_{j,k}$, we now study stochastic processes of the type

$$X_t := \sum_{j\geq 0}\sum_{k=-2^j}^{2^j} Z_{j,k}\psi_{j,k}(t), \quad t \in \mathbb{R}. \tag{5.38}$$

Without loss of generality, we truncated the series expansion in k since we always have to localize the signal in order to apply our results concerning RDEs, see the equations (5.11) and (5.26). Let us impose the following weak assumptions on $(Z_{j,k})_{j,k}$ and $(\psi_{j,k})_{j,k}$:

Assumption 3. Let $\{\psi_{j,k} : j \in \mathbb{N}, k \in \mathbb{Z}\}$ be an orthonormal and band limited wavelet basis of $L^2(\mathbb{R})$ and suppose $Z_{j,k} = A_{j,k}B_{j,k}$ for all $j \geq 0$ and $k = -2^j, \dots, 2^j$ where

- $(A_{j,k})_{j,k}$ are random variables satisfying $\mathbb{E}[A_{j,k}^p]^{1/p} \lesssim 2^{-js}$ for some $s > 0$ and $p \in \{2, 4\}$,
- $\mathbb{E}[A_{j,k}] = 0$ for all j, k and $\mathbb{E}[A_{j,k}A_{m,n}] = 0$ for $j \neq m$ or $k \neq n$,
- $(B_{j,k})_{j,k}$ are Bernoulli random variables with $\mathbb{P}(B_{j,k} = 1) = 2^{-jr}$ for some $r \in [0, 1)$,
- $\mathbb{E}[A_{j,k}B_{j,k}A_{m,n}B_{m,n}] = \mathbb{E}[A_{j,k}A_{m,n}]\mathbb{E}[B_{j,k}B_{m,n}]$ for all j, k, m, n.

The assumption allows for a quite flexible class of stochastic processes although it is chosen in a way to keep the required analysis simple. Having in mind the construction of Brownian motion via Schauder functions, as mentioned before, the process X behaves like a Wiener process if $(Z_{j,k})_{j,k}$ are i.i.d. standard normal distributed random variables with $s = 1$. In particular, the self-similar behavior of Brownian motion is then achieved because all wavelet coefficients at a level j are of the same order of magnitude (especially $r = 0$). If instead $r \in (0, 1)$, we expect only a number of $2 \cdot 2^{j(1-r)}$ non-zero wavelet coefficients at each level j and we consequently gain from measuring the regularity of X in a $B^{\alpha}_{p,q}$-norm for some finite p.

In order to profit from $(Z_{j,k})_{j,k}$ being uncorrelated we choose an even number p. Together with the requirement $p \geq 3$ in our uniqueness and existence theorem for RDEs (Theorem 5.4.8), we thus take $p = 4$. Keeping in mind that the Littlewood-Paley theory relies on decomposing functions into blocks with compact support in the Fourier domain, we postulate to take band limited wavelets, e.g. Meyer wavelets.

Note that X then is not compactly supported, but exponentially concentrated on a fixed interval for an appropriate choice of ψ. We obtain the following sample path regularity of X:

Lemma 5.5.3. *If X is defined as in (5.38) and satisfies Assumption 3, then $X \in B^{\alpha}_{p,1}$ almost surely for any $\alpha < s + \frac{r}{p} - \frac{1}{2}$ and for $p \in \{2, 4\}$.*

Proof. Applying formally the Littlewood-Paley decomposition, one has

$$X = \sum_{j \geq -1} \Delta_j X$$

and for the sake of brevity we introduce the multi-indices $\lambda = (j, k)$ with $|\lambda| := j$. Noting that by the assumption on the wavelet basis $\operatorname{supp} \mathcal{F}\psi_\lambda \subset 2^{|\lambda|}\mathcal{A}$ for some annulus $\mathcal{A}$ independent of λ, we obtain $\Delta_j \psi_\lambda = 0$ if $|j - |\lambda||$ is larger than some fixed integer. Therefore, the Littlewood-Paley blocks are well-defined and given by

$$\Delta_j X = \sum_{\lambda : |\lambda| \sim j} Z_\lambda \Delta_j \psi_\lambda \quad \text{for} \quad j \geq -1.$$

Further, let us remark that X as given in (5.38) exists in $B^{\alpha}_{p,1}$ if $\sum_j \Delta_j X$ exists as limit in $B^{\alpha}_{p,1}$.

In order to show the claimed Besov regularity, we have to verify

$$\|\Delta_j X\|_{L^p} \lesssim 2^{-j(s+r/p-1/2)} \quad \text{for} \quad j \geq -1, \quad p \in \{2, 4\}.$$

Let us focus on $p = 2$. The case $p = 4$ can be proved similarly relying on the estimates for the forth moments of (Z_λ), see also Lemma 5.5.5 below. For $j \geq -1$ we have

$$\begin{aligned}
\mathbb{E}[\|\Delta_j X\|_{L^2}^2] &= \int_{\mathbb{R}} \mathbb{E}\Big[\Big(\sum_{\lambda} Z_\lambda \Delta_j \psi_{j,k}(t)\Big)^2\Big] \,\mathrm{d}t \\
&= \sum_{\lambda, \lambda'} \mathbb{E}[Z_\lambda Z_{\lambda'}] \int \Delta_j \psi_\lambda(t) \Delta_j \psi_{\lambda'}(t) \,\mathrm{d}t \lesssim \sum_{\lambda} 2^{-(2s+r)|\lambda|} \int (\Delta_j \psi_\lambda)^2(t) \,\mathrm{d}t,
\end{aligned}$$

where the last equality follows from (Z_λ) being mutually uncorrelated. Hence, we further estimate

$$\begin{aligned}
\mathbb{E}[\|\Delta_j X\|_{L^2}^2] &\lesssim \sum_{\lambda : |\lambda| \sim j} 2^{-(2s+r)|\lambda|} \|\Delta_j \psi_\lambda\|_{L^2}^2 \\
&\lesssim \sum_{j' \sim j} 2^{-(2s+r)j'} \sum_{k=-2^{j'}}^{2^{j'}} \|\psi_{j',k}\|_{L^2}^2 = 2 \sum_{j' \sim j} 2^{-2j'(s+r/2-1/2)}.
\end{aligned}$$

By the Littlewood-Paley characterization of the Besov norm we conclude

$$\mathbb{E}[\|X\|_{\alpha,p,1}] = \sum_{j \geq -1} 2^{j\alpha} \mathbb{E}[\|\Delta_j X\|_{L^p}] \lesssim \sum_{j \geq -1} 2^{j(\alpha - s - r/2 + 1/2)},$$

which is finite whenever $\alpha < s + r/2 - 1/2$. ☐

Remark 5.5.4. *With analogous estimates as in Lemma 5.5.3 it is easy to show that $X \in B^{\alpha}_{p,1}$ a.s. for any $\alpha < s + \frac{r}{p} - \frac{1}{2}$ for any even $p \geq 2$ provided $\mathbb{E}[A^p_{j,k}]^{1/p} \lesssim 2^{-s}$ still holds for these higher powers.*

The derivative of X is naturally given by $\mathrm{d}X_t = \sum_{j,k} Z_{j,k}\psi'_{j,k}(t)$ for $t \in \mathbb{R}$. The crucial point is now, that we can indeed verify that the resonant term $\pi(X, \mathrm{d}X)$ is in $B^{2\alpha-1}_{2,1}$ almost surely due to the probabilistic nature of X. The following lemma highlights how the stochastic setting nicely complements the analytical foundation.

Lemma 5.5.5. *Suppose X is given by (5.38) and satisfies Assumption 3, then*

$$X \in B^{\alpha}_{4,1} \quad \textit{and} \quad \pi(X, \mathrm{d}X) \in B^{2\alpha-1}_{2,1}$$

almost surely for any $\alpha < s + \frac{r}{4} - \frac{1}{2}$.

Proof. We start as in the classical proof of Bony's estimate (Lemma 5.1.1 (iii), cf. [BCD11, Thm. 2.85]), and decompose

$$\pi(X, \mathrm{d}X) = \sum_{j \geq -1} R_j \quad \text{with} \quad R_j := \sum_{|\nu| \leq 1} (\Delta_{j-\nu}X)(\Delta_j \,\mathrm{d}X).$$

By the properties of the Littlewood-Paley blocks the Fourier transform of R_j is supported in 2^j times some fixed ball. Consequently, $\Delta_{j'}R_j = 0$ if $j' \gtrsim j$ and thus

$$\big\|\Delta_{j'}\pi(X, \mathrm{d}X)\big\|_{L^2} = \Big\| \sum_{j \gtrsim j'} \Delta_{j'}R_j \Big\|_{L^2} \lesssim \sum_{j \gtrsim j'} \sum_{|\nu| \leq 1} \|(\Delta_{j-\nu}X)(\Delta_j \,\mathrm{d}X)\|_{L^2}.$$

Now we proceed similarly to Lemma 5.5.3 (using again the multi-indices $\lambda = (j,k)$):

$$\begin{aligned}
&\mathbb{E}\big[\|(\Delta_{j-\nu}X)(\Delta_j \,\mathrm{d}X)\|^2_{L^2}\big] \\
&\quad = \int_{\mathbb{R}} \mathbb{E}\Big[\Big(\sum_{\lambda_1,\lambda_2} Z_{\lambda_1}Z_{\lambda_2}(\Delta_{j-\nu}\psi_{\lambda_1})(\Delta_j\psi'_{\lambda_2})\Big)^2\Big]\,\mathrm{d}t \\
&\quad = \sum_{\substack{\lambda_1,\dots,\lambda_4: \\ |\lambda_\cdot| \sim j}} \mathbb{E}[Z_{\lambda_1}Z_{\lambda_2}Z_{\lambda_3}Z_{\lambda_3}] \int_{\mathbb{R}} (\Delta_{j-\nu}\psi_{\lambda_1})(\Delta_j\psi'_{\lambda_2})(\Delta_{j-\nu}\psi_{\lambda_3})(\Delta_j\psi'_{\lambda_4})\,\mathrm{d}t \\
&\quad \leq \sum_{\substack{\lambda_1 \neq \lambda_2: \\ |\lambda_\cdot| \sim j}} \mathbb{E}[Z^2_{\lambda_1}Z^2_{\lambda_2}] \int_{\mathbb{R}} \Big((\Delta_{j-\nu}\psi_{\lambda_1})^2(\Delta_j\psi'_{\lambda_2})^2 \\
&\qquad + (\Delta_{j-\nu}\psi_{\lambda_1})(\Delta_j\psi'_{\lambda_1})(\Delta_{j-\nu}\psi_{\lambda_2})(\Delta_j\psi'_{\lambda_2})\Big)\,\mathrm{d}t \\
&\qquad + \sum_{\lambda: |\lambda| \sim j} \mathbb{E}[Z^4_\lambda] \int_{\mathbb{R}} (\Delta_{j-\nu}\psi_\lambda)^2(\Delta_j\psi'_\lambda)^2\,\mathrm{d}t \\
&\quad \lesssim \sum_{\substack{\lambda_1 \neq \lambda_2: \\ |\lambda_\cdot| \sim j}} 2^{-(4s+2r)j}\|\psi_{\lambda_1}\|_{L^4}\|\psi'_{\lambda_2}\|_{L^4}\big(\|\psi_{\lambda_1}\|_{L^4}\|\psi'_{\lambda_2}\|_{L^4} + \|\psi_{\lambda_2}\|_{L^4}\|\psi'_{\lambda_1}\|_{L^4}\big) \\
&\qquad + \sum_{\lambda: |\lambda| \sim j} 2^{-(4s+r)j}\|\psi_\lambda\|^2_{L^4}\|\psi'_\lambda\|^2_{L^4}.
\end{aligned}$$

Plugging in $\psi_{j,k} = 2^{j/2}\psi(2^j \cdot -k)$, we obtain

$$\mathbb{E}\big[\|(\Delta_{j-\nu}X)(\Delta_j \,\mathrm{d}X)\|_{L^2}\big] \lesssim 2^{-j(2s+r/2-2)}.$$

The assertion follows from Lemma A.4.2 by the compact support of $\mathcal{F}R_j$ for $j \geq -1$. □

Combining the two previous lemmas, we conclude that stochastic models of the form (5.38) are prototypical examples of geometric Besov rough paths, which were introduced in Definition 5.4.1, and thus Theorem 5.4.8 can be applied to the corresponding stochastic differential equations.

Proposition 5.5.6. *Let φ satisfy Assumption 1 and $X = (X^1, \dots, X^n)$ be an n-dimensional stochastic process. Suppose each component X^d, $d = 1, \dots n$, is of the form (5.38), fulfills Assumption 3 for $\frac{5}{6} < s + \frac{r}{4}$ and the corresponding coefficients $(Z^d_{j,k})$ and $(Z^m_{j,k})$ are independent for $d \neq m$ and all j, k. Then, the localized process φX can be enhanced to a geometric Besov rough path, that is $\varphi X \in \mathcal{B}^{0,\alpha}_{4,1}$ almost surely for $\alpha \in (\frac{1}{3}, s + \frac{r}{4} - \frac{1}{2})$.*

Proof. The regularity for each component X^d, $d = 1, \dots, n$, is determined by Lemma 5.5.3 and thus $X \in B^{\alpha}_{4,1}$ for $\alpha \in (\frac{1}{3}, s+\frac{r}{4}-\frac{1}{2})$. Furthermore, a smooth approximation is given by the projection of X onto the first $J \geq 1$ Littlewood-Paley blocks as used in the proof of Lemma 5.5.3 or similarly by projecting on the first $J \geq 1$ wavelet resolution levels.

The resonant terms $\pi(X^d, \mathrm{d}X^d)$, $d = 1, \dots, n$, are constructed in Lemma 5.5.5 again by a smooth approximation in terms of Littlewood-Paley blocks. Due to the independence of the corresponding coefficients $(Z^d_{j,k})$ and $(Z^m_{j,k})$ for $d \neq m$, an analogous calculation shows that the resonant terms $\pi(X^d, \mathrm{d}X^m)$ for $d \neq m$ exists as limit of the same approximation in terms of Littlewood-Paley blocks, too.

It remains to deduce the above results for the localized process φX as well. The regularity and approximation of φX is implied by Lemma 5.1.2. For the resonant term $\pi(\varphi X, \mathrm{d}(\varphi X))$ we observe that

$$\pi(\varphi X, \mathrm{d}(\varphi X)) = \pi(\varphi X, \varphi' X) + \pi(\varphi X, \varphi \,\mathrm{d}X),$$

where the first term turns out to be no issue thanks to Lemma 5.1.1. For the second one we apply Bony's decomposition to φX and our commutator lemma (Lemma 5.3.4) to get

$$\begin{aligned}\pi(\varphi X, \varphi \,\mathrm{d}X) =& \varphi\pi(X, \varphi \,\mathrm{d}X) + \varphi\Gamma(\varphi, X, \varphi \,\mathrm{d}X) + \pi(\pi(\varphi, X), \varphi \,\mathrm{d}X) \\ &+ X\pi(\varphi, \varphi \,\mathrm{d}X) + \Gamma(X, \varphi, \varphi \,\mathrm{d}X).\end{aligned}$$

Due to the regularity of φ and X it remains to only handle the first term. By another analogous application of the commutator lemma, we finally see that the approximation of the resonant term of the localized process can be deduced from the above approximation of the non-localized process and therefore $\varphi X \in \mathcal{B}^{0,\alpha}_{4,1}$. □

6. An FBSDE approach to the Skorokhod embedding problem for Gaussian processes with non-linear drift

The Skorokhod embedding problem (SEP) stimulates research in probability theory now for over 50 years. The classical goal of the SEP consists in finding, for a given Brownian motion W and a probability measure ν, a stopping time τ such that W_τ possesses the law ν. It was first formulated and solved by Skorokhod [Sko61, Sko65] in 1961. Since then there appeared many different constructions for the stopping time τ and generalizations of the original problem in the literature. Just to name some of the most famous solutions to the SEP we refer to Root [Roo69], Rost [Ros71] and Azéma-Yor [AY79]. A comprehensive survey can be found in [Obł04].

Recently, the Skorokhod embedding raised additional interest because of some applications in financial mathematics, as for instance to obtain model-independent bounds on lookback options [Hob98] or on options on variance [CL10, CW13, OdR13]. An introduction to this close connection of the Skorokhod embedding problem and robust financial mathematics can be found in [Hob11].

In this chapter we construct a solution to the Skorokhod embedding problem for Gaussian process G of the form

$$G_t := G_0 + \int_0^t \alpha_s \,\mathrm{d}s + \int_0^t \beta_s \,\mathrm{d}W_s,$$

where $G_0 \in \mathbb{R}$ is a constant and $\alpha, \beta\colon [0,\infty) \to \mathbb{R}$ are suitable functions. Especially, this class of processes includes Brownian motions with non-linear drift. The SEP for Brownian motion with linear drift was first solved in the technical report [Hal68] and 30 years later again in [GF00] and [Pes00]. Techniques developed in these works can be extended to time-homogeneous diffusions, as done in [PP01], and can be seen as generalization of the Azéma-Yor solution. However, to the best of our knowledge there exists no solution so far for the case of a Brownian motion with non-linear drift.

The spirit of our approach is related to the one by Bass [Bas83], who employed martingale representation to find an alternative solution of the SEP for the Brownian motion. This approach was further developed for the Brownian motion with linear drift in [AHI08] and for time-homogeneous diffusion in [AHS15]. It rests upon the observation that the SEP may be viewed as the weak version of a stochastic control problem: the goal is to steer G in such a way that it takes the distribution of a prescribed law, which in case of zero drift is closely related to the martingale representation of a random variable with this law. We therefore propose in this chapter to formulate and solve the SEP for G in terms of a fully coupled Forward Backward Stochastic Differential Equation (FBSDE).

In general terms, the dynamics of a system of FBSDE is expressed by the equations

$$\begin{aligned} X_s &= X_0 + \int_0^s \mu(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_0^s \sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r, \\ Y_t &= \xi(X_T) - \int_t^T f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \quad t \in [0, T], \end{aligned}$$

with coefficient functions μ, σ of the forward part, terminal condition ξ and driver f of the backward component. In recent decades the theory of FBSDE with its close connection to quasi-linear partial differential equations and their viscosity solutions has been propagated extensively, in particular in its numerous areas of applications as stochastic control and mathematical finance (see [EPQ97] or [PW99]).

There are mainly three methods to show the existence of a solution for a system of FBSDE: the *contraction method* [Ant93, PT99], the *four step scheme* [MPY94] and the *method of continuation* [HP95, Yon97, PW99]. As a unified approach, [MWZZ15] (see also [Del02]) designed the theory of decoupling fields for FBSDE, which was significantly refined in [FI13]. It can primarily be seen as an extension of the contraction method. In our approach of the SEP via FBSDE, we shall focus on the subclass of Markovian ones for which all involved coefficient functions $(\xi, (\mu, \sigma, f))$ are deterministic. We, however, have to allow for not globally, but only locally Lipschitz continuous coefficients (μ, σ, f) in the control variable z, and therefore to develop an existence, uniqueness and regularity theory for FBSDE in this framework.

Equipped with these tools we solve the FBSDE system resulting from the SEP. We first construct a weak solution, i.e. we obtain a Gaussian process of the above form and an integrable random time such that, stopped at this time, the process possesses the given distribution ν. Under suitable regularity on the given measure ν and the process, this construction will be carried over to the originally given Gaussian process G. This solves the SEP for G.

The chapter is organized as follows: in Section 6.1 we relate the SEP to a fully coupled system of FBSDE, and in Section 6.2 we establish general results for decoupling fields of FBSDE. The Skorokhod embedding problem is solved in Section 6.3, in its weak and in its strong version. Section A.6 recalls some auxiliary results for BMO processes.

6.1. An FBSDE approach to the Skorokhod embedding problem

We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,\infty)}, \mathbb{P})$ large enough to carry a one-dimensional Brownian motion W. The filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$ is assumed to be generated by the Brownian motion and is assumed to be augmented by $\mathbb{P}$-null sets. We also assume that $\mathcal{F} = \sigma\left(\bigcup_{t=0}^{\infty} \mathcal{F}_t\right)$.

We start by formulating the Skorokhod embedding problem in the modified version **(SEP)**: For given probability measure ν on $\mathbb{R}$ and a Gaussian process X on $[0, \infty)$ of the form

$$X_t := X_0 + \int_0^t \alpha_s\,\mathrm{d}s + \int_0^t \beta_s\,\mathrm{d}W_s, \tag{6.1}$$

where $X_0 \in \mathbb{R}$ is some predetermined constant and $\alpha, \beta\colon [0,\infty) \to \mathbb{R}$ are deterministic measurable processes such that $\int_0^t |\alpha_s| \,\mathrm{d}s + \int_0^t \beta_s^2 \,\mathrm{d}s < \infty$ for all $t \geq 0$, find

- a $(\mathcal{F}_t)$-stopping time τ s.t. $\mathbb{E}[\tau] < \infty$ together with
- a starting point $c \in \mathbb{R}$

such that $c + X_\tau$ has the law ν.

In order to have a truly stochastic problem β should not vanish and ν should not be a Dirac measure. In fact we will assume that β is bounded away from zero later on.

Our method of solving this problem is based on the observation that it may be viewed as the weak version of a *stochastic control problem*: We want to steer X in such a way that it takes the distribution of a prescribed law. The spirit of our approach is related to an approach to the original Skorokhod embedding problem by Bass [Bas83] that was later extended to the Brownian motion with linear drift in [AHI08]. The procedure of both papers can be briefly summarized and divided into the following four steps.

(i) Construct a function $g\colon \mathbb{R} \to \mathbb{R}$ such that $g(W_1)$ has the given law ν.

(ii) Use the martingale representation property of the Brownian motion for $\alpha \equiv 0$ and $\beta \equiv 1$ or BSDE techniques for $\alpha \equiv \kappa \neq 0$ and $\beta \equiv 1$ to solve

$$Y_t = g(W_1) - \kappa \int_t^1 Z_s^2 \,\mathrm{d}s - \int_t^1 Z_s \,\mathrm{d}W_s, \quad t \in [0,1]. \tag{6.2}$$

(iii) Apply the random time-change of Dambis, Dubins and Schwarz in the quadratic variation scale $\int_0^\cdot Z_s^2 \,\mathrm{d}s$ to transform the martingale $\int_0^\cdot Z_s \,\mathrm{d}W_s$ into a Brownian motion B. This also provides a random time $\tilde{\tau} := \int_0^1 Z_s^2 \,\mathrm{d}s$ fulfilling $B_{\tilde{\tau}} + \kappa\tilde{\tau} + Y_0 = g(W_1)$, which is why $B_{\tilde{\tau}} + \kappa\tilde{\tau} + Y_0$ has the law ν.

(iv) Show that $\tilde{\tau}$ is a stopping time with respect to the filtration generated by B through an explicit characterization using the unique solution of an ordinary differential equation. With this description transform the embedding with respect to B into one with respect to the original Brownian motion W to obtain the stopping time τ as the analogue to $\tilde{\tau}$.

The **first step** of the algorithm just sketched is fairly easy. Let $F\colon \mathbb{R} \to [0,1]$ such that $F(x) := \nu((-\infty, x])$ is the cumulative distribution function associated with ν and define $F^{-1}\colon (0,1) \to \mathbb{R}$ via

$$F^{-1}(y) := \inf\{x \in \mathbb{R} \,:\, F(x) \geq y\}.$$

Denoting by Φ the distribution function of the standard normal distribution, we define $g\colon \mathbb{R} \to \mathbb{R}$ as $g(x) := F^{-1}(\Phi(x))$. It is straightforward to prove that g has the following properties.

Lemma 6.1.1. *The function g is measurable and non-decreasing. Moreover, if ν is not a Dirac measure, then g is not identically constant and $g(W_1)$ has the law ν.*

Proof. Since Φ and F^{-1} are measurable and non-decreasing, their composition g is also measurable and non-decreasing.

Clearly, g can only be constant if F^{-1} is constant, which can only happen if F assumes values in $\{0,1\}$. This only happens in case ν is a Dirac measure. In order to see that $g(W_1)$ has the law ν, note that

$$\mathbb{P}(g(W_1) \leq x) = \mathbb{P}(F^{-1}(\Phi(W_1)) \leq x) = \mathbb{P}(W_1 \leq \Phi^{-1}(F(x))) = \Phi(\Phi^{-1}(F(x))) = F(x)$$

for all $x \in \mathbb{R}$. □

Since we want to require as little regularity as possible for the processes involved, we use the concept of weak differentiability. We recall that a measurable $f\colon \Omega \times \mathbb{R}^n \to \mathbb{R}$ is *weakly differentiable* if there exists a mapping $\frac{\mathrm{d}}{\mathrm{d}\lambda} f\colon \Omega \times \mathbb{R}^n \to \mathbb{R}^{1\times n}$ such that

$$\int_{\mathbb{R}^n} \varphi(\lambda) \frac{\mathrm{d}}{\mathrm{d}\lambda} f(\omega,\lambda)\,\mathrm{d}\lambda = -\int_{\mathbb{R}^n} f(\omega,\lambda) \frac{\mathrm{d}}{\mathrm{d}\lambda}\varphi(\lambda)\,\mathrm{d}\lambda,$$

for any smooth test function $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ with compact support, for almost all $\omega \in \Omega$.

Now define a measurable function $\hat{\delta}\colon [0,\infty) \to \mathbb{R}$ via

$$\hat{\delta}(t) := X_0 + \int_0^t \alpha_s \,\mathrm{d}s$$

such that $X_t = \hat{\delta}(t) + \int_0^t \beta_s \,\mathrm{d}W_s$. Obviously, $\hat{\delta}$ is weakly differentiable. Conversely, for every weakly differentiable function $\hat{\delta}\colon [0,\infty) \to \mathbb{R}$ we can set $X_0 := \hat{\delta}(0)$ and $\alpha_s := \hat{\delta}'(s)$.

Furthermore, define $H\colon [0,\infty) \to [0,\infty)$ via

$$H(t) := \int_0^t \beta_s^2 \,\mathrm{d}s.$$

Note that H is weakly differentiable, monotonically increasing and starts at 0. If we assume that β is bounded away from 0, H becomes strictly increasing and invertible such that the inverse function H^{-1} is monotonically increasing and Lipschitz continuous. In this case we can define

$$\delta := \hat{\delta} \circ H^{-1}.$$

If $\beta \equiv 1$, then $H = \mathrm{Id}$ and thus $\delta = \hat{\delta}$.

For the **second step** we assume that β is bounded away from 0 and observe that the random time change, which turns the martingale $\int_0^\cdot Z_s \,\mathrm{d}W_s$ into a Gaussian process of the form $\int_0^\cdot \beta_s \,\mathrm{d}B_s$ simultaneously turns the scale process $\int_0^\cdot Z_s^2 \,\mathrm{d}s$ into $\int_0^\cdot \beta_s^2 \,\mathrm{d}s = H$. This means we have to modify the classical martingale representation of $g(W_1)$ to

$$g(W_1) + \hat{\delta}\Big(H^{-1}\Big(\int_0^1 Z_s^2 \,\mathrm{d}s\Big)\Big) - \mathbb{E}\Big[g(W_1) + \hat{\delta}\Big(H^{-1}\Big(\int_0^1 Z_s^2 \,\mathrm{d}s\Big)\Big)\Big] = \int_0^1 Z_s \,\mathrm{d}W_s,$$

which amounts to finding a solution (Y, Z) to the equation

$$Y_t = g(W_1) - \delta\Big(\int_0^1 Z_s^2 \,\mathrm{d}s\Big) - \int_t^1 Z_s \,\mathrm{d}W_s, \quad t \in [0,1]. \tag{6.3}$$

For $\delta(t) \equiv 0$ this would be just the usual martingale representation with respect to the Brownian motion. Also for a linear drift $\delta(t) = \kappa t$ and $\beta \equiv 1$ equation (6.3) can be rewritten as

$$\tilde{Y}_t := Y_t + \kappa \int_0^t Z_s^2 \,\mathrm{d}s = g(W_1) - \kappa \int_t^1 Z_s^2 \,\mathrm{d}s - \int_t^1 Z_s \,\mathrm{d}W_s, \quad t \in [0,1],$$

which is exactly the BSDE (6.2) related to the SEP as stated in [AHI08]. In the case of a Brownian motion with general drift equation (6.3) would be a BSDE with time-delayed terminal condition. Unfortunately, the theory of BSDE with time-delay as introduced by Delong and Imkeller in [DI10] and extended by Delong [Del12] for time-delayed terminal conditions reaches its limits in our situation. Alternatively, we will understand equation (6.3) as an FBSDE and develop new techniques to solve it. This will be done in Sections 6.2 and 6.3. Before we tackle the solvability of equation (6.3), we show that it really leads to the desired result in the **third step** of our algorithm. To be mathematically rigorous we introduce

- $\mathbb{S}^2(\mathbb{R})$ as the space of all progressively measurable processes $Y\colon \Omega \times [0,1] \to \mathbb{R}$ satisfying
  $\sup_{t\in[0,1]} \mathbb{E}[|Y_t|^2] < \infty$,
- $\mathbb{H}^2(\mathbb{R})$ as the space of all progressively measurable processes $Z\colon \Omega \times [0,1] \to \mathbb{R}$ satisfying
  $\mathbb{E}[\int_0^1 |Z_t|^2 \,\mathrm{d}t] < \infty$,

where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}$.

For the rest of the chapter we assume that β is **bounded away from** 0, i.e. $\inf_{s\in[0,\infty)} |\beta_s| > 0$.

Lemma 6.1.2. *Suppose* $(Y, Z) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R})$ *is a solution of* (6.3)*. Then there exist a Brownian motion B and a random time $\tilde{\tau}$ with $\mathbb{E}[\tilde{\tau}] < \infty$ such that*

$$Y_0 + X_0 + \int_0^{\tilde{\tau}} \alpha_s \,\mathrm{d}s + \int_0^{\tilde{\tau}} \beta_s \,\mathrm{d}B_s = g(W_1).$$

Proof. Note that Y is a martingale with quadratic variation process $\int_0^t Z_s^2 \,\mathrm{d}s$ for $t \in [0,1]$ since $Z \in \mathbb{H}^2(\mathbb{R})$. Now choose another Brownian motion $\tilde{B}$ which is independent of Y. If necessary we extend our probability space such that it accommodates the Brownian motion $\tilde{B}$. Set $\tilde{\tau} := H^{-1}\left(\int_0^1 Z_s^2 \,\mathrm{d}s\right)$, and define the time-change of the type of Dambis, Dubins and Schwarz by

$$\sigma_r := \begin{cases} \inf\left\{t \geq 0 \,:\, \int_0^t Z_s^2 \,\mathrm{d}s > \int_0^r \beta_s^2 \,\mathrm{d}s\right\} & \text{if } 0 \leq r < \tilde{\tau} \\ 1 & \text{if } r \geq \tilde{\tau}. \end{cases}$$

Observe that the condition $r < \tilde{\tau}$ is equivalent to $\int_0^r \beta_s^2 \,\mathrm{d}s < \int_0^1 Z_s^2 \,\mathrm{d}s$. Since Y_{σ_r} is a continuous martingale with quadratic variation $H(r) = \int_0^r \beta_s^2 \,\mathrm{d}s$, we can define a Brownian motion B by

$$B_r := \tilde{B}_r - \tilde{B}_{r\wedge\tilde{\tau}} + \int_0^{r\wedge\tilde{\tau}} \frac{1}{\beta_s} \,\mathrm{d}Y_{\sigma_s}, \quad 0 \le r < \infty.$$

We find

$$\int_0^{\tilde{\tau}} \beta_s \,\mathrm{d}B_s + \hat{\delta}(\tilde{\tau}) + Y_0 = Y_1 - Y_0 + \delta\Big(\int_0^1 Z_s^2 \,\mathrm{d}s\Big) + Y_0 = g(W_1),$$

and further

$$\mathbb{E}[\tilde{\tau}] = \mathbb{E}\Big[H^{-1}\Big(\int_0^1 Z_s^2 \,\mathrm{d}s\Big)\Big] < \infty,$$

where we used that $Z \in \mathbb{H}^2(\mathbb{R})$ and H^{-1} is Lipschitz continuous. $\square$

As an immediate consequence of the previous lemma we observe the following fact. If we have a solution $(Y, Z) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R})$ of equation (6.3), we obtain a *weak solution* to the Skorokhod embedding problem, i.e. a Gaussian process of the form (6.1), a starting point c, and an integrable random time such that our process stopped at this time possesses a given distribution.

At a first glance equation (6.3) might look easy. We, however, have to deal with a fully coupled FBSDE which in addition possesses a not globally Lipschitz continuous coefficient in the forward component.

6.2. Decoupling fields for fully coupled FBSDEs

The theory of FBSDE, closely connected to the theory of quasi-linear partial differential equations and their viscosity solutions, receives its general interest from numerous areas of application among which stochastic control and mathematical finance are the most vivid ones in recent decades (see [EPQ97] or [PW99]). Owing to their general significance, we treat the theory of FBSDEs and their decoupling fields in a more general framework than might be needed to obtain a solution to our equation (6.3).

Although in Section 6.2.2 we will focus on the Markovian case, which means that all involved coefficients are purely deterministic, let us dwell in a more general setting first.

6.2.1. General decoupling fields

For a fixed time horizon $T > 0$, we consider a complete filtered probability space

$$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}),$$

where $\mathcal{F}_0$ contains all null sets, $(W_t)_{t\in[0,T]}$ is a d-dimensional Brownian motion independent of $\mathcal{F}_0$, and $\mathcal{F}_t := \sigma(\mathcal{F}_0, (W_s)_{s\in[0,t]})$ with $\mathcal{F} := \mathcal{F}_T$. The dynamics of an

FBSDE is classically given by

$$
\begin{aligned}
X_s &= X_0 + \int_0^s \mu(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_0^s \sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r, \\
Y_t &= \xi(X_T) - \int_t^T f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r,
\end{aligned}
$$

for $s, t \in [0, T]$ and $X_0 \in \mathbb{R}^n$, where $(\xi, (\mu, \sigma, f))$ are measurable functions. More precisely,

$$
\begin{aligned}
&\xi\colon \Omega \times \mathbb{R}^n \to \mathbb{R}^m, && \mu\colon [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^n, \\
&\sigma\colon [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^{n\times d}, && f\colon [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m,
\end{aligned}
$$

for $n, m, d \in \mathbb{N}$. Throughout the whole section μ, σ and f are assumed to be progressively measurable with respect to $(\mathcal{F}_t)_{t\in[0,T]}$, i.e. $\mu\mathbf{1}_{[0,t]}, \sigma\mathbf{1}_{[0,t]}, f\mathbf{1}_{[0,t]}$ are $\mathcal{B}([0,T]) \otimes \mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^{m\times d})$-measurable for all $t \in [0, T]$.

A decoupling field comes with an even richer structure than just a classical solution.

Definition 6.2.1. Let $t \in [0, T]$. A function $u\colon [t, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ with $u(T, \cdot) = \xi$ a.e. is called *decoupling field* for $(\xi, (\mu, \sigma, f))$ on $[t, T]$ if for all $t_1, t_2 \in [t, T]$ with $t_1 \le t_2$ and any $\mathcal{F}_{t_1}$-measurable $X_{t_1}\colon \Omega \to \mathbb{R}^n$ there exist progressive processes (X, Y, Z) on $[t_1, t_2]$ such that

$$
\begin{aligned}
X_s &= X_{t_1} + \int_{t_1}^s \mu(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_{t_1}^s \sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r, \\
Y_s &= Y_{t_2} - \int_s^{t_2} f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_s^{t_2} Z_r\,\mathrm{d}W_r, \\
Y_s &= u(s, X_s), \qquad (6.4)
\end{aligned}
$$

for all $s \in [t_1, t_2]$. In particular, we want all integrals to be well-defined and (X, Y, Z) to have values in $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^{m\times d}$, respectively.

Some remarks about this definition are in place.

- The first equation in (6.4) is called the *forward equation*, the second the *backward equation* and the third will be referred to as the *decoupling condition*.
- The requirement that X should start at X_{t_1} is referred to as the *initial condition*. By a slight abuse of notation we will sometimes refer to X_{t_1} itself as the initial condition.
- Note that, if $t_2 = T$, we get $Y_T = \xi(X_T)$ a.s. as a consequence of the decoupling condition together with $u(T, \cdot) = \xi$. At the same time $Y_T = \xi(X_T)$ together with the decoupling condition implies $u(T, \cdot) = \xi$ a.e.
- If $t_2 = T$ we can say that a triplet (X, Y, Z) solves the FBSDE, meaning that it satisfies the forward and the backward equation, together with $Y_T = \xi(X_T)$. This relationship $Y_T = \xi(X_T)$ is referred to as the *terminal condition*.

  By an abuse of notation the function ξ itself is also sometimes referred to as the terminal condition. Sometimes we will describe the relationship $u(T, \cdot) = \xi$ a.e. with this term.

In contrast to classical solutions of FBSDE, decoupling fields on different intervals can be pasted together.

Lemma 6.2.2 (Lemma 1 in [FI13]). *Let u be a decoupling field for $(\xi, (\mu, \sigma, f))$ on $[t, T]$ and $\tilde{u}$ be a decoupling field for $(u(t, \cdot), (\mu, \sigma, f))$ on $[s, t]$, for $0 \leq s < t < T$. Then, the map $\hat{u}$ given by $\hat{u} := \tilde{u}\mathbf{1}_{[s,t]} + u\mathbf{1}_{(t,T]}$ is a decoupling field for $(\xi, (\mu, \sigma, f))$ on $[s, T]$.*

We want to remark that, if u is a decoupling field and $\tilde{u}$ is a modification of u, i.e. for each $s \in [t, T]$ the functions $u(s, \omega, \cdot)$ and $\tilde{u}(s, \omega, \cdot)$ coincide for almost all $\omega \in \Omega$, then $\tilde{u}$ is also a decoupling field to the same problem. So u could also be referred to as a class of modifications. Some of the representatives of the class might be progressively measurable, others not. As we see below a progressively measurable representative does exist if the decoupling field is Lipschitz continuous in x:

Lemma 6.2.3 (Lemma 2 in [FI13]). *Let $u\colon [t, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ be a decoupling field to $(\xi, (\mu, \sigma, f))$ which is Lipschitz continuous in $x \in \mathbb{R}^n$ in the sense that there exists a constant $L > 0$ s.t. for every $s \in [t, T]$:*

$$|u(s, \omega, x) - u(s, \omega, x')| \leq L|x - x'| \qquad \forall x, x' \in \mathbb{R}^n, \quad \text{for a.a. } \omega \in \Omega.$$

Then u has a modification $\tilde{u}$ which is progressively measurable and Lipschitz continuous in x in the strong sense

$$|\tilde{u}(s, \omega, x) - \tilde{u}(s, \omega, x')| \leq L|x - x'| \qquad \forall s \in [t, T],\, \omega \in \Omega,\, x, x' \in \mathbb{R}^n.$$

Let $I \subseteq [0, T]$ be an interval and $u : I \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ a map such that $u(s, \cdot)$ is measurable for every $s \in I$. We define

$$\begin{aligned} L_{u,x} := \sup_{s \in I} \inf\{L \geq 0 \,|\, &\text{for a.a. } \omega \in \Omega : |u(s, \omega, x) - u(s, \omega, x')| \leq L|x - x'| \\ &\text{for all } x, x' \in \mathbb{R}^n\}, \end{aligned} \tag{6.5}$$

where $\inf \emptyset := \infty$. We also set $L_{u,x} := \infty$ if $u(s, \cdot)$ is not measurable for every $s \in I$. One can show that $L_{u,x} < \infty$ is equivalent to u having a modification which is truly Lipschitz continuous in $x \in \mathbb{R}^n$.

We denote by $L_{\sigma,z}$ the Lipschitz constant of σ w.r.t. the dependence on the last component z and w.r.t. the Frobenius norms on $\mathbb{R}^{m \times d}$ and $\mathbb{R}^{n \times d}$. We set $L_{\sigma,z} = \infty$ if σ is not Lipschitz continuous in z.

By $L_{\sigma,z}^{-1} = \frac{1}{L_{\sigma,z}}$ we mean $\frac{1}{L_{\sigma,z}}$ if $L_{\sigma,z} > 0$ and ∞ otherwise.

Definition 6.2.4. Let $u\colon [t, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ be a decoupling field to $(\xi, (\mu, \sigma, f))$. We say u to be *weakly regular* if $L_{u,x} < L_{\sigma,z}^{-1}$ and $\sup_{s \in [t,T]} \|u(s, \cdot, 0)\|_\infty < \infty$.

This is a natural definition due to Lemma 6.2.3. In practice, however, it is important to have explicit knowledge about the regularity of (X, Y, Z). For instance, it is important to know in which spaces the processes live, and how they react to changes in the initial value. Specifically, it can be very useful to have differentiability of (X, Y, Z) w.r.t. the initial value.

In the following we need further notation. For an integrable real valued random variable F the expression $\mathbb{E}_t[F]$ refers to $\mathbb{E}[F|\mathcal{F}_t]$, while $\mathbb{E}_{\hat{t},\infty}[F]$ refers to $\operatorname{ess\,sup} \mathbb{E}[F|\mathcal{F}_t]$, which might be ∞, but is always well defined as the infimum of all constants $c \in [-\infty, \infty]$ such that $\mathbb{E}[F|\mathcal{F}_t] \leq c$ a.s. Additionally, we write $\|F\|_\infty$ for the essential supremum of $|F|$.

Definition 6.2.5. Let $u\colon [t,T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ be a weakly regular decoupling field to $(\xi, (\mu, \sigma, f))$. We call u *strongly regular* if for all fixed $t_1, t_2 \in [t,T]$, $t_1 \leq t_2$, the processes (X, Y, Z) arising in (6.4) are a.e unique and satisfy

$$\sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}[|X_s|^2] + \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}[|Y_s|^2] + \mathbb{E}_{t_1,\infty}\left[\int_{t_1}^{t_2} |Z_s|^2 \,\mathrm{d}s\right] < \infty, \tag{6.6}$$

for each constant initial value $X_{t_1} = x \in \mathbb{R}^n$. In addition they must be measurable as functions of (x, s, ω) and even weakly differentiable w.r.t. $x \in \mathbb{R}^n$ such that for every $s \in [t_1, t_2]$ the mappings X_s and Y_s are measurable functions of (x, ω) and even weakly differentiable w.r.t. x such that

$$\begin{aligned}
&\operatorname{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}\left[\left|\frac{\mathrm{d}}{\mathrm{d}x} X_s\right|_v^2\right] < \infty,\\
&\operatorname{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}\left[\left|\frac{\mathrm{d}}{\mathrm{d}x} Y_s\right|_v^2\right] < \infty,\\
&\operatorname{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \mathbb{E}_{t_1,\infty}\left[\int_{t_1}^{t_2} \left|\frac{\mathrm{d}}{\mathrm{d}x} Z_s\right|_v^2 \mathrm{d}s\right] < \infty.
\end{aligned} \tag{6.7}$$

We say that a decoupling field on $[t,T]$ is *strongly regular* on a subinterval $[t_1, t_2] \subseteq [t,T]$ if u restricted to $[t_1, t_2]$ is a strongly regular decoupling field for $(u(t_2, \cdot), (\mu, \sigma, f))$.

Under certain conditions a rich existence, uniqueness and regularity theory for decoupling fields can be developed. We will summarize the main results, which are proven in [FI13]:

Assumption (SLC): $(\xi, (\mu, \sigma, f))$ satisfies *standard Lipschitz conditions* (SLC) if

(i) (μ, σ, f) are Lipschitz continuous in (x, y, z) with Lipschitz constant L,

(ii) $\|(|\mu| + |f| + |\sigma|)(\cdot, \cdot, 0, 0, 0)\|_\infty < \infty$,

(iii) $\xi\colon \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ is measurable such that $\|\xi(\cdot, 0)\|_\infty < \infty$ and $L_{\xi,x} < L_{\sigma,z}^{-1}$.

Theorem 6.2.6 (Theorem 1 in [FI13])**.** *Suppose* $(\xi, (\mu, \sigma, f))$ *satisfies* (SLC)*. Then there exists a time* $t \in [0,T)$ *such that* $(\xi, (\mu, \sigma, f))$ *has a unique (up to modification) decoupling field* u *on* $[t,T]$ *with* $L_{u,x} < L_{\sigma,z}^{-1}$ *and* $\sup_{s\in[t,T]} \|u(s, \cdot, 0)\|_\infty < \infty$.

A brief discussion of existence and uniqueness of classical solutions can be found in Remark 3 in [FI13]. For later reference we give the following remarks (cf. Remark 1 and 2 in [FI13]).

Remark 6.2.7. *It can be observed from the proof that the supremum of all* $h = T - t$*, with* t *satisfying the properties required in Theorem 6.2.6 can be bounded away from 0 by a bound, which only depends on*

- *the Lipschitz constant of* (μ, σ, f) *w.r.t. the last* 3 *components,* T*,*
- L_ξ *and* $L_\xi \cdot L_{\sigma,z} < 1$*,*

and which is monotonically decreasing in these values.

Remark 6.2.8. *It can be observed from the proof that our decoupling field* u *on* $[t,T]$ *satisfies* $L_{u(s,\cdot),x} \leq L_{\xi,x} + C(T-s)^{\frac{1}{4}}$*, where* C *is some constant which does not depend on* $s \in [t,T]$*. More precisely,* C *depends only on* T*,* L*,* $L_{\xi,x}$*,* $L_{\xi,x}L_{\sigma,z}$ *and is monotonically increasing in these values.*

We can systematically extend this local theory to obtain global results. This is based on a simple argument which we will refer to as *small interval induction.*

Lemma 6.2.9 (Lemma 11 and 12 in [FI13])**.** *Let* $T_1 < T_2$ *be real numbers and let* $S \subseteq [T_1, T_2]$*.*

(i) Forward: *If* $T_1 \in S$ *and there exists an* $h > 0$ *s.t.* $[s, s+h] \cap [T_1, T_2] \subseteq S$ *for all* $s \in S$*, then* $S = [T_1, T_2]$ *and in particular* $T_2 \in S$*.*

(ii) Backward: *If* $T_2 \in S$ *and there exists an* $h > 0$ *s.t.* $[s-h, s] \cap [T_1, T_2] \subseteq S$ *for all* $s \in S$*, then* $S = [T_1, T_2]$ *and in particular* $T_1 \in S$*.*

Using these simple results we obtain *global uniqueness* and *global regularity* of a decoupling field.

Theorem 6.2.10 (Corollary 1 and 2 in [FI13])**.** *Suppose that* $(\xi, (\mu, \sigma, f))$ *satisfies* (SLC)*.*

(i) Global uniqueness: *If there are two weakly regular decoupling fields* $u^{(1)}, u^{(2)}$ *to the corresponding problem on some interval* $[t,T]$*, then we have* $u^{(1)} = u^{(2)}$ *up to modifications.*

(ii) Global regularity: *If there exists a weakly regular decoupling field* u *to this problem on some interval* $[t,T]$*, then* u *is strongly regular.*

Notice that Theorem 6.2.10 only provides uniqueness of weakly regular decoupling fields, not uniqueness of processes (X,Y,Z) solving the FBSDE in the classical sense. However, using global regularity in Theorem 6.2.10 one can show:

Corollary 6.2.11 (Corollary 3 in [FI13])**.** *Let* $(\xi, (\mu, \sigma, f))$ *fulfill* (SLC)*. If there exists a weakly regular decoupling field* u *of the corresponding FBSDE on some interval* $[t,T]$*, then for any initial condition* $X_t = x \in \mathbb{R}^n$ *there is a unique solution* (X,Y,Z) *of the FBSDE on* $[t,T]$ *satisfying*

$$\sup_{s\in[t,T]} \mathbb{E}[|X_s|^2] + \sup_{s\in[t,T]} \mathbb{E}[|Y_s|^2] + \mathbb{E}\left[\int_t^T |Z_s|^2 \,\mathrm{d}s\right] < \infty.$$

6.2.2. Markovian decoupling fields

A system of FBSDE given by $(\xi, (\mu, \sigma, f))$ is said to be *Markovian* if these four coefficient functions are deterministic, that is, if they depend only on (t, x, y, z). In the Markovian situation we can somewhat relax the Lipschitz continuity assumption and still obtain local existence together with uniqueness. What makes the Markovian case so special is the property

$$"Z_s = u_x(s, X_s) \cdot \sigma(s, X_s, Y_s, Z_s)",$$

which comes from the fact that u will also be deterministic. This property allows us to bound Z by a constant if we assume that σ is bounded.

Lemma 6.2.12 (Lemma 14 in [FI13])**.** *Let μ, σ, f, ξ satisfy* (SLC) *and assume in addition that they are deterministic. Assume that we have a weakly regular decoupling field u on an interval $[t, T]$. Then u is deterministic in the sense that it has a modification which is a function of $(r, x) \in [t, T] \times \mathbb{R}^n$ only.*

An application of Lemma 6.2.12 is the following very fundamental result.

Lemma 6.2.13 (Lemma 15 in [FI13])**.** *Let $(\xi, (\mu, \sigma, f))$ satisfy* (SLC) *and suppose that these coefficient functions are deterministic. Let u be a weakly regular decoupling field on an interval $[t, T]$. Choose $t_1 < t_2$ from $[t, T]$ and an initial condition X_{t_1}. Then the corresponding Z satisfies $\|Z\|_\infty \leq L_{u,x} \cdot \|\sigma\|_\infty$.*

If $\|Z\|_\infty < \infty$, we also have $\|Z\|_\infty \leq L_{u,x} \|\sigma(\cdot, \cdot, \cdot, 0)\|_\infty (1 - L_{u,x} L_{\sigma,z})^{-1}$.

Next we investigate the continuity of u as a function of time and space.

Lemma 6.2.14 (Lemma 16 in [FI13])**.** *Assume that (μ, σ, f) have linear growth in (x, y) in the sense*

$$(|\mu| + |\sigma| + |f|)\,(t, \omega, x, y, z) \leq C\,(1 + |x| + |y|) \quad \forall (t, x, y, z) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d},$$

for a.a. $\omega \in \Omega$, where $C \in [0, \infty)$ is some constant.

If u is a strongly regular and deterministic decoupling field to $(\xi, (\mu, \sigma, f))$ on an interval $[t, T]$, then u is continuous in the sense that it has a modification which is a continuous function on $[t, T] \times \mathbb{R}^n$.

This boundedness of Z in the Markovian case motivates the following definition. It will allow us to develop a theory for non-Lipschitz problems via truncation.

Definition 6.2.15. Let $t \in [0, T]$. We call a function $u \colon [t, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ with $u(T, \omega, \cdot) = \xi(\omega, \cdot)$ for a.a. $\omega \in \Omega$ a *Markovian decoupling field* for $(\xi, (\mu, \sigma, f))$ on $[t, T]$ if for all $t_1, t_2 \in [t, T]$ with $t_1 \leq t_2$ and any $\mathcal{F}_{t_1}$-measurable $X_{t_1} \colon \Omega \to \mathbb{R}^n$ there exist progressive processes (X, Y, Z) on $[t_1, t_2]$ such that the equations in (6.4) hold a.s. for all $s \in [t_1, t_2]$, and additionally $\|Z\|_\infty < \infty$.

We remark that a Markovian decoupling field is always a decoupling field in the standard sense as well. The only difference is that we are only interested in triplets (X, Y, Z), where Z is a.e. bounded.

Regularity for Markovian decoupling fields is defined very similarly to standard regularity.

Definition 6.2.16. Let $u\colon [t,T]\times\Omega\times\mathbb{R}^n\to\mathbb{R}^m$ be a Markovian decoupling field to $(\xi,(\mu,\sigma,f))$.

- We call u *weakly regular* if $L_{u,x} < L_{\sigma,z}^{-1}$ and $\sup_{s\in[t,T]}\|u(s,\cdot,0)\|_\infty < \infty$.
- We call a weakly regular u *strongly regular* if for all fixed $t_1,t_2\in[t,T]$, $t_1\le t_2$, the processes (X,Y,Z) arising in the defining property of a Markovian decoupling field are a.e. unique for each *constant* initial value $X_{t_1}=x\in\mathbb{R}^n$ and satisfy (6.6). In addition they must be measurable as functions of (x,s,ω) and even weakly differentiable w.r.t. $x\in\mathbb{R}^n$ such that for every $s\in[t_1,t_2]$ the mappings X_s and Y_s are measurable functions of (x,ω), and even weakly differentiable w.r.t. x such that (6.7) holds.
- We say that a Markovian decoupling field on $[t,T]$ is strongly regular on a subinterval $[t_1,t_2]\subseteq[t,T]$ if u restricted to $[t_1,t_2]$ is a strongly regular Markovian decoupling field for $(u(t_2,\cdot),(\mu,\sigma,f))$.

Now we define a class of problems for which an existence and uniqueness theory will be developed.

Assumption (MLLC):
$(\xi,(\mu,\sigma,f))$ fulfills a *modified local Lipschitz condition* (MLLC) if

(i) the functions (μ,σ,f) are

   a) deterministic,

   b) Lipschitz continuous in (x,y,z) on sets of the form $[0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times B$, where $B\subset\mathbb{R}^{m\times d}$ is an arbitrary bounded set,

   c) and fulfill $\|\mu(\cdot,0,0,0)\|_\infty, \|f(\cdot,0,0,0)\|_\infty, \|\sigma(\cdot,\cdot,\cdot,0)\|_\infty, L_{\sigma,z}<\infty$,

(ii) $\xi\colon\mathbb{R}^n\to\mathbb{R}^m$ satisfies $L_{\xi,x}<L_{\sigma,z}^{-1}$.

We start a providing a local existence result.

Theorem 6.2.17. *Let $(\xi,(\mu,\sigma,f))$ satisfy* (MLLC). *Then there exists a time $t\in[0,T)$ such that $(\xi,(\mu,\sigma,f))$ has a unique weakly regular Markovian decoupling field u on $[t,T]$. This u is also strongly regular, deterministic, continuous and satisfies $\sup_{t_1,t_2,X_{t_1}}\|Z\|_\infty<\infty$, where $t_1<t_2$ are from $[t,T]$ and X_{t_1} is an initial value (see the definition of a Markovian decoupling field for the meaning of these variables).*

Proof. For any constant $H>0$ let $\chi_H\colon\mathbb{R}^{m\times d}\to\mathbb{R}^{m\times d}$ be defined as

$$\chi_H(z) := \mathbf{1}_{\{|z|<H\}}z + \frac{H}{|z|}\mathbf{1}_{\{|z|\ge H\}}z.$$

It is easy to check that χ_H is Lipschitz continuous with Lipschitz constant $L_{\chi_H}=1$ and bounded by H. Furthermore, we have $\chi_H(z)=z$ if $|z|\le H$. We implement an "inner cutoff" by defining (μ_H,σ_H,f_H) via $\mu_H(t,x,y,z):=\mu(t,x,y,\chi_H(z))$, etc.

The boundedness of χ_H together with its Lipschitz continuity makes (μ_H,σ_H,f_H) Lipschitz continuous with some Lipschitz constant L_H. Furthermore, $L_{\sigma_H,z}\le L_{\sigma,z}$.

Also (μ_H, σ_H, f_H) have linear growth in (y, z) as required by Lemma 6.2.14. According to Theorem 6.2.6 we know that the problem given by $(\xi, (\mu_H, \sigma_H, f_H))$ has a unique weakly regular decoupling field u on some small interval $[t', T]$ where $t' \in [0, T)$. We also know that this u is strongly regular, u is deterministic (by Lemma 6.2.12), and continuous (by Lemma 6.2.14).

We will show that for sufficiently large H and $t \in [t', T)$ it will also be a Markovian decoupling field to the problem $(\xi, (\mu, \sigma, f))$. Using Remark 6.2.8

$$L_{u(t,\cdot),x} \leq L_{\xi,x} + C_H(T-t)^{\frac{1}{4}} \quad \forall t \in [t', T],$$

where $C_H < \infty$ is a constant which does not depend on $t \in [t', T]$. For any $t_1 \in [t', T]$ and $\mathcal{F}_{t_1}$-measurable initial value X_{t_1} consider the corresponding unique X, Y, Z on $[t_1, T]$ satisfying the forward equation, the backward equation and the decoupling condition for μ_H, σ_H, f_H and u. Using Lemma 6.2.13 we have $\|Z\|_\infty \leq L_{u,x}\|\sigma_H\|_\infty \leq L_{u,x}\left(\|\sigma(\cdot,\cdot,\cdot,0)\|_\infty + L_{\sigma,z}H\right) < \infty$ and, therefore,

$$\begin{aligned}\|Z\|_\infty &\leq \frac{\sup_{s\in[t_1,T]} L_{u(s,\cdot),x} \cdot \|\sigma(\cdot,\cdot,\cdot,0)\|_\infty}{1 - \sup_{s\in[t_1,T]} L_{u(s,\cdot),x} L_{\sigma,z}} \leq \frac{\left(L_{\xi,x} + C_H(T-t_1)^{\frac{1}{4}}\right) \cdot \|\sigma(\cdot,\cdot,\cdot,0)\|_\infty}{1 - L_{\xi,x}L_{\sigma,z} - L_{\sigma,z}C_H(T-t_1)^{\frac{1}{4}}} \\ &= \frac{L_{\xi,x}\|\sigma(\cdot,\cdot,\cdot,0)\|_\infty}{1 - L_{\xi,x}L_{\sigma,z} - L_{\sigma,z}C_H(T-t_1)^{\frac{1}{4}}} + \frac{C_H(T-t_1)^{\frac{1}{4}} \cdot \|\sigma(\cdot,\cdot,\cdot,0)\|_\infty}{1 - L_{\xi,x}L_{\sigma,z} - L_{\sigma,z}C_H(T-t_1)^{\frac{1}{4}}} \qquad (6.8)\end{aligned}$$

for $T - t_1$ small enough.

Now we only need to

- choose H large enough such that $\frac{L_{\xi,x}\|\sigma(\cdot,\cdot,\cdot,0)\|_\infty}{1-L_{\xi,x}L_{\sigma,z}}$ becomes smaller than $\frac{H}{4}$,
- and then in the second step choose t close enough to T, such that
  - $L_{\sigma,z}C_H(T-t)^{\frac{1}{4}}$ becomes smaller than $\frac{1}{2}(1 - L_{\xi,x}L_{\sigma,z})$,
  - $\frac{C_H\|\sigma(\cdot,\cdot,\cdot,0)\|_\infty(T-t)^{\frac{1}{4}}}{1-L_{\xi,x}L_{\sigma,z}}$ becomes smaller than $\frac{H}{4}$.

Considering (6.8) this implies that if $t_1 \in [t, T]$ the process Z a.e. does not leave the region in which the cutoff is "passive", i.e. the ball of radius H. Therefore, u restricted to the interval $[t, T]$ is a decoupling field to $(\xi, (\mu, \sigma, f))$, not just to $(\xi, (\mu_H, \sigma_H, f_H))$. It is even a Markovian decoupling field due to the boundedness of Z. As a Markovian decoupling field it is weakly regular, because it is weakly regular as a decoupling field to $(\xi, (\mu_H, \sigma_H, f_H))$.

Uniqueness: Assume than there is another weakly regular Markovian decoupling field $\tilde{u}$ to $(\xi, (\mu, \sigma, f))$ on $[t, T]$. Choose a $t_1 \in [t, T]$ and an $x \in \mathbb{R}^n$ as initial condition $X_{t_1} = x$, and consider the corresponding processes $(\tilde{X}, \tilde{Y}, \tilde{Z})$ that satisfy the corresponding FBSDE on $[t_1, T]$, together with the decoupling condition via $\tilde{u}$. At the same time consider (X, Y, Z) solving the same FBSDE on $[t_1, T]$, but associated with the Markovian decoupling field u. Since $\tilde{Z}, Z$ are bounded, the two triplets $(\tilde{X}, \tilde{Y}, \tilde{Z})$ and (X, Y, Z) also solve the Lipschitz FBSDE given by $(\xi, (\mu_H, \sigma_H, f_H))$ on $[t_1, T]$ for H large enough. The two conditions $\tilde{Y}_s = \tilde{u}(s, \tilde{X}_s)$ and $Y_s = u(s, X_s)$ imply by Remark 3 in [FI13] that both triplets are progressively measurable processes on $[t_1, T] \times \Omega$ s.t.

$$\sup_{s\in[t_1,T]} \mathbb{E}_{0,\infty}\left[|X_s|^2\right] + \sup_{s\in[t_1,T]} \mathbb{E}_{0,\infty}\left[|Y_s|^2|\right] + \mathbb{E}_{0,\infty}\left[\int_{t_1}^{T} |Z_s|^2 \, \mathrm{d}s\right] < \infty$$

and coincide. In particular, $\tilde{u}(t_1, x) = \tilde{Y}_{t_1} = Y_{t_1} = u(t_1, x)$.

Strong regularity of u as a Markovian decoupling field to $(\xi, (\mu, \sigma, f))$ follows directly from the above argument about uniqueness of (X, Y, Z) for deterministic initial values and bounded Z, and the strong regularity of u as decoupling field to $(\xi, (\mu_H, \sigma_H, f_H))$. □

Remark 6.2.18. *We observe from the proof that the supremum of all $h = T - t$ with t satisfying the hypotheses of Theorem 6.2.17 can be bounded away from* 0 *by a bound, which only depends on*

- $L_{\xi,x}$, $L_{\xi,x} \cdot L_{\sigma,z}$,
- $\|\sigma(\cdot, \cdot, \cdot, 0)\|_\infty$, T, $L_{\sigma,z}$,
- *the values $(L_H)_{H\in[0,\infty)}$ where L_H is the Lipschitz constant of (μ, σ, f) on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times B_H$ w.r.t. to the last* 3 *components, where $B_H \subset \mathbb{R}^{m\times d}$ denotes the ball of radius H with center* 0*,*

and which is monotonically decreasing in these values.

The following natural concept introduces a type of Markovian decoupling fields for non-Lipschitz problems (non-Lipschitz in z), to which nevertheless standard Lipschitz results can be applied.

Definition 6.2.19. Let u be a Markovian decoupling field for $(\xi, (\mu, \sigma, f))$.

- We call u *controlled in z* if there exists a constant $C > 0$ such that for all $t_1, t_2 \in [t, T]$, $t_1 \le t_2$, and all initial values X_{t_1}, the corresponding processes (X, Y, Z) from the definition of a Markovian decoupling field satisfy $|Z_s(\omega)| \le C$, for almost all $(s, \omega) \in [t, T] \times \Omega$. If for a fixed triplet (t_1, t_2, X_{t_1}) there are different choices for (X, Y, Z), then all of them are supposed to satisfy the above control.
- We say that a Markovian decoupling field on $[t, T]$ is *controlled in z* on a subinterval $[t_1, t_2] \subseteq [t, T]$ if u restricted to $[t_1, t_2]$ is a Markovian decoupling field for $(u(t_2, \cdot), (\mu, \sigma, f))$ that is controlled in z.
- A Markovian decoupling field u on an interval $(s, T]$ is said to be *controlled in z* if it is controlled in z on every compact subinterval $[t, T] \subseteq (s, T]$ with C possibly depending on t.

Remark 6.2.20. *Our Markovian decoupling field from Theorem 6.2.17 is obviously controlled in z: consider* (6.8) *together with the choice of $t \le t_1$ made in the proof.*

Remark 6.2.21. *Let* $(\xi, (\mu, \sigma, f))$ *satisfy* (MLLC)*, and assume that we have a Markovian decoupling field* u *on some interval* $[t, T]$*, which is weakly regular and controlled in* z*. Then* u *is also a solution to a Lipschitz problem obtained through a cutoff as in Theorem 6.2.17. As such it is strongly regular (Theorem 6.2.10) and* <u>*deterministic*</u> *(Lemma 6.2.12). But now Lemma 6.2.14 is also applicable, since due to the use of a cutoff we can assume the type of linear growth required there. So* u *is also* <u>*continuous*</u>*.*

Lemma 6.2.22. *Let* $(\xi, (\mu, \sigma, f))$ *satisfy* (MLLC)*. For* $0 \leq s < t < T$ *let* u *be a weakly regular Markovian decoupling field for* $(\xi, (\mu, \sigma, f))$ *on* $[s, T]$*. If* u *is controlled in* z *on* $[s, t]$ *and* $T - t$ *is small enough as required in Theorem 6.2.17 resp. Remark 6.2.18, then* u *is controlled in* z *on* $[s, T]$*.*

Proof. Clearly, u is not just controlled in z on $[s, t]$, but also on $[t, T]$ (with a possibly different constant), according to Remark 6.2.20. Define C as the maximum of these two constants.

We only need to control Z by C for the case $s \leq t_1 \leq t \leq t_2 \leq T$, the other two cases being trivial. For this purpose consider the processes (X, Y, Z) on the interval $[t_1, t_2]$ corresponding to some initial value X_{t_1} and fulfilling the forward equation, the backward equation and the decoupling condition. Since the restrictions of these processes to $[t_1, t]$ still fulfill these three properties we obtain $|Z_r(\omega)| \leq C$ for almost all $r \in [t_1, t]$, $\omega \in \Omega$.

At the same time, if we restrict (X, Y, Z) to $[t, t_2]$, we observe that these restrictions satisfy the forward equation, the backward equation and the decoupling condition for the interval $[t, t_2]$ with X_t as initial value. Therefore $|Z_r(\omega)| \leq C$ holds for a.a. $r \in [t, t_2]$, $\omega \in \Omega$ as well. $\square$

The following important result allows us to connect the (MLLC)-case to (SLC).

Theorem 6.2.23. *Let* $(\xi, (\mu, \sigma, f))$ *be such that* (MLLC) *is satisfied and assume that there exists a weakly regular Markovian decoupling field* u *to this problem on some interval* $[t, T]$*. Then* u *is controlled in* z*.*

Proof. Let $S \subseteq [t, T]$ be the set of all times $s \in [t, T]$, s.t. u is controlled in z on $[t, s]$.

- Clearly $t \in S$: For the interval $[t, t] = \{t\}$ one can only choose $t_1 = t_2 = t$ and so $Z \colon [t, t] \times \Omega \to \mathbb{R}^{m \times d}$ is $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-a.e. 0, independently of the initial value X_{t_1}. So we can take for C any positive value.

- Let $s \in S$ be arbitrary. According to Lemma 6.2.22 there exists an $h > 0$ s.t. u is controlled in z on $[t, (s + h) \wedge T]$ since $\|u((s + h) \wedge T, \cdot)\|_\infty < \infty$ and $L_{u((s+h)\wedge T, \cdot)} < L_{\sigma,z}^{-1}$. Considering Remark 6.2.18 and the requirements $\|u\|_\infty < \infty$, $L_{u,x} < L_{\sigma,z}^{-1}$, we can choose h independently of s.

This shows $S = [t, T]$ by small interval induction. $\square$

Note that Theorem 6.2.23 implies together with Remark 6.2.21 that a weakly regular Markovian decoupling field to an (MLLC) problem is deterministic and continuous.

Such a u will be a standard decoupling field to an (SLC) problem if we truncate μ, σ, f appropriately. We can thereby extend the whole theory to (MLLC) problems:

Theorem 6.2.24. *Let* $(\xi, (\mu, \sigma, f))$ *satisfy* (MLLC)*.*

(i) Global uniqueness: *If there are two weakly regular Markovian decoupling fields* $u^{(1)}, u^{(2)}$ *to this problem on some interval* $[t, T]$*, then* $u^{(1)} = u^{(2)}$*.*

(ii) Global regularity: *If that there exists a weakly regular Markovian decoupling field* u *to this problem on some interval* $[t, T]$*, then* u *is strongly regular.*

Proof. 1. We know that $u^{(1)}$ and $u^{(2)}$ are controlled in z. Choose a passive cutoff (see proof of Theorem 6.2.17) and apply 1. of Theorem 6.2.10.

2. u is controlled in z. Choose a passive cutoff (see proof of Theorem 6.2.17) and apply 2. of Theorem 6.2.10. □

Lemma 6.2.25. *Let* $(\xi, \mu, \sigma, f))$ *satisfy* (MLLC) *and assume that there exists a weakly regular Markovian decoupling field* u *of the corresponding FBSDE on some interval* $[t, T]$*.*

Then for any initial condition $X_t = x \in \mathbb{R}^n$ *there is a unique solution* (X, Y, Z) *of the FBSDE on* $[t, T]$ *such that*

$$\sup_{s \in [t,T]} \mathbb{E}[|X_s|^2] + \sup_{s \in [t,T]} \mathbb{E}[|Y_s|^2] + \|Z\|_\infty < \infty.$$

Proof. Existence follows from the fact that u is also strongly regular according to 2. of Theorem 6.2.24 and controlled in z according to Theorem 6.2.23.

Uniqueness follows from Corollary 6.2.11: Assume there are two solutions (X, Y, Z) and $(\tilde{X}, \tilde{Y}, \tilde{Z})$ to the FBSDE on $[t, T]$ both satisfying the aforementioned bound. But then they both solve an (SLC)-conform FBSDE obtained through a passive cutoff. So they must coincide according to Corollary 6.2.11. □

Definition 6.2.26. Let $I^M_{\max} \subseteq [0, T]$ for $(\xi, (\mu, \sigma, f))$ be the union of all intervals $[t, T] \subseteq [0, T]$ such that there exists a weakly regular Markovian decoupling field u on $[t, T]$.

Unfortunately, the maximal interval might very well be open to the left. Therefore, we need to make our notions more precise in the following definitions.

Definition 6.2.27. Let $0 \le t < T$.

- We call a function $u\colon (t, T] \times \mathbb{R}^n \to \mathbb{R}^m$ a Markovian decoupling field for $(\xi, (\mu, \sigma, f))$ on $(t, T]$ if u restricted to $[t', T]$ is a Markovian decoupling field for all $t' \in (t, T]$.
- A Markovian decoupling field u on $(t, T]$ is said to be *weakly regular* if u restricted to $[t', T]$ is a weakly regular Markovian decoupling field for all $t' \in (t, T]$.
- A Markovian decoupling field u on $(t, T]$ is said to be *strongly regular* if u restricted to $[t', T]$ is strongly regular for all $t' \in (t, T]$.

Theorem 6.2.28 (Global existence in weak form)**.** *Let* $(\xi, (\mu, \sigma, f))$ *satisfy* (MLLC)*. Then there exists a unique weakly regular Markovian decoupling field* u *on* $I^M_{\max}$*. This* u *is also deterministic, controlled in* z *and strongly regular.*

Moreover, either $I^M_{\max} = [0, T]$ *or* $I^M_{\max} = (t^M_{\min}, T]$*, where* $0 \le t^M_{\min} < T$*.*

Proof. Let $t \in I^M_{\max}$. Obviously, there exists a Markovian decoupling field $\check{u}^{(t)}$ on $[t,T]$ satisfying $L_{\check{u}^{(t)},x} < L^{-1}_{\sigma,z}$ and $\sup_{s\in[t,T]} \|\check{u}^{(t)}(s,\cdot,0)\|_\infty < \infty$. $\check{u}^{(t)}$ is controlled in z and strongly regular due to Theorems 6.2.23 and 6.2.24. We can further assume w.l.o.g. that $\check{u}^{(t)}$ is a continuous function on $[t,T]\times\mathbb{R}^n$ according to Remark 6.2.21. There is only one such $\check{u}^{(t)}$ according to Theorem 6.2.24. Furthermore, for $t,t' \in I^M_{\max}$ the functions $\check{u}^{(t)}$ and $\check{u}^{(t')}$ coincide on $[t\vee t',T]$ because of Theorem 6.2.24.

Define $u(t,\cdot) := \check{u}^{(t)}(t,\cdot)$ for all $t \in I^M_{\max}$. This function u is a Markovian decoupling field on $[t,T]$, since it coincides with $\check{u}^{(t)}$ on $[t,T]$. Therefore, u is a Markovian decoupling field on the whole interval $I^M_{\max}$ and satisfies $L_{u|_{[t,T]},x} < L^{-1}_{\sigma,z}$, $\sup_{s\in[t,T]} \|u|_{[t,T]}(s,\cdot,0)\|_\infty < \infty$ for all $t \in I^M_{\max}$.

Uniqueness of u follows directly from Theorem 6.2.24 applied to every interval $[t,T] \subseteq I^M_{\max}$.

Addressing the form of $I^M_{\max}$, we see that $I^M_{\max} = [t,T]$ with $t \in (0,T]$ is not possible: Assume otherwise. According to the above there exists a Markovian decoupling field u on $[t,T]$ s.t. $L_{u,x} < L^{-1}_{\sigma,z}$ and $\sup_{s\in[t,T]} \|u(s,\cdot,0)\|_\infty < \infty$. But then u can be extended a little bit to the left using Theorem 6.2.17 and Lemma 6.2.2, thereby contradicting the definition of $I^M_{\max}$. □

The following result basically states that for a singularity $t^M_{\min}$ to occur u_x has to "explode" at $t^M_{\min}$.

Lemma 6.2.29. *Let* $(\xi,(\mu,\sigma,f))$ *satisfy* (MLLC). *If* $I^M_{\max} = (t^M_{\min},T]$, *then*

$$\lim_{t\downarrow t^M_{\min}} L_{u(t,\cdot),x} = L^{-1}_{\sigma,z},$$

where u is the Markovian decoupling field according to Theorem 6.2.28.

Proof. We argue indirectly. Assume otherwise. Then we can select times $t_n \downarrow t^M_{\min}$, $n\to\infty$ such that

$$\sup_{n\in\mathbb{N}} L_{u(t_n,\cdot),x} < L^{-1}_{\sigma,z}.$$

But then we may choose an $h>0$ according to Remark 6.2.18 which does not depend on n and then choose n large enough to have $t_n - t^M_{\min} < h$. So u can be extended to the left to a larger interval $[(t_n-h)\vee 0,T]$ contradicting the definition of $I^M_{\max}$. □

6.3. Solution to the Skorokhod embedding problem

In this section we present a solution to the Skorokhod embedding problem as stated in (SEP) at the beginning of Section 6.1 based on solutions of the associated system of FBSDE.

6.3.1. Weak solution

Let us therefore return to our FBSDE (6.3) that can be rewritten slightly more generally as

$$
\begin{aligned}
X_s^{(1)} &= x^{(1)} + \int_t^s 1 \,\mathrm{d}W_r, & X_s^{(2)} &= x^{(2)} + \int_t^s Z_r^2 \,\mathrm{d}r, \\
Y_s &= g(X_T^{(1)}) - \delta(X_T^{(2)}) - \int_s^T Z_r \,\mathrm{d}W_r, & u(s, X_s^{(1)}, X_s^{(2)}) &= Y_s,
\end{aligned}
\tag{6.9}
$$

for $s \in [t,T]$ and $x = \left(x^{(1)}, x^{(2)}\right)^\top \in \mathbb{R}^2$. So using the notations of Section 6.2 we have

$$
\begin{aligned}
\mu(t,\omega,x,y,z) &= (0, z^2)^\top, & \sigma(t,\omega,x,y,z) &= (1,0)^\top, \\
f(t,\omega,x,y,z) &= 0, & \xi(\omega,x) &= g(x^{(1)}) - \delta(x^{(2)}),
\end{aligned}
$$

for all $(t,\omega,x,y,z) \in [0,T] \times \Omega \times \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}$ and $d = 1$, $n = 2$ and $m = 1$. In particular, the problem satisfies (MLLC).

Notice that by choosing $x := (x^{(1)}, x^{(2)})^\top := (0,0)^\top$ and $T = 1$ we will have $X_1^{(1)} = W_1$ and $X_1^{(2)} = \int_0^1 Z_s^2 \,\mathrm{d}s$, which makes the FBSDE equivalent to (6.3).

With the general results of Section 6.2.2 at hand we are capable to solve this system of equations. In other words, we are able to perform the **second step** of our algorithm to solve the SEP.

Lemma 6.3.1. *Assume δ and g are Lipschitz continuous. Then for the FBSDE* (6.9) *there exists a unique weakly regular Markovian decoupling field u on $[0,T]$. This u is strongly regular, controlled in z, deterministic and continuous.*

In particular, equation (6.3) *has a unique solution (Y,Z) such that $\|Z\|_\infty < \infty$.*

Proof. Using Theorem 6.2.28 we know that there exists a unique weakly regular Markovian decoupling field u on $I_{\max}^M$. This u is furthermore strongly regular, controlled in z, deterministic and continuous. It remains to prove $I_{\max}^M = [0,T]$. Due to Lemma 6.2.29 it is sufficient to show the existence of a constant $C \in [t,\infty]$ such that $L_{u(t,\cdot),x} \le C < L_{\sigma,z}^{-1}$ for all $t \in I_{\max}^M$. In our case $L_{\sigma,z}^{-1} = \infty$, so we have to prove that the weak partial derivatives of u with respect to $x^{(1)}$ and $x^{(2)}$ are both uniformly bounded.

Fix $t \in I_{\max}^M$ and consider the corresponding FBSDE on $[t,T]$: First notice that the associated triplet (X,Y,Z) depends on the initial value $x = (x^{(1)}, x^{(2)})^\top \in \mathbb{R}^2$, even in a weakly differentiable way with respect to the initial value x, according to the strong regularity of u. For more on rules regarding working with weak derivatives consult Section 2.2 of [FI13].

Let us first look at the matrix $\frac{\mathrm{d}}{\mathrm{d}x}X$. We have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_s^{(1)} &= 1, & \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_s^{(2)} &= \int_t^s 2Z_r \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Z_r \,\mathrm{d}r, \\
\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(1)} &= 0, & \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)} &= 1 + \int_t^s 2Z_r \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Z_r \,\mathrm{d}r,
\end{aligned}
$$

a.s. for $s \in [t, T]$, for almost all $x = (x^{(1)}, x^{(2)})^\top \in \mathbb{R}^2$. In particular, the 2×2-matrix $\frac{\mathrm{d}}{\mathrm{d}x} X_s$ is invertible if and only if $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}$ ist not 0. We will see later that it remains positive on the whole interval allowing us to apply the chain rule of Lemma A.6.8 in order to write $\frac{\mathrm{d}}{\mathrm{d}x} u(s, X_s) \frac{\mathrm{d}}{\mathrm{d}x} X_s$. But let us first proceed by differentiating the backward equation in (6.9) with respect to $x^{(2)}$:

$$\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_s = -\delta'(X_T^{(2)}) \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_T^{(2)} - \int_s^T \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Z_r \,\mathrm{d}W_r.$$

To be precise the above holds a.s. for every $s \in [t, T]$, for almost all $x = (x^{(1)}, x^{(2)})^\top \in \mathbb{R}^2$.

Now define a stopping time τ via

$$\tau := \inf \left\{ s \in [t, T] \,:\, \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)} \leq 0 \right\} \wedge T.$$

For $s \in [t, \tau)$ we have $\frac{\mathrm{d}}{\mathrm{d}x} u(s, X_s) \frac{\mathrm{d}}{\mathrm{d}x} X_s$ according to the chain rule of Lemma A.6.8 and in particular $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u(s, X_s^{(1)}, X_s^{(2)}) \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)} = \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_s$. Let us set

$$V_s := \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u(s, X_s^{(1)}, X_s^{(2)}),\ s \in [t, T] \quad \text{and} \quad \tilde{Z}_r := \frac{\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Z_r}{\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_r^{(2)}} \mathbf{1}_{\{r \in [t, \tau)\}}.$$

Then the dynamics of $\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}\right)^{-1}$ can be expressed by

$$\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_{s\wedge\tilde{\tau}}^{(2)}\right)^{-1} = 1 - \int_t^{s\wedge\tilde{\tau}} 2 Z_r \tilde{Z}_r \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_r^{(2)}\right)^{-1} \mathrm{d}r, \tag{6.10}$$

for an arbitrary stopping time $\tilde{\tau} < \tau$ with values in $[t, T]$. We also have $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_s = V_s \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}$ and therefore

$$V_s = \frac{\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_s}{\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}}, \quad s \in [t, \tau).$$

Applying Itô's formula and using the dynamics of $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y$ and $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}$ we easily obtain an equation describing the dynamics of $V_{s\wedge\tilde{\tau}}$:

$$\begin{aligned}
V_{s\wedge\tilde{\tau}} &= V_t + \int_t^{s\wedge\tilde{\tau}} -2 Z_r \tilde{Z}_r \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_r^{(2)}\right)^{-1} \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_r \,\mathrm{d}r \\
&\quad + \int_t^{s\wedge\tilde{\tau}} \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Z_r \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_r^{(2)}\right)^{-1} \mathrm{d}W_r \\
&= V_t + \int_t^{s\wedge\tilde{\tau}} (-2 Z_r V_r) \tilde{Z}_r \,\mathrm{d}r + \int_t^{s\wedge\tilde{\tau}} \tilde{Z}_r \,\mathrm{d}W_r
\end{aligned} \tag{6.11}$$

for any stopping time $\tilde{\tau} < \tau$ with values in $[t, T]$.

Note that, since V and $(-2ZV)$ are bounded processes, $\tilde{Z}\mathbf{1}_{[\cdot \leq \tilde{\tau}]}$ is in $BMO(\mathbb{P})$ according to Theorem A.6.7 with a $BMO(\mathbb{P})$-norm which does not depend on $\tilde{\tau} < \tau$, and so in particular $\mathbb{E}[\int_t^\tau |2Z_r\tilde{Z}_r|^2 \,\mathrm{d}r] < \infty$. From (6.10) we can actually deduce that $\tau = T$ must hold almost surely. Indeed, (6.10) implies that

$$\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}_{s\wedge\tilde{\tau}}\right)^{-1} = \exp\left(-\int_t^{s\wedge\tilde{\tau}} 2Z_r\tilde{Z}_r \,\mathrm{d}r\right)$$

or equivalently

$$\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}_{s\wedge\tilde{\tau}} = \exp\left(\int_t^{s\wedge\tilde{\tau}} 2Z_r\tilde{Z}_r \,\mathrm{d}r\right)$$

for all stopping times $\tilde{\tau} < \tau$ with values in $[t, T]$. Using continuity of $s \mapsto \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}_s$ we obtain

$$\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}_\tau = \exp\left(\int_t^\tau 2Z_r\tilde{Z}_r \,\mathrm{d}r\right) > 0,$$

which gives us $\tau = T$ a.s. because $\{\tau < T\} \subset \left\{\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}_\tau = 0\right\}$, due to continuity of $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}$.

So we have $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}$ is positive on the whole $[t, T]$ and therefore $\frac{\mathrm{d}}{\mathrm{d}x} X$ is invertible on $[t, T]$.

Setting $\tilde{W}_s := W_s - \int_t^s 2Z_rV_r \,\mathrm{d}r$, $s \in [t, T]$ we can reformulate (6.11) to

$$V_s = V_t + \int_t^s \tilde{Z}_r \,\mathrm{d}\tilde{W}_r.$$

This means that V_s can be viewed as the conditional expectation of

$$V_T = \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u(T, X^{(1)}_T, X^{(2)}_T) = -\delta'(X^{(2)}_T)$$

with respect to $\mathcal{F}_s$ and some probability measure, which turns $\tilde{W}$ into a Brownian motion on $[t, T]$. Note here that $2Z_rV_r$ is bounded on $[t, T]$ because $||Z||_\infty < \infty$. Hence, we conclude that V_t and therefore $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u(t, x^{(1)}, x^{(2)})$ is bounded by $\|\delta'\|_\infty$ for almost all $x = (x^{(1)}, x^{(2)})^\top \in \mathbb{R}^2$. This value is independent of t.

Secondly, we have to bound $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u(t, x^{(1)}, x^{(2)})$. To this end we differentiate the equations in (6.9) with respect to $x^{(1)}$:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X^{(1)}_s = 1, \\
&\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X^{(2)}_s = \int_t^s 2Z_r \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Z_r \,\mathrm{d}r, \\
&\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y_s = g'(X^{(1)}_T) - \delta'(X^{(2)}_T) \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X^{(2)}_T - \int_s^T \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Z_r \,\mathrm{d}W_r, \\
&\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u(s, X^{(1)}_s, X^{(2)}_s) + \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u(s, X^{(1)}_s, X^{(2)}_s) \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X^{(2)}_s = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y_s,
\end{aligned}$$

and define

$$U_s := \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u(s, X^{(1)}_s, X^{(2)}_s), \quad \check{Z}_r := \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Z_r - \tilde{Z}_r \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X^{(2)}_r.$$

Note that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_s^{(2)} &= \int_t^s 2Z_r\left(\check{Z}_r + \tilde{Z}_r\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_r^{(2)}\right)\mathrm{d}r,\\ U_s &= \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}Y_s - V_s\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_s^{(2)},\end{aligned}$$

which allows us to deduce the dynamics of U from the dynamics of $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}Y$, $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X^{(2)}$ and V using Itô formula:

$$\begin{aligned}U_s =&U_t + \int_t^s 1\,\mathrm{d}\Big(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}Y_r\Big) - \int_t^s V_r\,\mathrm{d}\Big(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_r^{(2)}\Big) - \int_t^s \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_r^{(2)}\,\mathrm{d}V_r\\ =&U_t + \int_t^s \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}Z_r\,\mathrm{d}W_r - 2\int_t^s V_rZ_r\left(\check{Z}_r + \tilde{Z}_r\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_r^{(2)}\right)\mathrm{d}r\\ &- \int_t^s \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_r^{(2)}\left(-2Z_rV_r\tilde{Z}_r\,\mathrm{d}r + \tilde{Z}_r\,\mathrm{d}W_r\right)\end{aligned}$$

where the marked terms either merge into one or cancel out and the equation simplifies to

$$\begin{aligned}U_s =&U_t + \int_t^s(-2Z_rV_r\check{Z}_r)\,\mathrm{d}r + \int_t^s \check{Z}_r\,\mathrm{d}W_r. &(6.12)\\ =&U_t + \int_t^s \check{Z}_r\,\mathrm{d}\tilde{W}_r.\end{aligned}$$

By the same argument as for the process V we deduce that U and therefore

$$\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(t,x^{(1)},x^{(2)})$$

is bounded by $\|g'\|_\infty = L_g$ for almost all $x^{(1)}, x^{(2)}$, where L_g is the Lipschitz constant of g, i.e. the infimum of all Lipschitz constants.

This shows that $I^M_{\max} = [0,T]$.

Finally, Lemma 6.2.25 shows that there is a unique solution (X,Y,Z) to the FBSDE on $[0,T]$ for any initial value $(X_0^{(1)},X_0^{(2)})^\top = (x^{(1)},x^{(2)})^\top \in \mathbb{R}^2$ such that

$$\sup_{s\in[0,T]}\mathbb{E}[|X_s|^2] + \sup_{s\in[0,T]}\mathbb{E}[|Y_s|^2] + \|Z\|_\infty < \infty,$$

which is equivalent to the simpler condition $\|Z\|_\infty < \infty$ as we claim:

If $\|Z\|_\infty < \infty$, then according to the forward equation

$$\begin{aligned}&\|X^{(2)}\|_\infty \le |x^{(2)}| + T\|Z\|_\infty^2 < \infty,\\ &\sup_{s\in[0,T]}\mathbb{E}[|X_s|^2] = |x^{(1)}|^2 + \sup_{s\in[0,T]}\mathbb{E}[|W_s|^2] = |x^{(1)}|^2 + T < \infty,\end{aligned}$$

and according to the backward equation together with the Minkowski inequality

$$\begin{aligned}\Big(\sup_{s\in[0,T]}\mathbb{E}[|Y_s|^2]\Big)^{\frac12} &= \left(\sup_{s\in[0,T]}\mathbb{E}\left[\left|\mathbb{E}\left[g(X_T^{(1)}) - \delta(X_T^{(2)})\middle|\mathcal{F}_s\right]\right|^2\right]\right)^{\frac12}\\ &\le \left(\mathbb{E}\left[\left|g(X_T^{(1)}) - \delta(X_T^{(2)})\right|^2\right]\right)^{\frac12} \le \left(\mathbb{E}\left[\left|g(X_T^{(1)})\right|^2\right]\right)^{\frac12} + \left(\mathbb{E}\left[\left|\delta(X_T^{(2)})\right|^2\right]\right)^{\frac12}\\ &\le |g(0)| + L_g\left(\mathbb{E}\left[\left|X_T^{(1)}\right|^2\right]\right)^{\frac12} + |\delta(0)| + L_\delta\left(\mathbb{E}\left[\left|X_T^{(2)}\right|^2\right]\right)^{\frac12} < \infty,\end{aligned}$$

where L_g, L_δ are Lipschitz constants of g, δ. $\square$

For the following result we use the notations of Section 6.1. As before we assume that β is bounded away from 0. Under this condition H^{-1} is well defined and Lipschitz continuous. Therefore, $\delta = \hat{\delta} \circ H^{-1}$ is Lipschitz continuous if $\hat{\delta}$ is Lipschitz continuous, which is equivalent to α being bounded.

Lemma 6.3.2. *Suppose g and δ are both Lipschitz continuous with Lipschitz constants L_g and L_δ. Then there exist a Brownian motion B, a random time $\tilde{\tau} \leq H^{-1}(L_g^2)$ and a constant $c \in \mathbb{R}$ such that $c + \int_0^{\tilde{\tau}} \alpha_s \, \mathrm{d}s + \int_0^{\tilde{\tau}} \beta_s \, \mathrm{d}B_s$ has law ν.*

Proof. First we solve FBSDE (6.3) using Lemma 6.3.1 such that the corresponding Z is bounded. According to Lemma 6.3.5, which we prove a bit later, we can even assume that Z is bounded by L_g. Now we set $c := Y_0$ and construct B and $\tilde{\tau}$ as in the proof of Lemma 6.1.2.

Moreover, $\tilde{\tau} = H^{-1}\left(\int_0^1 Z_s^2 \, \mathrm{d}s\right)$ is bounded by $H^{-1}(L_g^2)$ since Z is bounded by L_g and H^{-1} is increasing. $\square$

Remark 6.3.3. *It is a priori not clear that the random time $\tilde{\tau}$ is also a stopping time with respect to*

$$\left(\mathcal{F}_s^B\right)_{s\in[0,\infty)} := \left(\sigma\left(B_r, r \in [0,s]\right)_{s\in[0,\infty)}\right)$$

as also mentioned in Remark 1.2 in [AHI08]. However, we will prove a sufficient criterion for this in terms of regularity properties of the Markovian decoupling field u.

Remark 6.3.4. *The boundedness of the stopping time solving the Skorokhod embedding problem has not been investigated so frequently. However, very recently it gained attention in [AS11] and [AHS15]. Especially, its economic interest comes from its applications in the context of game theory (see [SS09]).*

6.3.2. Strong solution

This subsection is devoted to the **fourth step** of our algorithm, i.e. to translate the results of the preceding section into a solution of the Skorokhod embedding problem in the strong sense. Our main goal is to show that if g, δ are sufficiently smooth, then $\tilde{\tau}$ and B constructed so far will have the property that $\tilde{\tau}$ is indeed a stopping time w.r.t. filtration $\left(\mathcal{F}_s^B\right)_{s\in[0,\infty)}$ generated by the Brownian motion B, and thus a functional of the trajectories of B. The same functional applied to the trajectories of the original Brownian motion W will then provide the required strong solution. For this purpose, we will assume that g and δ are three times weakly differentiable with bounded derivatives. We will also require that g is non-decreasing and not constant. Our arguments shall be based on a deep analysis of regularity properties of the associated decoupling field u.

First let us first prove the following very useful result about the solution (Y, Z) to FBSDE (6.3) constructed in Lemma 6.3.1.

Lemma 6.3.5. *Assume δ and g are Lipschitz continuous. Let u be the unique weakly regular Markovian decoupling field associated to the problem (6.9) on $[0,T]$ constructed in Lemma 6.3.1. Then for any $t \in [0,T)$ and initial condition $(X_t^{(1)}, X_t^{(2)})^\top = (x^{(1)}, x^{(2)})^\top \in \mathbb{R}^2$ the associated process Z on $[t,T]$ satisfies $\|Z\|_\infty \leq L_g = \|g'\|_\infty$.*

Furthermore, if the weak derivative $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ has a version which is continuous in the first two components $(s, x^{(1)})$ on $[t,T) \times \mathbb{R}^2$ then

$$Z_s(\omega) = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u\left(s, X_s^{(1)}(\omega), X_s^{(2)}(\omega)\right)$$

for almost all $(s,\omega) \in [t,T] \times \Omega$.

Proof. We already know that Z is bounded according to Lemma 6.3.1, but not in the form of the more explicit bound $\|Z\|_\infty \leq L_g$.

Notice that $\lim_{h\downarrow 0} \frac{1}{h}\int_s^{s+h} Z_r(\omega)\,\mathrm{d}r = Z_s(\omega)$ for almost all $(\omega, s) \in \Omega \times [t,T)$ due to the fundamental theorem of Lebesgue integral calculus.

Now take some $s \in [t,T)$ s.t. $\lim_{h\downarrow 0} \frac{1}{h}\int_s^{s+h} Z_r\,\mathrm{d}r = Z_s$ almost surely. Almost all $s \in [t,T)$ have this property. Choose any $h > 0$ s.t $s + h < T$ and consider the expression

$$\frac{1}{h}\mathbb{E}[Y_{s+h}(W_{s+h} - W_s)|\mathcal{F}_s]$$

for small $h > 0$. On the one hand we can write using Itô's formula

$$Y_{s+h}(W_{s+h} - W_s) = \int_s^{s+h} Y_r\,\mathrm{d}W_r + \int_s^{s+h} (W_r - W_s)Z_r\,\mathrm{d}W_r + \int_s^{s+h} Z_r\,\mathrm{d}r,$$

which leads to

$$\frac{1}{h}\mathbb{E}[Y_{s+h}(W_{s+h} - W_s)|\mathcal{F}_s] = \frac{1}{h}\mathbb{E}\left[\int_s^{s+h} Z_r\,\mathrm{d}r\middle|\mathcal{F}_s\right] \to Z_s \quad \text{for} \quad h \to 0.$$

On the other hand we can use the decoupling condition to write

$$\begin{aligned} Y_{s+h}(W_{s+h} - W_s) =& u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}\right)(W_{s+h} - W_s) \\ =& u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}\right)(W_{s+h} - W_s) \\ &+ \left(u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}\right) - u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}\right)\right)(W_{s+h} - W_s). \end{aligned}$$

After applying conditional expectations to both sides of the above equation we investigate the two summands on the right hand side separately.

First summand: Recall:

- $X_s^{(1)}$ and $X_s^{(2)}$ are $\mathcal{F}_s$-measurable,
- $X_{s+h}^{(1)} = X_s^{(1)} + (W_{s+h} - W_s)$,
- $W_{s+h} - W_s$ is independent of $\mathcal{F}_s$,
- u is deterministic, i.e. can be assumed to be a function of $\left(s, x^{(1)}, x^{(2)}\right) \in [0,T] \times \mathbb{R} \times \mathbb{R}$ only.

These properties imply

$$
\begin{aligned}
\mathbb{E}\Big[u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\Big]&\\
&= \int_{\mathbb{R}} u\left(s+h, X_s^{(1)}+z\sqrt{h}, X_s^{(2)}\right) z\sqrt{h}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z\\
&= \int_{\mathbb{R}} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\left(s+h, X_s^{(1)}+z\sqrt{h}, X_s^{(2)}\right) h\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z,
\end{aligned}
$$

which means

$$
\lim_{h\downarrow 0}\frac{1}{h}\mathbb{E}\left[u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\right] = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\left(s, X_s^{(1)}, X_s^{(2)}\right),
$$

if $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ is continuous in the first two components on $[0,T)\times\mathbb{R}^2$. Here we use that $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ is bounded by $\|g'\|_\infty$ according to the proof of Lemma 6.3.1. But even if $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ is not continuous in the first two components, we can still at least control the value

$$
\left|\frac{1}{h}\mathbb{E}\left[u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\right]\right|
$$

by $\|g'\|_\infty$.

Second summand: Recall:

- u is also Lipschitz continuous in the last component and $\|\delta'\|_\infty$ serves as a Lipschitz constant,
- $X_{s+h}^{(2)} = X_s^{(2)} + \int_s^{s+h} Z_r^2\,\mathrm{d}r$.

These properties allow us to estimate

$$
\begin{aligned}
&\frac{1}{h}\left|\mathbb{E}\left[\left(u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}\right) - u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}\right)\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\right]\right|\\
&\quad\le \frac{1}{h}\mathbb{E}\left[\left|u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}\right) - u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}\right)\right|\cdot|W_{s+h}-W_s|\Big|\mathcal{F}_s\right]\\
&\quad\le \frac{1}{h}\mathbb{E}\left[\|\delta'\|_\infty\left(\int_s^{s+h} Z_r^2\,\mathrm{d}r\right)\cdot|W_{s+h}-W_s|\Big|\mathcal{F}_s\right] \le \frac{1}{h}\|\delta'\|_\infty h\|Z\|_\infty^2\mathbb{E}[|W_{s+h}-W_s|],
\end{aligned}
$$

which clearly tends to 0 as $h\to 0$.

Conclusion: We have shown

$$
Z_s = \lim_{h\downarrow 0}\frac{1}{h}\mathbb{E}[Y_{s+h}(W_{s+h}-W_s)|\mathcal{F}_s] = \lim_{h\downarrow 0}\frac{1}{h}\mathbb{E}\left[u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\right],
$$

which is identical with $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\left(s, X_s^{(1)}, X_s^{(2)}\right)$ a.s. if $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ is continuous in the first two components on $[0,T)\times\mathbb{R}^2$ and bounded by $\|g'\|_\infty$ otherwise. □

For the sequel let u be the unique weakly regular Markovian decoupling field to the problem (6.9) constructed in Lemma 6.3.1. At least for the following result we assume for convenience $T=1$. We also use definitions and notations from the proof of Lemma 6.1.2.

Theorem 6.3.6. *Assume that* $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ *is*

- *Lipschitz continuous in the first two components on compact subsets of* $[0,1)\times\mathbb{R}^2$,
- $\mathbb{R}\backslash\{0\}$ *- valued on* $[0,1)\times\mathbb{R}^2$.

Then $\tilde{\tau}$ *is a stopping time with respect to the filtration* $(\mathcal{F}^B_\cdot) = (\mathcal{F}^B_s)_{s\in[0,\infty)}$.

Proof. We consider the system (6.9) for $t=0$ and $x^{(1)} = x^{(2)} = 0$. According to Lemma 6.3.5 we can assume

$$Z = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\left(\cdot, X^{(1)}_\cdot, X^{(2)}_\cdot\right)$$

and, thereby, have

$$X^{(2)}_s = \int_0^s Z_r^2\,\mathrm{d}r = \int_0^s \left(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\left(r, X^{(1)}_r, X^{(2)}_r\right)\right)^2 \mathrm{d}r$$

for all $s\in[0,T]$. So, we can assume that $X^{(1)}$

- is Lipschitz continuous and strictly increasing in s due to positivity of $\left(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\right)^2$ on $[0,1)\times\mathbb{R}^2$,
- starts in 0.

Therefore, for every $\omega\in\Omega$ the path

$$H^{-1}\left(X^{(2)}_\cdot(\omega)\right) : [0,1]\to[0,\infty)$$

is also Lipschitz continuous and strictly increasing in time and, therefore, has a continuous and strictly increasing inverse function on the interval

$$\left[0, H^{-1}\left(X^{(2)}_1(\omega)\right)\right] = [0,\tilde{\tau}(\omega)].$$

It is straightforward to see that this inverse is given by the process σ from the proof of Lemma 6.1.2. We can now calculate the weak derivative of σ: Firstly, note $\left(H^{-1}\right)'(x) = \frac{1}{H'(H^{-1}(x))}$ and also $H^{-1}(X^{(2)}_{\sigma_r}(\omega)) = r$ or equivalently $X^{(2)}_{\sigma_r}(\omega) = H(r)$. So, we can calculate

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}r}\sigma_r &= \frac{1}{\frac{\mathrm{d}}{\mathrm{d}s}\left(H^{-1}\left(X^{(2)}_s\right)\right)\Big|_{s=\sigma_r}} = \frac{1}{\left(H^{-1}\right)'\left(X^{(2)}_{\sigma_r}\right)Z^2_{\sigma_r}}\\ &= \frac{H'(r)}{\left(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\left(\sigma_r, X^{(1)}_{\sigma_r}, X^{(2)}_{\sigma_r}\right)\right)^2} = \frac{\beta_r^2}{\left(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\right)^2(\sigma_r, W_{\sigma_r}, H(r))}\end{aligned} \tag{6.13}$$

on $\{\sigma_r < 1\}$. Observe at this point that

$$\{\sigma_r<1\} = \left\{r < H^{-1}\left(X^{(2)}_1\right)\right\} = \{r<\tilde{\tau}\}.$$

If we define $\sigma_r := 1$ for $r > \tilde{\tau}$, then σ is still continuous and we have

$$\tilde{\tau} = \inf\{r \in [0,\infty) \,|\, \sigma_r \geq 1\}.$$

It is also straightforward to see $Z_{\sigma_r} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u(\sigma_r, W_{\sigma_r}, H(r))$ for $r \in [0,\tilde{\tau})$.

Now, remember $B_r = \int_0^r \frac{1}{\beta_s} \mathrm{d}Y_{\sigma_s}$ for $r \in [0,\tilde{\tau}]$ and also $Y_s - Y_0 = \int_0^s Z_r \,\mathrm{d}W_r$ for $s \in [0,1]$, so

$$\int_0^r \frac{\beta_s}{Z_{\sigma_s}} \mathrm{d}B_s = \int_0^r \frac{\beta_s}{Z_{\sigma_s}} \frac{1}{\beta_s} \mathrm{d}Y_{\sigma_s} = \int_0^r \frac{1}{Z_{\sigma_s}} Z_{\sigma_s} \mathrm{d}W_{\sigma_s} = W_{\sigma_r}.$$

So, if we define $\Sigma_r := W_{\sigma_r}$, we have dynamics

$$\Sigma_r = \int_0^r \frac{\beta_s}{\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u(\sigma_s, \Sigma_s, H(s))} \mathrm{d}B_s,$$

for $r \in [0,\tilde{\tau})$. So, to sum up σ, Σ fulfill on $[0,\tilde{\tau})$ the dynamics

$$\sigma_r = 0 + \int_0^r \frac{\beta_s^2}{\left(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u\right)^2 (\sigma_s, \Sigma_s, H(s))} \mathrm{d}s + \int_0^r 0 \,\mathrm{d}B_s,$$
$$\Sigma_r = 0 + \int_0^r 0 \,\mathrm{d}s + \int_0^r \frac{\beta_s}{\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u(\sigma_s, \Sigma_s, H(s))} \mathrm{d}B_s,$$

where $r \in [0,\tilde{\tau})$. Note that this dynamical system is locally Lipschitz continuous in (σ, Σ).

Now, for any $K_1, K_2 > 0$ and $K_3 \in (0,1)$ define a bounded random variable τ_{K_1,K_2,K_3} via

$$\tau_{K_1,K_2,K_3} := K_1 \wedge \inf\{r \in [0,\infty) \,|\, |\Sigma_r| \geq K_2\} \wedge \inf\{r \in [0,\infty) \,|\, \sigma_r \geq K_3\}.$$

Note that σ and Σ both remain bounded on $[0, \tau_{K_1,K_2,K_3}]$. Therefore, on $[0, \tau_{K_1,K_2,K_3}]$ the pair (σ, Σ) coincides with the unique solution $(\sigma^{K_1,K_2,K_3}, \Sigma^{K_1,K_2,K_3})$ to a Lipschitz problem, which is automatically progressively measurable w.r.t. the filtration $(\mathcal{F}^B_\cdot)$. Note that

$$\tau_{K_1,K_2,K_3} = K_1 \wedge \inf\left\{r \in [0,\infty \,|\, |\Sigma_r^{K_1,K_2,K_3}| \geq K_2\right\} \wedge \inf\left\{r \in [0,\infty) \,|\, \sigma_r^{K_1,K_2,K_3} \geq K_3\right\},$$

which is clearly a stopping time w.r.t. $(\mathcal{F}^B_\cdot)$. Furthermore, due to continuity of Σ and σ

$$\tilde{\tau} = \sup_{K_3 \in (0,1), K_1, K_2 > 0} \tau_{K_1,K_2,K_3},$$

which makes it a stopping time w.r.t. $(\mathcal{F}^B_\cdot)$ as well. □

In order to deduce sufficient conditions for Theorem 6.3.6 to hold we need to investigate higher order derivatives of u.

Assume that g, δ, g' and δ' are Lipschitz continuous, and consider the following (MLLC) system with $d = 1$, $n = 2$ and $m = 3$:

$$X_s^{(1)} = x^{(1)} + \int_t^s 1 \,\mathrm{d}W_r, \qquad X_s^{(2)} = x^{(2)} + \int_t^s \left(Z_r^{(0)}\right)^2 \mathrm{d}r,$$

$$
\begin{aligned}
Y_s^{(0)} &= g(X_T^{(1)}) - \delta(X_T^{(2)}) - \int_s^T Z_r^{(0)}\,\mathrm{d}W_r,\\
Y_s^{(1)} &= g'(X_T^{(1)}) - \int_s^T Z_r^{(1)}\,\mathrm{d}W_r - \int_s^T \left(-2Z_r^{(0)}Y_r^{(2)}\right) Z_r^{(1)}\,\mathrm{d}r,\\
Y_s^{(2)} &= -\delta'(X_T^{(2)}) - \int_s^T Z_r^{(2)}\,\mathrm{d}W_r - \int_s^T \left(-2Z_r^{(0)}Y_r^{(2)}\right) Z_r^{(2)}\,\mathrm{d}r,\\
Y_s^{(0)} &= u^{(0)}(s,X_s^{(1)},X_s^{(2)}), \quad Y_s^{(1)} = u^{(1)}(s,X_s^{(1)},X_s^{(2)}), \quad Y_s^{(2)} = u^{(2)}(s,X_s^{(1)},X_s^{(2)}).
\end{aligned}
\tag{6.14}
$$

Theorem 6.3.7. *For the above problem* (6.14) *we have $I_{\max}^M = [0,T]$. Furthermore,*

$$
u^{(0)} = u, \quad u^{(1)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u \quad \textit{and} \quad u^{(2)} = \frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u, \quad \textit{a.e.},
$$

where u is the unique weakly regular Markovian decoupling field to the problem (6.9).

In particular, u is twice weakly differentiable w.r.t. x with uniformly bounded derivatives.

Proof. The proof is in parts akin to the proof of Lemma 6.3.1 and we will seek to keep these parts short.

Let $u^{(i)}$, $i = 0,1,2$ be the unique weakly regular Markovian decoupling field on $I_{\max}^M$. We can assume $u^{(i)}$ to be continuous functions on $I_{\max}^M \times \mathbb{R}^2$ (Theorem 6.2.28).

Let $t \in I_{\max}^M$. For an arbitrary initial condition $x \in \mathbb{R}^2$ consider the corresponding processes

$$
X^{(1)}, X^{(2)}, Y^{(0)}, Y^{(1)}, Y^{(2)}, Z^{(0)}, Z^{(1)}, Z^{(2)}
$$

on $[t,T]$. Note that $X^{(1)}, X^{(2)}, Y^{(0)}, Z^{(0)}$ solve FBSDE (6.9), which implies that they coincide with the processes $X^{(1)}, X^{(2)}, Y, Z$ from (6.9) if we assume

$$
\sum_{i=1}^{2} \sup_{s\in[t,T]} \mathbb{E}_{0,\infty}[|X_s^{(i)}|^2] + \sup_{s\in[t,T]} \mathbb{E}_{0,\infty}[|Y_s|^2] \\
+ \|Z\|_\infty + \sum_{i=0}^{2} \sup_{s\in[t,T]} \mathbb{E}_{0,\infty}[|Y_s^{(i)}|^2] + \sum_{i=0}^{2} \|Z^i\|_\infty < \infty,
$$

according to Lemma 6.2.25. This condition is fulfilled due to strong regularity and the fact that we work with Markovian decoupling fields.

Now, $Y^{(0)} = Y$ implies $u(t,x) = u^{(0)}(t,x)$ for all $t \in I_{\max}^M, x \in \mathbb{R}^2$, where $I_{\max}^M$ is the maximal interval for the problem given by (6.14). We now claim that $Y^{(1)}, Y^{(2)}$ are bounded processes: Using the backward equation we have

$$
Y_s^{(2)} = \mathbb{E}_s\left[-\delta'(X_T^{(2)})\right] - \mathbb{E}_s\left[\int_s^T \left(-2Z_r^{(0)}Y_r^{(2)}\right) Z_r^{(2)}\,\mathrm{d}r\right]
$$

and, therefore,

$$
|Y_s^{(2)}| \leq \|\delta'\|_\infty + \int_s^T 2\|Z^{(0)}\|_\infty \|Z^{(2)}\|_\infty \mathbb{E}_s\left[\left|Y_r^{(2)}\right|\right]\,\mathrm{d}r,
$$

for $s \in [t, T]$, which using Gronwall's lemma implies

$$|Y_s^{(2)}| = \mathbb{E}_s \left[\left| Y_s^{(2)} \right| \right] \leq \|\delta'\|_\infty \exp\left(2T \|Z^{(0)}\|_\infty \|Z^{(2)}\|_\infty \right).$$

This in turn automatically implies boundedness of $Y^{(1)}$ according to its dynamics. Furthermore, $Y^{(1)}, Z^{(1)}$ and $Y^{(2)}, Z^{(2)}$ satisfy the BSDE which is also fulfilled by the processes $U, \check{Z}$ and $V, \tilde{Z}$ from the proof of Lemma 6.3.1 (see (6.11) and (6.12)) and so in particular

$$\begin{aligned} Y_s^{(2)} - V_s &= 0 - \int_s^T \left(Z_r^{(2)} - \tilde{Z}_r \right) \mathrm{d}W_r - \int_s^T \left(-2Z_r^{(0)} \right) \left(Y_r^{(2)} Z_r^{(2)} - V_r \tilde{Z}_r \right) \mathrm{d}r \\ &= 0 - \int_s^T \left(Z_r^{(2)} - \tilde{Z}_r \right) \mathrm{d}W_r - \int_s^T \left(-2Z_r^{(0)} \right) \left(\left(Y_r^{(2)} V_r \right) Z_r^{(2)} + V_r \left(Z_r^{(2)} - \tilde{Z}_r \right) \right) \mathrm{d}r. \end{aligned}$$

Using the boundedness of $Z^{(0)}$, $Z^{(2)}$ and V this implies using Lemma A.6.6 that $Y^{(2)} - V$ is 0 almost everywhere. Therefore, after setting $\tilde{W}_s := W_s - \int_t^s 2Z_r^{(0)} V_r \,\mathrm{d}r$, $s \in [t, T]$ we get from the above equation $\int_s^T \left(Z_r^{(2)} - \tilde{Z}_r \right) \mathrm{d}\tilde{W}_r = 0$ a.s. for $s \in [t, T]$. Since $\tilde{W}$ is a Brownian motion under some probability measure equivalent to $\mathbb{P}$ we also have $Z^{(2)} - \tilde{Z} = 0$ a.e.

Similarly, one shows that $Y^{(1)}$ and U as well as $Z^{(1)}$ and $\check{Z}$ coincide so

$$Y^{(1)} = U, \quad Y^{(2)} = V, \quad Z^{(1)} = \check{Z} \quad \text{and} \quad Z^{(2)} = \tilde{Z} \qquad \text{a.e.}$$

Now, remember $U_s = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u(s, X_s^{(1)}, X_s^{(2)})$. Together with $u^{(1)}(s, X_s^{(1)}, X_s^{(2)}) = Y_s^{(1)}$ and $Y^{(1)} = U$ this yields $u^{(1)}(t, \cdot) = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u(t, \cdot)$ and, therefore, $u^{(1)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u$ a.e. on $I_{\max}^M$. Similarly, we get $u^{(2)} = \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u$. Now, note that $u^{(1)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u$ is continuous. This makes Lemma 6.3.5 applicable, so

$$Z^{(0)} = Z = U = Y^{(1)} \text{ a.e.} \tag{6.15}$$

Thereby $Y^{(1)}$, $Y^{(2)}$ satisfy the following dynamics:

$$Y_s^{(1)} = g'(X_T^{(1)}) - \int_s^T Z_r^{(1)} \,\mathrm{d}W_r - \int_s^T \left(-2Y_r^{(1)} Y_r^{(2)} \right) Z_r^{(1)} \,\mathrm{d}r, \tag{6.16}$$

$$Y_s^{(2)} = -\delta'(X_T^{(2)}) - \int_s^T Z_r^{(2)} \,\mathrm{d}W_r - \int_s^T \left(-2Y_r^{(1)} Y_r^{(2)} \right) Z_r^{(2)} \,\mathrm{d}r, \quad s \in [t, T], \tag{6.17}$$

which implies using the chain rule of Lemma A.6.8:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y_s^{(1)} &= g''(X_T^{(1)}) \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} X_T^{(1)} - \int_s^T \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Z_r^{(1)} \,\mathrm{d}W_r \\ &\quad - \int_s^T (-2) \left(\left(\frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y_r^{(1)} Y_r^{(2)} + Y_r^{(1)} \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y_r^{(2)} \right) Z_r^{(1)} + Y_r^{(1)} Y_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Z_r^{(1)} \right) \mathrm{d}r, \end{aligned}$$

and

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y_s^{(2)} &= -\delta''(X_T^{(2)}) \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} X_T^{(2)} - \int_s^T \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Z_r^{(2)} \,\mathrm{d}W_r \\ &\quad - \int_s^T (-2) \left(\left(\frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y_r^{(1)} Y_r^{(2)} + Y_r^{(1)} \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y_r^{(2)} \right) Z_r^{(2)} + Y_r^{(1)} Y_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Z_r^{(2)} \right) \mathrm{d}r, \end{aligned}$$

for $i = 1, 2$. Let us recall some statements about the forward process obtained in the proof of Lemma 6.3.1:

$$\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)} > 0, \quad \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X^{(1)} = 1, \quad \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(1)} = 0, \quad \text{a.e.,}$$

and

$$\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}\right)^{-1} = 1 - \int_t^s 2Y_r^{(1)} Z_r^{(2)} \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_r^{(2)}\right)^{-1} \mathrm{d}r, \tag{6.18}$$

$$\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_s^{(2)} = \int_t^s 2Y_r^{(1)} \left(Z_r^{(1)} + Z_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_r^{(2)}\right) \mathrm{d}r. \tag{6.19}$$

Using the chain rule of Lemma A.6.8 and the decoupling condition, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y_s^{(i)} &= \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u^{(i)}(s, X_s^{(1)}, X_s^{(2)}) + \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u^{(i)}(s, X_s^{(1)}, X_s^{(2)}) \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_s^{(2)}, \\
\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_s^{(i)} &= \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u^{(i)}(s, X_s^{(1)}, X_s^{(2)}) \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}, \quad i = 1, 2.
\end{aligned}$$

Now, define

$$Y_s^{(12)} := \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u^{(1)}(s, X_s^{(1)}, X_s^{(2)}) = \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_s^{(1)}\right) \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}\right)^{-1}, \tag{6.20}$$

$$Y_s^{(22)} := \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} u^{(2)}(s, X_s^{(1)}, X_s^{(2)}) = \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_s^{(2)}\right) \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}\right)^{-1},$$

$$Y_s^{(11)} := \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u^{(1)}(s, X_s^{(1)}, X_s^{(2)}) = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y_s^{(1)} - Y_s^{(12)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_s^{(2)}, \tag{6.21}$$

$$Y_s^{(21)} := \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u^{(2)}(s, X_s^{(1)}, X_s^{(2)}) = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y_s^{(2)} - Y_s^{(22)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_s^{(2)}.$$

We can apply the Itô formula to deduce dynamics of $Y^{(12)}$ and $Y^{(11)}$ from dynamics of $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y^{(1)}$, $\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)}\right)^{-1}$, $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y^{(1)}$ and $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X^{(2)}$:

Let us define $Z_s^{(12)} := \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Z_s^{(1)}\right) \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_s^{(2)}\right)^{-1}$, so we can write using (6.20)

$$\begin{aligned}
Y_s^{(12)} = 0 - \int_s^T Z_r^{(12)} \,\mathrm{d}W_r - \int_s^T \Bigg\{ (-2) \Bigg(\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_r^{(1)} Y_r^{(2)} + Y_r^{(1)} \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_r^{(2)} \right) Z_r^{(1)} \\
+ Y_r^{(1)} Y_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Z_r^{(1)} \Bigg) \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_r^{(2)} \right)^{-1} - 2 \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} Y_s^{(1)} Y_r^{(1)} Z_r^{(2)} \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X_r^{(2)} \right)^{-1} \Bigg\} \mathrm{d}r.
\end{aligned}$$

Using the definitions of $Y^{(12)}$, $Y^{(22)}$, $Z^{(12)}$ we can simplify this to

$$\begin{aligned}
Y_s^{(12)} &= 0 - \int_s^T Z_r^{(12)} \,\mathrm{d}W_r \\
&\quad - \int_s^T (-2) \left(\left(Y_r^{(12)} Y_r^{(2)} + Y_r^{(1)} Y_r^{(22)} \right) Z_r^{(1)} + Y_r^{(1)} Y_r^{(2)} Z_r^{(12)} + Y_r^{(12)} Y_r^{(1)} Z_r^{(2)} \right) \mathrm{d}r.
\end{aligned}$$

Let us now define $Z_s^{(11)} := \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Z_s^{(1)} - Z_s^{(12)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_s^{(2)}$, so we can write using (6.21)

$$\begin{aligned}
Y_s^{(11)} &= g''(X_T^{(1)}) - \int_s^T Z_r^{(11)} \,\mathrm{d}W_r \\
&\quad - \int_s^T \Bigg\{ (-2) \Bigg(\Bigg(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y_r^{(1)} \, Y_r^{(2)} + Y_r^{(1)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y_r^{(2)} \Bigg) Z_r^{(1)} + Y_r^{(1)} Y_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Z_r^{(1)} \Bigg) \\
&\quad - (-2) \Big(\Big(Y_r^{(12)} \, Y_r^{(2)} + Y_r^{(1)} Y_r^{(22)} \Big) Z_r^{(1)} + Y_r^{(1)} Y_r^{(2)} Z_r^{(12)} + Y_r^{(12)} Y_r^{(1)} Z_r^{(2)} \Big) \\
&\quad \times \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_r^{(2)} - Y_r^{(12)} \cdot 2 \cdot Y_r^{(1)} \left(Z_r^{(1)} + Z_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_r^{(2)} \right) \Bigg\} \,\mathrm{d}r.
\end{aligned}$$

The two marked terms above can be effectively merged into one using (6.21):

$$\begin{aligned}
Y_s^{(11)} =& g''(X_T^{(1)}) - \int_s^T Z_r^{(11)} \,\mathrm{d}W_r \\
& - \int_s^T \Bigg\{ (-2) \Bigg(\Bigg(Y_r^{(11)} Y_r^{(2)} + Y_r^{(1)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Y_r^{(2)} \Bigg) Z_r^{(1)} + Y_r^{(1)} Y_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} Z_r^{(1)} \Bigg) \\
& - (-2) \Big(Y_r^{(1)} \, Y_r^{(22)} \, Z_r^{(1)} + Y_r^{(1)} Y_r^{(2)} \, Z_r^{(12)} + Y_r^{(12)} Y_r^{(1)} Z_r^{(2)} \Big) \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_r^{(2)} \\
& - Y_r^{(12)} \cdot 2 \cdot Y_r^{(1)} \left(Z_r^{(1)} + Z_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_r^{(2)} \right) \Bigg\} \,\mathrm{d}r.
\end{aligned}$$

Similarly, the four marked terms can be merged into only two using the structure of $Y^{(21)}$ and $Z^{(11)}$ s.t.

$$\begin{aligned}
Y_s^{(11)} &= g''(X_T^{(1)}) - \int_s^T Z_r^{(11)} \,\mathrm{d}W_r \\
&\quad - \int_s^T \Bigg\{ (-2) \Big(\Big(Y_r^{(11)} Y_r^{(2)} + Y_r^{(1)} Y_r^{(21)} \Big) Z_r^{(1)} + Y_r^{(1)} Y_r^{(2)} Z_r^{(11)} \Big) \\
&\quad + 2 \Big(Y_r^{(12)} Y_r^{(1)} \, Z_r^{(2)} \Big) \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_r^{(2)} - Y_r^{(12)} \cdot 2 \cdot Y_r^{(1)} \left(Z_r^{(1)} + \, Z_r^{(2)} \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X_r^{(2)} \right) \Bigg\} \,\mathrm{d}r,
\end{aligned}$$

where the two marked terms effectively cancel each other out:

$$\begin{aligned}
Y_s^{(11)} &= g''(X_T^{(1)}) - \int_s^T Z_r^{(11)} \,\mathrm{d}W_r \\
&\quad - \int_s^T (-2) \Big(\Big(Y_r^{(11)} Y_r^{(2)} + Y_r^{(1)} Y_r^{(21)} \Big) Z_r^{(1)} + Y_r^{(1)} Y_r^{(2)} Z_r^{(11)} + Y_r^{(12)} Y_r^{(1)} Z_r^{(1)} \Big) \,\mathrm{d}r.
\end{aligned}$$

Analogously to $Y^{(12)}$ we can deduce dynamics of $Y^{(22)}$:

$$\begin{aligned}
Y_s^{(22)} &= -\delta''(X_T^{(2)}) - \int_s^T Z_r^{(22)} \,\mathrm{d}W_r \\
&\quad - \int_s^T (-2) \Big(\Big(Y_r^{(12)} Y_r^{(2)} + Y_r^{(1)} Y_r^{(22)} \Big) Z_r^{(2)} + Y_r^{(1)} Y_r^{(2)} Z_r^{(22)} + Y_r^{(22)} Y_r^{(1)} Z_r^{(2)} \Big) \,\mathrm{d}r.
\end{aligned}$$

From here we can, analogously to $Y^{(11)}$, deduce dynamics of $Y^{(21)}$:

$$Y_s^{(21)} = 0 - \int_s^T Z_r^{(21)} \,\mathrm{d}W_r$$
$$- \int_s^T (-2) \left(\left(Y_r^{(11)} Y_r^{(2)} + Y_r^{(1)} Y_r^{(21)} \right) Z_r^{(2)} + Y_r^{(1)} Y_r^{(2)} Z_r^{(21)} + Y_r^{(22)} Y_r^{(1)} Z_r^{(1)} \right) \mathrm{d}r.$$

And so we have finally obtained the complete dynamics of the 4-dimensional process $(Y^{(ij)})$, $i, j = 1, 2$, which are clearly linear in it. Furthermore, remember:

- $Y^{(1)}$, $Y^{(2)}$ are uniformly bounded independently of (t, x) due to the decoupling condition, $u^{(i)} = \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} u$, $i = 1, 2$ and Lemma 6.3.1,
- $Z^{(1)}$, $Z^{(2)}$ are $BMO(\mathbb{P})$ processes with uniformly bounded $BMO(\mathbb{P})$-norms independently of (t, x) due to (6.16)), (6.17) and Theorem A.6.7,
- $(Y^{(ij)})$, $i, j = 1, 2$ are bounded according to their definition (with a bound which may depend on t, x at this point),
- $(Z^{(ij)})$, $i, j = 1, 2$ are in $BMO(\mathbb{P})$ according to Theorem A.6.7,
- $(Y_T^{(ij)})_{i,j=1,2}$ is uniformly bounded by $\|g''\|_\infty + \|\delta''\|_\infty < \infty$.

Therefore, Lemma A.6.6 is applicable and $(Y^{(ij)})_{i,j=1,2}$ is uniformly bounded, independently of (t, x). In particular, $Y_t^{(ij)} = \frac{\mathrm{d}}{\mathrm{d}x^{(j)}} u^{(i)}(t, x)$, $i, j = 1, 2$ can be controlled independently of $t \in I_{\max}^M$, $x \in \mathbb{R}^2$, while $\frac{\mathrm{d}}{\mathrm{d}x^{(j)}} u^{(0)}(t, x)$, $j = 1, 2$ has the same property as we already know. This shows $I_{\max}^M = [0, T]$ using Lemma 6.2.29. $\square$

Lemma 6.3.8. *Assume that g, δ, g', δ' are Lipschitz continuous. Let $\left(u^{(i)}\right)_{i=0,1,2}$ be the unique weakly regular Markovian decoupling field to the problem* (6.14) *constructed in Theorem 6.3.7.*

Assume that $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u^{(i)}$, $i = 0, 1, 2$, has a version which is continuous in the first two components $(s, x^{(1)})$ on $[t, T) \times \mathbb{R}^2$ for some $t \in [0, T)$. Then for any initial condition $(X_t^{(1)}, X_t^{(2)})^\top = (x^{(1)}, x^{(2)})^\top = x \in \mathbb{R}^2$ the associated processes $Z^{(i)}$, $i = 0, 1, 2$, on $[t, T]$ satisfy

$$Z_s^{(i)}(\omega) = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u^{(i)} \left(s, X_s^{(1)}(\omega), X_s^{(2)}(\omega) \right), \quad i = 0, 1, 2,$$

for almost all $(s, \omega) \in [t, T] \times \Omega$.

Furthermore, in this case the processes

$$\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} X^{(2)}, \quad \frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)} \quad \textit{and} \quad \left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}} X^{(2)} \right)^{-1} \quad \textit{on } [t, T],$$

can be bounded uniformly, i.e. independently of (t, x).

Proof. The first part of the proof works analogously to the proof of Lemma 6.3.5. So we keep our arguments short. For $i = 0, 1, 2$ we consider

$$\frac{1}{h}\mathbb{E}[Y_{s+h}^{(i)}(W_{s+h} - W_s)|\mathcal{F}_s]$$

for small $h > 0$. As in the proof of Lemma 6.3.5, we use Itô's formula applied to (6.14) to obtain

$$\begin{aligned} Y_{s+h}^{(i)}(W_{s+h} - W_s) = &\int_s^{s+h} Y_r^{(i)} \,\mathrm{d}W_r + \int_s^{s+h} (W_r - W_s) Z_r^{(i)} \,\mathrm{d}W_r \\ &+ \int_s^{s+h} (W_r - W_s) \left(-2Z_r^{(0)} Y_r^{(2)}\right) Z_r^{(i)} \,\mathrm{d}r + \int_s^{s+h} Z_r^{(i)} \,\mathrm{d}r, \end{aligned}$$

and also

$$Y_{s+h}^{(0)}(W_{s+h} - W_s) = \int_s^{s+h} Y_r^{(0)} \,\mathrm{d}W_r + \int_s^{s+h} (W_r - W_s) Z_r^{(0)} \,\mathrm{d}W_r + \int_s^{s+h} Z_r^{(0)} \,\mathrm{d}r,$$

which leads to

$$\frac{1}{h}\mathbb{E}[Y_{s+h}^{(0)}(W_{s+h} - W_s)|\mathcal{F}_s] = \frac{1}{h}\mathbb{E}\left[\int_s^{s+h} Z_r^{(0)} \,\mathrm{d}r \middle| \mathcal{F}_s\right] \to Z_s^{(0)} \quad \text{for} \quad h \to 0,$$

and

$$\begin{aligned} &\frac{1}{h}\mathbb{E}[Y_{s+h}^{(i)}(W_{s+h} - W_s)|\mathcal{F}_s] \\ &\qquad = \frac{1}{h}\mathbb{E}\left[\int_s^{s+h} Z_r^{(i)} \left(1 + (W_r - W_s)\left(-2Z_r^{(0)} Y_r^{(2)}\right)\right) \mathrm{d}r \middle| \mathcal{F}_s\right] \to Z_s^{(i)} \end{aligned}$$

as $h \to 0$ for $i = 1, 2$. The arguments are valid for almost all $s \in [t, T]$.

On the other hand we can use the decoupling condition to rewrite

$$\begin{aligned} &Y_{s+h}^{(i)}(W_{s+h} - W_s) \\ &\quad = u^{(i)}\left(s + h, X_{s+h}^{(1)}, X_s^{(2)}\right)(W_{s+h} - W_s) \\ &\qquad + \left(u^{(i)}\left(s + h, X_{s+h}^{(1)}, X_{s+h}^{(2)}\right) - u^{(i)}\left(s + h, X_{s+h}^{(1)}, X_s^{(2)}\right)\right)(W_{s+h} - W_s). \end{aligned}$$

Let us deal separately with the two summands. For the first one recall that

- $X_s^{(1)}$ and $X_s^{(2)}$ are $\mathcal{F}_s$-measurable,
- $X_{s+h}^{(1)} = X_s^{(1)} + (W_{s+h} - W_s)$,
- $W_{s+h} - W_s$ is independent of $\mathcal{F}_s$,
- u is deterministic, i.e. is assumed to be a function of $\left(s, x^{(1)}, x^{(2)}\right) \in [0, T] \times \mathbb{R}^2$.

A combination of these properties leads to

$$\lim_{h\downarrow 0}\frac{1}{h}\mathbb{E}\left[u^{(i)}\left(s+h,X^{(1)}_{s+h},X^{(2)}_s\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\right]=\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(i)}\left(s,X^{(1)}_s,X^{(2)}_s\right),$$

if $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(i)}$ is continuous in the first two components on $[t,T)\times\mathbb{R}^2$, where we use that $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(i)}$ is bounded.

For the second summand recall that

- $u^{(i)}$ is also Lipschitz continuous in the last component with some Lipschitz constant L,
- $X^{(2)}_{s+h}=X^{(2)}_s+\int_s^{s+h}(Z^{(0)}_r)^2\,\mathrm{d}r$.

These properties allow us to estimate

$$\begin{aligned}&\frac{1}{h}\left|\mathbb{E}\left[\left(u^{(i)}\left(s+h,X^{(1)}_{s+h},X^{(2)}_{s+h}\right)-u^{(i)}\left(s+h,X^{(1)}_{s+h},X^{(2)}_s\right)\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\right]\right|\\&\quad\le\frac{1}{h}\mathbb{E}\left[L\cdot\left(\int_s^{s+h}(Z^{(0)}_r)^2\,\mathrm{d}r\right)\cdot|W_{s+h}-W_s|\Big|\mathcal{F}_s\right]\le\frac{1}{h}L\cdot h\|Z^{(0)}\|^2_\infty\mathbb{E}[|W_{s+h}-W_s|],\end{aligned}$$

which tends to 0 as $h\to 0$. Therefore, we can conclude

$$Z^{(i)}_s=\lim_{h\downarrow 0}\frac{1}{h}\mathbb{E}[Y^{(i)}_{s+h}(W_{s+h}-W_s)|\mathcal{F}_s]=\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(i)}\left(s,X^{(1)}_s,X^{(2)}_s\right)$$

if $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(i)}$ is continuous in the first two components on $[t,T)\times\mathbb{R}^2$, for $i=0,1,2$.

Now recall (6.18)) and (6.19) from the proof of Theorem 6.3.7:

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}X^{(2)}_s\right)^{-1}&=1-\int_t^s 2Y^{(1)}_rZ^{(2)}_r\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}X^{(2)}_r\right)^{-1}\mathrm{d}r,\\ \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X^{(2)}_s&=\int_t^s 2Y^{(1)}_r\left(Z^{(1)}_r+Z^{(2)}_r\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X^{(2)}_r\right)\mathrm{d}r,\end{aligned}$$

a.s. for $s\in[t,T]$. The first equation implies

$$\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}X^{(2)}_s\right)^{-1}=\exp\left(-\int_t^s 2Y^{(1)}_rZ^{(2)}_r\,\mathrm{d}r\right).$$

Using $Z^{(2)}=\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(2)}(\cdot,X^{(1)}_\cdot,X^{(2)}_\cdot)$, $Y^{(1)}=Z^{(0)}=\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(0)}(\cdot,X^{(1)}_\cdot,X^{(2)}_\cdot)$ (see (6.15) in the proof of Theorem 6.3.7) and uniform boundedness of $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(i)}$ for $i=0,1,2$ we see that this implies uniform boundedness of $\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}X^{(2)}_s\right)^{-1}$ and its inverse $\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}X^{(2)}_s$.

Furthermore,

$$\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X^{(2)}_s\right|\le 2T\|Y^{(1)}Z^{(1)}\|_\infty+\int_t^s 2\|Y^{(1)}Z^{(2)}\|_\infty\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X^{(2)}_r\right|\mathrm{d}r.$$

By Gronwall's lemma together with uniform boundedness of $Z^{(2)}$, $Y^{(1)}$ and $Z^{(1)}=\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(1)}(\cdot,X^{(1)}_\cdot,X^{(2)}_\cdot)$ this implies uniform boundedness of $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X^{(2)}$. □

For the subsequent result we employ the following notation:

- For a real number $H > 0$ let $\chi_H : \mathbb{R} \to \mathbb{R}$ be defined via $\chi_H(x) := (-H) \vee (x \wedge H)$ for $x \in \mathbb{R}$. In particular, χ_H is bounded, Lipschitz continuous and coincides with the identity function on the interval $[-H, H]$.
- For real numbers $y^{(ij)}$, $i, j = 1, 2$ and $y^{(i)}$, $i = 1, 2$ we denote by $y^{(ij)\wedge H}$ and $y^{(i)\wedge H}$ the values $\chi_H(y^{(ij)})$ and $\chi_H(y^{(i)})$.

Now assume that g, δ, g', δ', g'', δ'' are all Lipschitz continuous and consider for $H > 0$ the following (MLLC) system with $d = 1$, $n = 2$ and $m = 6$:

$$X_s^{(1)} = x^{(1)} + \int_t^s 1 \, \mathrm{d}W_r, \qquad X_s^{(2)} = x^{(2)} + \int_t^s \left(Z_r^{(0)}\right)^2 \mathrm{d}r,$$
$$Y_s^{(0)} = g(X_T^{(1)}) - \delta(X_T^{(2)}) - \int_s^T Z_r^{(0)} \, \mathrm{d}W_r,$$
$$Y_s^{(1)} = g'(X_T^{(1)}) - \int_s^T Z_r^{(1)} \, \mathrm{d}W_r - \int_s^T \left(-2Z_r^{(0)} Y_r^{(2)}\right) Z_r^{(1)} \, \mathrm{d}r,$$
$$Y_s^{(2)} = -\delta'(X_T^{(2)}) - \int_s^T Z_r^{(2)} \, \mathrm{d}W_r - \int_s^T \left(-2Z_r^{(0)} Y_r^{(2)}\right) Z_r^{(2)} \, \mathrm{d}r,$$
$$Y_s^{(0)} = u^{(0)}(s, X_s^{(1)}, X_s^{(2)}), \quad Y_s^{(1)} = u^{(1)}(s, X_s^{(1)}, X_s^{(2)}), \quad u^{(2)}(s, X_s^{(1)}, X_s^{(2)}) = Y_s^{(2)}.$$

and

$$\begin{aligned}
Y_s^{(11)} =& g''(X_T^{(1)}) - \int_s^T Z_r^{(11)} \, \mathrm{d}W_r + \int_s^T 2\Big\{ \left(Y_r^{(11)\wedge H} Y_r^{(2)\wedge H} + Y_r^{(1)\wedge H} Y_r^{(21)\wedge H}\right) Z_r^{(1)} \\
&+ Y_r^{(1)\wedge H} Y_r^{(2)\wedge H} Z_r^{(11)} + Y_r^{(12)\wedge H} Y_r^{(1)\wedge H} Z_r^{(1)} \Big\} \mathrm{d}r, \\
Y_s^{(12)} =& 0 - \int_s^T Z_r^{(12)} \, \mathrm{d}W_r + \int_s^T 2\Big\{ \left(Y_r^{(12)\wedge H} Y_r^{(2)\wedge H} + Y_r^{(1)\wedge H} Y_r^{(22)\wedge H}\right) Z_r^{(1)} \\
&+ Y_r^{(1)\wedge H} Y_r^{(2)\wedge H} Z_r^{(12)} + Y_r^{(12)\wedge H} Y_r^{(1)\wedge H} Z_r^{(2)} \Big\} \mathrm{d}r, \\
Y_s^{(21)} =& 0 - \int_s^T Z_r^{(21)} \, \mathrm{d}W_r + \int_s^T 2\Big\{ \left(Y_r^{(11)\wedge H} Y_r^{(2)\wedge H} + Y_r^{(1)\wedge H} Y_r^{(21)\wedge H}\right) Z_r^{(2)} \\
&+ Y_r^{(1)\wedge H} Y_r^{(2)\wedge H} Z_r^{(21)} + Y_r^{(22)\wedge H} Y_r^{(1)\wedge H} Z_r^{(1)} \Big\} \mathrm{d}r, \\
Y_s^{(22)} =& -\delta''(X_T^{(2)}) - \int_s^T Z_r^{(22)} \, \mathrm{d}W_r + \int_s^T 2\Big\{ \left(Y_r^{(12)\wedge H} Y_r^{(2)\wedge H} + Y_r^{(1)\wedge H} Y_r^{(22)\wedge H}\right) Z_r^{(2)} \\
&+ Y_r^{(1)\wedge H} Y_r^{(2)\wedge H} Z_r^{(22)} + Y_r^{(22)\wedge H} Y_r^{(1)\wedge H} Z_r^{(2)} \Big\} \mathrm{d}r,
\end{aligned}$$

with the decoupling conditions

$$\begin{aligned}
&u^{(11)}(s, X_s^{(1)}, X_s^{(2)}) = Y_s^{(11)}, \qquad u^{(12)}(s, X_s^{(1)}, X_s^{(2)}) = Y_s^{(12)}, \\
&u^{(21)}(s, X_s^{(1)}, X_s^{(2)}) = Y_s^{(21)}, \qquad u^{(22)}(s, X_s^{(1)}, X_s^{(2)}) = Y_s^{(22)}.
\end{aligned} \tag{6.22}$$

With (6.22) we will always refer to all the above equations belonging to the current (MLLC) system.

Theorem 6.3.9. *For sufficiently large $H > 0$ the above problem* (6.22) *satisfies $I^M_{\max} = [0,T]$ and in addition*

$$u^{(0)} = u, \quad u^{(1)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u, \quad u^{(2)} = \frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u, \quad u^{(11)} = \frac{\mathrm{d}^2}{\left(\mathrm{d}x^{(1)}\right)^2}u,$$
$$u^{(12)} = \frac{\mathrm{d}}{\mathrm{d}x^{(2)}}\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u, \quad u^{(21)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u, \quad u^{(22)} = \frac{\mathrm{d}^2}{\left(\mathrm{d}x^{(2)}\right)^2}u, \quad a.e.,$$

where u is the unique weakly regular Markovian decoupling field to the problem (6.9).

In particular, u is three times weakly differentiable w.r.t. x with uniformly bounded derivatives.

Proof. The proof is in parts akin to the proof of Lemma 6.3.1, and we will again seek to keep these parts short.

Assume $I^M_{\max} = (t^M_{\min}, T]$ and $t \in I^M_{\max}$. Let $u^{(i)}$ and $u^{(jk)}$, $i = 0,1,2,\ j,k = 1,2$ be the associated weakly regular decoupling field on $I^M_{\max}$. We want to control $\frac{\mathrm{d}}{\mathrm{d}x}u^{(i)}u(t,\cdot)$, $\frac{\mathrm{d}}{\mathrm{d}x}u^{(jk)}(t,\cdot)$, $i = 0,1,2,\ j,k = 1,2$ independently of t to create a contradiction according to Lemma 6.2.29.

For this purpose consider the first three components of the decoupling field. Since $\left(u^{(i)}\right)_{i=0,1,2}$ is clearly a weakly regular Markovian decoupling field to the problem (6.14)

- the mappings $\left(u^{(i)}\right)_{i=0,1,2}$ in (6.14) and in (6.22) are identical according to Theorem 6.2.24,
- the processes $X^{(1)}$, $X^{(2)}$, $Y^{(i)}$, $Z^{(i)}$, $i = 0,1,2$ in (6.14) must coincide with the identically denoted processes in (6.22) according to strong regularity. This is true for every $t \in I^M_{\max}$ and initial condition $x \in \mathbb{R}^2$.

So we can apply Theorem 6.3.7 and get

$$u^{(0)} = u, \quad u^{(1)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u, \quad u^{(2)} = \frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u \quad \text{on} \quad I^M_{\max}.$$

In particular, the last two functions are uniformly bounded.

Furthermore, we saw in the proof of Theorem 6.3.7 that

- $Y^{(1)}$ and $Y^{(2)}$ are uniformly bounded independently of (t,x),
- $Z^{(1)}$ and $Z^{(2)}$ are $BMO(\mathbb{P})$ processes with uniformly bounded $BMO(\mathbb{P})$-norms independently of (t,x).

Especially, $Y^{(i)\wedge H} = Y^{(i)}$ for $i = 1,2$ if we make H large enough. We will make this assumption from now on.

The processes $Y^{(jk)}$, $j,k = 1,2$ satisfy

$$\begin{aligned} Y^{(jk)}_s =& Y^{(jk)}_T - \int_s^T Z^{(jk)}_r \,\mathrm{d}W_r \\ &- \int_s^T \left(\sum_{l_1,l_2,l_3,l_4=1,2} \alpha^{(jk)}_{l_1,l_2,l_3,l_4} Y^{(l_1)}_r Z^{(l_2)}_r Y^{(l_3 l_4)\wedge H}_r + Y^{(1)}_r Y^{(2)}_r Z^{(jk)}_r \right) \mathrm{d}r, \end{aligned}$$

where $\alpha^{(jk)}_{l_1,l_2,l_3,l_4}$ is always either 0 or -2. Since due to the structure of the terminal condition $Y^{(jk)}_T$ are uniformly bounded, we can apply Lemma A.6.6 to obtain uniform boundedness of $Y^{(jk)}$ as processes on $[t,T]$ independently of (t,x).

In particular, $Y^{(jk)\wedge H} = Y^{(jk)}$ for $jk = 1,2$ if we make H large enough. We will make this assumption from now on.

This implies that the processes $Y^{(jk)}$, $j,k = 1,2$ must coincide with the identically denoted processes in the proof of Theorem 6.3.7, since

- they satisfy the same stochastic differential equations,
- they satisfy the same terminal condition and
- we can apply Lemma A.6.6 to the difference of these four-dimensional processes obtaining that this difference must vanish.

This implies however that $Y^{(jk)}_t = \frac{\mathrm{d}}{\mathrm{d}x^{(k)}} u^{(j)}\left(t, x^{(1)}, x^{(2)}\right)$ for almost all $x^{(1)}, x^{(2)}$. So we obtain $u^{(jk)} = \frac{\mathrm{d}}{\mathrm{d}x^{(k)}} u^{(j)}$, $j,k = 1,2$ a.e and these functions are uniformly bounded according to Theorem 6.3.7.

According to Remark 6.2.21, the functions $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u = u^{(1)}$, $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u^{(i)} = u^{(i1)}$, $i = 1,2$ are continuous on $[t,T]\times\mathbb{R}^2$ and we can apply Lemma 6.3.8 to get

$$Z^{(i)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}} u^{(i)}\left(\cdot, X^{(1)}_\cdot, X^{(2)}_\cdot\right), \quad i = 0,1,2.$$

Hence, $Z^{(i)}$, $i = 0,1,2$ are uniformly bounded.

Let us now analyze higher order derivatives $\frac{\mathrm{d}}{\mathrm{d}x^{(i)}} u^{(jk)}$, $i,j,k = 1,2$. As usual this is done by investigating equations characterizing the dynamics of $\frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(jk)}$, $i,j,k = 1,2$. Using strong regularity we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(jk)}_s =& \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(jk)}_T - \int_s^T \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Z^{(jk)}_r \,\mathrm{d}W_r \\
& - \int_s^T \left(G^{(jk)}_r + \sum_{l_1,l_2,l_3,l_4=1,2} \alpha^{(jk)}_{l_1,l_2,l_3,l_4} H^{(jk),l_1,l_2,l_3,l_4}_r \right) \mathrm{d}r,
\end{aligned}$$

where

$$H^{i,(jk),l_1,l_2,l_3,l_4}_r = \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(l_1)}_r Z^{(l_2)}_r Y^{(l_3 l_4)}_r + Y^{(l_1)}_r \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Z^{(l_2)}_r Y^{(l_3 l_4)}_r + Y^{(l_1)}_r Z^{(l_2)}_r \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(l_3 l_4)}_r,$$

$$G^{i,(jk)}_r = \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(1)}_r Y^{(2)}_r Z^{(jk)}_r + Y^{(1)}_r \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(2)}_r Z^{(jk)}_r + Y^{(1)}_r Y^{(2)}_r \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Z^{(jk)}_r.$$

This already implies that $\frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(jk)}$, $i,j,k = 1,2$, is uniformly bounded according to Lemma A.6.6. The lemma is applicable since

- $\frac{\mathrm{d}}{\mathrm{d}x^{(i)}} Y^{(jk)}_T$ is either 0 or has the structure

  $$g^{(3)}(X^{(1)}_T) \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} X^{(1)}_T \text{ or } -\delta^{(3)}(X^{(2)}_T) \frac{\mathrm{d}}{\mathrm{d}x^{(i)}} X^{(2)}_T$$

  which is uniformly bounded according to the Lipschitz continuity of g'', δ'' and Lemma 6.3.8,

- $\frac{\mathrm{d}}{\mathrm{d}x^{(i)}}Y_r^{(l)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(l)}(r, X_r^{(1)}, X_r^{(2)})\frac{\mathrm{d}}{\mathrm{d}x^{(i)}}X_r^{(1)} + \frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u^{(l)}(r, X_r^{(1)}, X_r^{(2)})\frac{\mathrm{d}}{\mathrm{d}x^{(i)}}X_r^{(2)}$ is also uniformly bounded according to Theorem 6.3.7 and Lemma 6.3.8,
- $\frac{\mathrm{d}}{\mathrm{d}x^{(i)}}Y_r^{(jk)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(jk)}(r, X_r^{(1)}, X_r^{(2)})\frac{\mathrm{d}}{\mathrm{d}x^{(i)}}X_r^{(1)} + \frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u^{(jk)}(r, X_r^{(1)}, X_r^{(2)})\frac{\mathrm{d}}{\mathrm{d}x^{(i)}}X_r^{(2)}$ is a bounded processes on $[t, T]$ according to Lemma 6.3.8 (but not necessarily uniformly in t at this point),
- $\frac{\mathrm{d}}{\mathrm{d}x^{(i)}}Z_r^{(l)} = \frac{\mathrm{d}}{\mathrm{d}x^{(i)}}u^{(l)}\left(r, X_r^{(1)}, X_r^{(2)}\right) = \frac{\mathrm{d}}{\mathrm{d}x^{(i)}}Y_r^{(l1)}$ for all $l = 1, 2$,
- $Y^{(l_1l_2)}$, $Y^{(l)}$, $Z^{(l)}$ are always uniformly bounded as was already mentioned,
- $Z^{(l_1l_2)}$ are $BMO(\mathbb{P})$-processes with uniformly bounded $BMO(\mathbb{P})$ - norms according to the equations describing $Y^{(l_1l_2)}$ and Theorem A.6.7.

Let $j, k \in \{1, 2\}$. As a consequence of the decoupling condition together with the chain rule of Lemma A.6.8 we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}Y_r^{(jk)} &= \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(jk)}(r, X_r^{(1)}, X_r^{(2)}) + \frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u^{(jk)}(r, X_r^{(1)}, X_r^{(2)})\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_r^{(2)},\\
\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}Y_r^{(jk)} &= \frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u^{(jk)}(r, X_r^{(1)}, X_r^{(2)})\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}X_r^{(2)}.
\end{aligned}$$

Using the boundedness of $\left(\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}X^{(2)}\right)^{-1}$, the second equation implies boundedness of

$$\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u^{(jk)}(t, x^{(1)}, x^{(2)})$$

for almost all $x^{(1)}, x^{(2)}$ by a uniform constant. Now the first equation together with uniform boundedness of $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}X_r^{(2)}$ and $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}Y_r^{(jk)}$ implies uniform boundedness of $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(jk)}$ as well.

Considering Lemma 6.2.29 we have a contradiction and the proof is complete. □

Lemma 6.3.10. *Let $T = 1$ and assume that*

- *g, δ, g', δ', g'', δ'' are all Lipschitz continuous,*
- *g is increasing and not constant.*

Then the Markovian decoupling field u from Lemma 6.3.1 fulfills the requirements of Theorem 6.3.6.

For the proof of Lemma 6.3.10 we need the following auxiliary lemma.

Lemma 6.3.11. *Let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be twice weakly differentiable s.t. $\varphi(0) = 0$ and $\|\varphi''\|_\infty < \infty$. Then*

$$\left|\int_{\mathbb{R}} \varphi(\sigma \cdot z)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z\right| \leq \frac{1}{2}\sigma^2\|\varphi''\|_\infty,$$

for all $\sigma \in [0, \infty)$.

Proof. Using weak differentiability of φ we can write for any $x \in \mathbb{R}$:

$$\begin{aligned}\varphi(x) &= \int_0^1 \varphi'(sx)x\,\mathrm{d}s \\ &= x\int_0^1 \left(\varphi'(0) + \int_0^1 \varphi''(tsx)sx\,\mathrm{d}t\right)\mathrm{d}s = x\varphi'(0) + x^2\int_0^1 s\int_0^1 \varphi''(tsx)\,\mathrm{d}t\,\mathrm{d}s,\end{aligned}$$

and so

$$\begin{aligned}&\int_{\mathbb{R}} \varphi(\sigma\cdot z)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z \\ &= \int_{\mathbb{R}} \sigma z\varphi'(0)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z + \int_{\mathbb{R}}\sigma^2 z^2\left(\int_0^1 s\int_0^1 \varphi''(ts\sigma z)\,\mathrm{d}t\,\mathrm{d}s\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z.\end{aligned}$$

The first summand clearly vanishes and we can finally estimate:

$$\begin{aligned}\left|\int_{\mathbb{R}} \varphi(\sigma\cdot z)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z\right| &\le \sigma^2\int_{\mathbb{R}} z^2\int_0^1 s\int_0^1 |\varphi''(ts\sigma z)|\,\mathrm{d}t\,\mathrm{d}s\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z \\ &\le \sigma^2\int_{\mathbb{R}} z^2\int_0^1 s\|\varphi''\|_\infty\,\mathrm{d}s\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z \\ &= \sigma^2\int_{\mathbb{R}} z^2\frac{1}{2}\|\varphi''\|_\infty\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z = \sigma^2\frac{\|\varphi''\|_\infty}{2}.\end{aligned}$$

□

Proof of Lemma 6.3.10. Denote by $\left(u^{(0)}, u^{(1)}, u^{(2)}, u^{(11)}, u^{(12)}, u^{(21)}, u^{(22)}\right)$ the unique Markovian decoupling field to the problem (6.22) on $[0,T]$. We have $u^{(0)} = u$, $u^{(1)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$, etc. according to Theorem 6.3.9.

Let us show that $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ is Lipschitz continuous in the *first* component (i. e. time). For this purpose, consider for a starting time $t \in [0,T]$ and initial condition $x \in \mathbb{R}^2$ the associated FBSDE (6.22) on $[t,1]$. Recall that

$$Y_s^{(1)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s, X_s^{(1)}, X_s^{(2)}), \quad s \in [t,1], \tag{6.23}$$

satisfies

$$Y_s^{(1)} = Y_t^{(1)} + \int_t^s \left(-2Z_r^{(0)}Y_r^{(2)}\right)Z_r^{(1)}\,\mathrm{d}r + \int_t^s Z_r^{(1)}\,\mathrm{d}W_r, \quad s \in [t,1], \tag{6.24}$$

where

- $Z^{(0)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(0)}\left(\cdot, X_\cdot^{(1)}, X_\cdot^{(2)}\right) = Y^{(1)}$ a.e. according to Lemma 6.3.8, which is applicable since $\left(\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(i)}\right)_{i=1,2} = \left(u^{(i1)}\right)_{i=1,2}$ and $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(0)} = u^{(1)}$ are continuous on $[t,1]$ according to Remark 6.2.21,
- $Z^{(0)} = Y^{(1)}$ and $Y^{(2)}$ are bounded by $\left\|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\right\|_\infty$ and $\left\|\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u\right\|_\infty$,
- $Z^{(1)} = \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(1)}\left(\cdot, X_\cdot^{(1)}, X_\cdot^{(2)}\right)$ a.e. according to Lemma 6.3.8, which is applicable as already mentioned. So $Z^{(1)}$ is bounded by $\left\|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(1)}\right\|_\infty$.

Let $s \in (t, 1]$. Using the triangular inequality we obtain

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x) - \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(t,x)\right| \leq &\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x) - \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},X_s^{(2)})\right]\right| \\ &+ \left|\mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},X_s^{(2)})\right] - \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(t,x)\right|.\end{aligned}$$

Applying the triangular inequality for a second time together with (6.23) we get

$$\begin{aligned}&\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x) - \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(t,x)\right| \\ &\leq \left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x^{(1)},x^{(2)}) - \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},x^{(2)})\right]\right| \\ &+ \left|\mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},x^{(2)})\right] - \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},X_s^{(2)})\right]\right| + \left|\mathbb{E}\left[Y_s^{(1)} - Y_t^{(1)}\right]\right|.\end{aligned}$$

Let us now control the three summands on the right-hand-side separately.
First summand: Define

$$\varphi(z) := \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x^{(1)},x^{(2)}) - \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x^{(1)}+z,x^{(2)}), \quad z \in \mathbb{R},$$

and note:

- $$\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x^{(1)},x^{(2)}) - \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},x^{(2)})\right]\right| = \left|\int_{\mathbb{R}} \varphi(\sqrt{s-t}z)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z\right|,$$
  since
  $$X_s^{(1)} = x^{(1)} + W_s - W_t \sim \mathcal{N}\left(x^{(1)}, s-t\right),$$
- φ is Lipschitz continuous with Lipschitz constant $L_{u^{(1)}}$, which is the Lipschitz constant of $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u = u^{(1)}$ w.r.t. the last two components,
- φ' is Lipschitz continuous with Lipschitz constant $L_{u^{(11)}}$, which is the Lipschitz constant of $\frac{\mathrm{d}^2}{\left(\mathrm{d}x^{(1)}\right)^2}u = u^{(11)}$ w.r.t. the last two components,
- $\varphi(0) = 0$.

And so using Lemma 6.3.11 we obtain

$$\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x^{(1)},x^{(2)}) - \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},x^{(2)})\right]\right| \leq \frac{1}{2}(s-t)\cdot L_{u^{(11)}}.$$

Second summand: We have

$$\begin{aligned}&\left|\mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},x^{(2)})\right] - \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},X_s^{(2)})\right]\right| \\ &\quad\leq \mathbb{E}\left[\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},x^{(2)}) - \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,X_s^{(1)},X_s^{(2)})\right|\right] \leq L_{u^{(1)}}\mathbb{E}\left[\left|X_s^{(2)} - x^{(2)}\right|\right],\end{aligned}$$

while

$$\left|X_s^{(2)} - x^{(2)}\right| = \left|\int_t^s \left(Z_r^{(0)}\right)^2 \mathrm{d}r\right| \leq (s-t) \cdot \left\|Y^{(1)}\right\|_\infty^2 \leq (s-t) \cdot \left\|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\right\|_\infty^2 \text{ a.s.,}$$

where we used $Z^{(0)} = Y^{(1)}$ a.e.

Third summand: We have using (6.24):

$$\begin{aligned}\left|\mathbb{E}\left[Y_s^{(1)} - Y_t^{(1)}\right]\right| &= \left|\mathbb{E}\left[-2\int_t^s Y_r^{(1)} Y_r^{(2)} Z_r^{(1)} \,\mathrm{d}r\right]\right| \\ &\leq 2 \cdot (t-s) \cdot \left\|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u\right\|_\infty \cdot \left\|\frac{\mathrm{d}}{\mathrm{d}x^{(2)}}u\right\|_\infty \cdot \left\|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u^{(1)}\right\|_\infty .\end{aligned}$$

Conclusion: We have shown

$$\left|\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x) - \frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(t,x)\right| \leq C|s-t|,$$

with some constant $C \in [0,\infty)$, which does not depend on t, x or s. In other words $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ is Lipschitz continuous in time. Since it is also Lipschitz continuous in space, it is a Lipschitz continuous function on its whole domain $[0,T] \times \mathbb{R}^2$.

It remains to show that $\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u$ is $\mathbb{R}\backslash\{0\}$-valued on $[0,1) \times \mathbb{R}^2$:

Clearly g' is non-negative and does not vanish. Let $t \in [0,1)$, $x \in \mathbb{R}^2$. Consider the associated FBSDE on $[t,1]$. Using (6.24) we can write

$$\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x) = g'(X_T^{(1)}) - \int_t^T Z_r^{(1)} \,\mathrm{d}\left(W_r + \int_t^r \left(-2Y_\kappa^{(1)} Y_\kappa^{(2)}\right) \mathrm{d}\kappa\right).$$

So there is a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x) = \mathbb{E}_{\mathbb{Q}}\left[g'\left(X_T^{(1)}\right)\right] \geq 0.$$

Now note that $X_T^{(1)} = x^{(1)} + W_T - W_t$ has a non-degenerate normal distribution w.r.t. $\mathbb{P}$. Therefore its distribution is equivalent to the Lebesgue measure. But since $\mathbb{Q} \sim \mathbb{P}$ the distribution of $X_T^{(1)}$ w.r.t. $\mathbb{Q}$ must also be equivalent to the Lebesgue measure. This shows

$$\frac{\mathrm{d}}{\mathrm{d}x^{(1)}}u(s,x) = \mathbb{E}_{\mathbb{Q}}\left[g'\left(X_T^{(1)}\right)\right] > 0$$

since otherwise $g' = 0$ a.e. would hold. $\square$

A. Appendix

A.1. Pathwise Hoeffding inequality

In the construction of the pathwise Itô integral for typical price processes as done in Chapter 2, we needed the following result, a pathwise formulation of the Hoeffding inequality which is due to Vovk. Here we present a slightly adapted version.

Lemma A.1.1 ([Vov12], Theorem A.1)**.** *Let $(\tau_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence of stopping times with $\tau_0 = 0$, such that for every $\omega \in \Omega$ we have $\tau_n(\omega) = \infty$ for all but finitely many $n \in \mathbb{N}$. Let for $n \in \mathbb{N}$ the function $h_n \colon \Omega \to \mathbb{R}^d$ be $\mathcal{F}_{\tau_n}$-measurable, and suppose that there exists a $\mathcal{F}_{\tau_n}$-measurable bounded function $b_n \colon \Omega \to \mathbb{R}$, such that*

$$\sup_{t\in[0,T]} |h_n(\omega) S_{\tau_n\wedge t,\tau_{n+1}\wedge t}(\omega)| \leq b_n(\omega) \tag{A.1}$$

for all $\omega \in \Omega$. Then for every $\lambda \in \mathbb{R}$ there exists a simple strategy $H^\lambda \in \mathcal{H}_1$ such that

$$1 + (H^\lambda \cdot S)_t \geq \exp\left(\lambda \sum_{n=0}^{\infty} h_n S_{\tau_n\wedge t,\tau_{n+1}\wedge t} - \frac{\lambda^2}{2} \sum_{n=0}^{N_t} b_n^2\right)$$

for all $t \in [0,T]$, where $N_t := \max\{n \in \mathbb{N} : \tau_n \leq t\}$.

Proof. Let $\lambda \in \mathbb{R}$. The proof is based on the following deterministic inequality: if (A.1) is satisfied, then for all $\omega \in \Omega$ and all $t \in [0,T]$ we have that

$$\begin{aligned}
&\exp\left(\lambda h_n(\omega) S_{\tau_n\wedge t,\tau_{n+1}\wedge t}(\omega) - \frac{\lambda^2}{2} b_n^2(\omega)\right) - 1 \\
&\quad \leq \exp\left(-\frac{\lambda^2}{2} b_n^2(\omega)\right) \frac{e^{\lambda b_n(\omega)} - e^{-\lambda b_n(\omega)}}{2 b_n(\omega)} h_n(\omega) S_{\tau_n\wedge t,\tau_{n+1}\wedge t}(\omega) \\
&\quad =: f_n(\omega) S_{\tau_n\wedge t,\tau_{n+1}\wedge t}(\omega).
\end{aligned} \tag{A.2}$$

This inequality is shown in (A.1) of [Vov12]. We define $H_t^\lambda := \sum_n F_n 1_{(\tau_n,\tau_{n+1}]}(t)$, with F_n that have to be specified. We choose $F_0 := f_0$, which is bounded and $\mathcal{F}_{\tau_0}$-measurable, and on $[0,\tau_1]$ we obtain

$$1 + (H^\lambda \cdot S)_t \geq \exp\left(\lambda h_0 S_{\tau_n\wedge t,\tau_{n+1}\wedge t} - \frac{\lambda^2}{2} b_0^2\right).$$

Observe also that $1 + (H^\lambda \cdot S)_{\tau_1} = 1 + f_0 S_{\tau_0,\tau_1}$ is bounded, because by assumption $h_0 S_{\tau_0,\tau_1}$ is bounded by the bounded random variable b_0.

Assume now that F_k has been defined for $k = 0, \dots, m-1$, that

$$1 + (H^\lambda \cdot S)_t \geq \exp\left(\lambda \sum_{n=0}^{\infty} h_n S_{\tau_n \wedge t, \tau_{n+1} \wedge t} - \frac{\lambda^2}{2} \sum_{n=0}^{N_t} b_n^2\right)$$

for all $t \in [0, \tau_m]$, and that $1 + (H^\lambda \cdot S)_{\tau_m}$ is bounded. We define $F_m := (1 + (H^\lambda \cdot S)_{\tau_m}) f_m$, which is $\mathcal{F}_{\tau_m}$-measurable and bounded. From (A.2) we obtain for $t \in [\tau_m, \tau_{m+1}]$

$$\begin{aligned} 1 + (H^\lambda \cdot S)_t &= 1 + (H^\lambda \cdot S)_{\tau_m} + (1 + (H^\lambda \cdot S)_{\tau_m}) f_m S_{\tau_m \wedge t, \tau_{m+1} \wedge t} \\ &\geq (1 + (H^\lambda \cdot S)_{\tau_m}) \exp\left(\lambda h_m S_{\tau_m \wedge t, \tau_{m+1} \wedge t} - \frac{\lambda^2}{2} b_m^2\right) \\ &\geq \exp\left(\lambda \sum_{n=0}^{\infty} h_n S_{\tau_n \wedge t, \tau_{n+1} \wedge t} - \frac{\lambda^2}{2} \sum_{n=0}^{N_t} b_n^2\right), \end{aligned}$$

where in the last step we used the induction hypothesis. From the first line of the previous equation we also obtain that $1 + (H^\lambda \cdot S)_{\tau_{m+1}}$ is bounded because $f_m S_{\tau_m, \tau_{m+1}}$ is bounded for the same reason that $f_0 S_{\tau_0, \tau_1}$ is bounded. □

A.2. Davie's criterion

It was already observed by Davie [Dav07] that in certain situations the rough path integral can be constructed as limit of Riemann sums and not just compensated Riemann sums. Davie shows that under suitable conditions, the usual Euler scheme (without "area compensation") converges to the solution of a given rough differential equation. But from there it is easily deduced that then also the rough path integral is given as limit of Riemann sums. Here we show that Davie's criterion implies our assumption (RIE) as required in Section 2.3.3.

Let $p \in (2,3)$ and let $\mathbb{S} = (S, A)$ be a $1/p$-Hölder continuous rough path, that is $|S_{s,t}| \lesssim |t-s|^{1/p}$ and $|A(s,t)| \lesssim |t-s|^{2/p}$. Write $\alpha := 1/p$ and let $\beta \in (1-\alpha, 2\alpha)$. Davie assumes that there exists $C > 0$ such that the area process A satisfies

$$\left| \sum_{j=k}^{\ell-1} A(jh, (j+1)h) \right| \leq C(\ell - k)^\beta h^{2\alpha}, \tag{A.3}$$

whenever $0 < k < \ell$ are integers and $h > 0$ such that $\ell h \leq T$. Under these conditions, Theorem 7.1 of [Dav07] implies that for $F \in C^\gamma$ with $\gamma > p$ and for $t_k^n = kT/n$, $n, k \in \mathbb{N}$, the Riemann sums

$$\sum_{k=0}^{n-1} F(S_{t_k^n}) S_{t_k^n \wedge t, t_{k+1}^n \wedge t}, \quad t \in [0, T],$$

converge uniformly to the rough path integral. But it can be easily deduced from (A.3) that the area process A is given as limit of non-anticipating Riemann sums along $(t^n)_n$.

Indeed, letting $h = T/n$,

$$\begin{aligned}
\Big|\int_0^t S_s \,\mathrm{d}S_s - \sum_{k=0}^{n-1} S_{t_k^n} S_{t_k^n\wedge t, t_{k+1}^n\wedge t}\Big| &= \Big|\sum_{k=0}^{n-1}\Big(\int_{t_k^n\wedge t}^{t_{k+1}^n\wedge t} S_s\,\mathrm{d}S_s - S_{t_k^n\wedge t} S_{t_k^n\wedge t, t_{k+1}^n\wedge t}\Big)\Big| \\
&= \Big|\sum_{k=0}^{n-1} A(t_k^n\wedge t, t_{k+1}^n\wedge t)\Big| \le \Big|\sum_{k=0}^{\lfloor t/h\rfloor -1} A_{kh,(k+1)h}\Big| + |A(\lfloor t/h\rfloor, t)| \\
&\lesssim C\lfloor t/h\rfloor^\beta h^{2\alpha} + h^{2\alpha}\|A\|_{2\alpha} \lesssim C t h^{2\alpha-\beta} + h^{2\alpha}\|A\|_{2\alpha}.
\end{aligned}$$

Since $\beta < 2\alpha$, the right hand side converges to 0 as n goes to ∞ (and thus h goes to 0). Furthermore, (A.3) implies the "uniformly bounded $p/2$-variation" condition (2.19):

$$\begin{aligned}
&\Big|(S^n\cdot S)_{t_k^n,t_\ell^n} - S_{t_k^n} S_{t_k^n,t_\ell^n}\Big| \\
&\quad\le \Big|\int_{t_k^n}^{t_\ell^n} S_s\,\mathrm{d}S_s - S_{t_k^n} S_{t_k^n,t_\ell^n}\Big| + \Big|\sum_{j=k}^{\ell-1}\Big(\int_{t_j^n}^{t_{j+1}^n} S_s\,\mathrm{d}S_s - S_{t_j^n} S_{t_j^n,t_{j+1}^n}\Big)\Big| \\
&\quad\le \|A\|_{2\alpha}|t_\ell^n - t_k^n|^{2\alpha} + \Big|\sum_{j=k}^{\ell-1} A_{t_k^n,t_{k+1}^n}\Big| \le \|A\|_{2\alpha}|t_\ell^n - t_k^n|^{2\alpha} + C(\ell-k)^\beta h^{2\alpha} \\
&\quad\le \|A\|_{2\alpha}|t_\ell^n - t_k^n|^{2\alpha} + C|t_\ell^n - t_k^n|^{2\alpha}.
\end{aligned}$$

A.3. Existence of local times via time change

A remarkable result in [Vov12] is a pathwise Dambis Dubins-Schwarz theorem, which allows to link results for the one-dimensional Wiener process to typical price paths. As already indicated at the end of Chapter 3, this opens another way to show the existence of local times, which we will briefly sketch here.

For that purpose let us briefly recall Vovk's outer measure and relate it to ours. For $\lambda \in (0,\infty)$ we define the set of processes

$$\mathcal{S}_\lambda := \Big\{\sum_{k=0}^\infty H^k : H^k \in \mathcal{H}_{\lambda_k}, \lambda_k > 0, \sum_{k=0}^\infty \lambda_k = \lambda\Big\}.$$

For every $G = \sum_{k\ge 0} H^k \in \mathcal{S}_\lambda$, all $\omega\in\Omega$, and all $t\in[0,\infty)$, the integral

$$(G\cdot S)_t(\omega) := \sum_{k\ge 0}(H^k\cdot S)_t(\omega) = \sum_{k\ge0}(\lambda_k + (H^k\cdot S)_t(\omega)) - \lambda$$

is well defined and takes values in $[-\lambda,\infty]$. Vovk then defines

$$\overline{Q}(A) := \inf\Big\{\lambda > 0 : \exists\, G\in\mathcal{S}_\lambda \text{ s.t. } \lambda + \liminf_{t\to\infty}(G\cdot S)_t(\omega) \ge \mathbf{1}_A(\omega)\,\forall\omega\in\Omega\Big\},\quad A\subseteq\Omega.$$

It is fairly easy to show that $\overline{P}(A) \le \overline{Q}(A)$ for all $A\subseteq\Omega$, see Section 2.1.3. In other words, all results which hold true outside of a $\overline{Q}$-null set are also true outside of a $\overline{P}$-null set.

To state Vovk's pathwise Dambis Dubins-Schwarz theorem, we need to define time-superinvariant sets.

Definition A.3.1. A continuous non-decreasing function $f\colon [0,\infty) \to [0,\infty)$ satisfying $f(0) = 0$ is said to be a *time change*. A subset $A \subseteq \Omega$ is called *time-superinvariant* if for each $\omega \in \Omega$ and each time change f it is true that $\omega \circ f \in A$ implies $\omega \in A$.

For $x \in \mathbb{R}$ we denote by μ_x the Wiener measure on $(\Omega, \mathcal{F})$ with $\mu_x(\omega(0) = x) = 1$.

Lemma A.3.2. *For every time-superinvariant set $A \subseteq \Omega$ satisfying $\omega(0) = x$ for all $\omega \in A$ and $\mu_x(A) = 0$, we have $\overline{P}(A) = 0$.*

Proof. Using Theorem 1 in [Vov12], we obtain $\overline{P}(A) \le \overline{Q}(A) = \mu_x(A) = 0$. □

First we investigate in the next lemma the behavior of local times under a time change. Recall that $\mathcal{L}_c$ is the set of those paths S which are in $\mathcal{L}_c(\pi^n)$ for the dyadic Lebesgue partition (π^n) constructed from S.

Lemma A.3.3. *Let $S \in \mathcal{Q}$ and assume that for all $t \ge 0$ the occupation measure*

$$\mu_t(A) = \int_0^t \mathbf{1}_A(S(s))\,\mathrm{d}S(s), \quad A \in \mathcal{B}(\mathbb{R}),$$

is absolutely continuous with density $2L_t(S)$. Let f be a time change. Then $S \circ f \in \mathcal{Q}$ and the occupation measure of $S \circ f$ is absolutely continuous with density $2L_{f(t)}(S)$ for all $t \ge 0$.

Proof. Recall that $\langle S \rangle$ is constructed along the dyadic Lebesgue partition, which yields $\langle S \circ f \rangle_t = \langle S \rangle_{f(t)}(\omega)$. The result then follows by considering the push forward of the occupation measure of S under f. □

With the previous lemma at hand we can reduce the existence and continuity of local times for typical price paths to the case of the Wiener process. For $p \ge 1$ let us define the events

$$\begin{aligned} A_c &:= \{\omega \in \Omega \,:\, S(\omega) \in \mathcal{L}_c\} \quad \text{and} \\ A_{c,p} &:= \{\omega \in A_c \,:\, u \mapsto L_t(S(\omega), u) \text{ has finite } p\text{-variation for all } t \in [0,\infty)\}. \end{aligned}$$

Theorem A.3.4. *Typical price paths are in $A_{c,p}$ for all $p > 2$.*

Proof. Define $\Omega_x := \{\omega \in \Omega \,:\, \omega(0) = x\}$ for $x \in \mathbb{R}$. Lemma A.3.2 and Lemma A.3.3 in combination with classical results for the Wiener process (see [KS88], Theorem 3.6.11 or [MP10], Theorem 6.19) show that typical price paths $\omega \in \Omega_x$ have an absolutely continuous occupation measure with jointly continuous density

$$\{2L_t(S,u), (t,u) \in [0,\infty) \times \mathbb{R}\}.$$

In [MP10], Theorem 6.19 it is also shown that $u \mapsto L_t(S,u)$ has finite p-variation for all $t \ge 0$, $p > 2$. It remains to show the uniform convergence of the discrete local times to L and to get rid of the restriction $\omega \in \Omega_x$.

Recall that $U_t(S,a,b)$ and $D_t(S,a,b)$ denote the number of up- respectively down-crossings of the interval $[a,b]$ completed by S up to time t. First observe that

$$\left|L_t^{\pi^n}(S,u) - 2^{-n} D_t(S, u - 2^{-n}, u)\right| \le 2^{-n} \tag{A.4}$$

for all $t \in [0, \infty)$ and $u \in \mathbb{D}^n$. For $u \in \mathbb{R}$ we define $\{u\}_n := \min\{k \in \mathbb{D}^n : k \geq u\}$ and by the triangle inequality we read

$$\begin{aligned}
&\sup_{(t,u)\in[0,T]\times\mathbb{R}} \big|L_t^{\pi^n}(S,u) - L_t(S,u)\big| \\
&\leq \sup_{(t,u)\in[0,T]\times\mathbb{R}} \big|L_t^{\pi^n}(S,u) - L_t^{\pi^n}(S,\{u\}_n)\big| \\
&\quad + \sup_{(t,u)\in[0,T]\times\mathbb{R}} \big|L_t^{\pi^n}(S,\{u\}_n) - L_t(S,\{u\}_n)\big| + \sup_{(t,u)\in[0,T]\times\mathbb{R}} \big|L_t(S,\{u\}_n) - L_t(S,u)\big|.
\end{aligned}$$

Now we separately deal with the three summands. The discrete Tanaka formula (3.7) yields

$$\big|L_t^{\pi^n}(S,u) - L_t^{\pi^n}(S,\{u\}_n)\big| \leq 3 \cdot 2^{-n}$$

for all $(t,u) \in [0,T] \times \mathbb{R}$.

For the second summand we remark that the event

$$\begin{aligned}
E := \Big\{\omega \in \Omega_x \,:\, \limsup_{n\to\infty} \sup_{(t,u)\in[0,T]\times\mathbb{R}} &\big|2^{-n} D_t(S, u - 2^{-n}, u) - L_t(S,u)\big| > 0 \\
&\text{for some } T \in [0,\infty)\Big\}
\end{aligned}$$

is time-superinvariant. Therefore, it suffices to combine Theorem 2 in [CLPT81] with (A.4) to obtain that the second summand converges to zero for typical price paths.

That the last ones goes to zero simply follows from the joint continuity of the compactly supported occupation density $L(S)$ in (t,u).

Finally, we indicate how to get rid of the assumption $\omega \in \Omega_x$ for some $x \in \mathbb{R}$. For $\varepsilon > 0$ it suffices to fix a sequence of simple trading strategies $(H^n) \subset \mathcal{H}_\varepsilon$ with

$$\liminf_{n\to\infty} (\varepsilon + (H^n \cdot S)_T(\omega)) \geq 1$$

for all $\omega \in \Omega_0$ for which the local time does not exist. Applying these simple trading strategies to $\omega - \omega(0)$ achieves the same aim but without the restriction $\omega(0) = 0$. $\square$

Remark A.3.5. *(i) For Theorem A.3.4, the dyadic points $\mathbb{D}^n$ in the definition of (π^n) can be replaced by any other increasing sequence of partition $(\mathcal{P}^n)$ of $\mathbb{R}$ such that* $\lim_{n\to\infty} |\mathcal{P}^n| = 0$. *See [CLPT81].*

(ii) While Theorem A.3.4 gives us the uniform convergence to a jointly continuous local time which is of finite p-variation in u, it does not give us the uniform boundedness in p-variation of the approximating sequence (L^{π^n}). Therefore, we can use Theorem A.3.4 only to prove an abstract version of Theorem 3.1.8, where the pathwise stochastic integral $\int_0^t g(S(s))\,\mathrm{d}S(s)$ *is defined by approximating g with smooth functions for which the Föllmer-Itô formula Theorem 3.1.2 holds (see [FZ06] for similar arguments in a semimartingale context). Since we want the Riemann sum interpretation of the pathwise integral, we need Theorem 2.2.5 to obtain the full strength of Theorem 3.1.8.*

A.4. Nonhomogeneous Besov spaces

In this part of the appendix we collect for the reader's convenience some results which allow to estimate the Besov norm of a function. For a general introduction to Littlewood-Paley theory and Besov spaces we recommend Triebel [Tri10] as well as Bahouri et al. [BCD11].

Lemma A.4.1. *[BCD11, Lem. 2.69] Let $\mathcal{A} \subset \mathbb{R}^d$ be an annulus, $\alpha \in \mathbb{R}$ and $p, q \in [1, \infty]$. Suppose that (f_j) is a sequence of smooth functions such that*

$$\operatorname{supp} \mathcal{F} f_j \subset 2^j \mathcal{A} \quad \textit{and} \quad \big\| \big(2^{\alpha j} \|f_j\|_{L^p}\big)_j \big\|_{\ell^q} < \infty.$$

Then $f := \sum_j f_j$ satisfies

$$f \in B^\alpha_{p,q}(\mathbb{R}^d) \quad \textit{and} \quad \|f\|_{\alpha,p,q} \lesssim \big\| \big(2^{\alpha j} \|f_j\|_{L^p}\big)_j \big\|_{\ell^q}.$$

Lemma A.4.2. *[BCD11, Lem. 2.84] Let $\mathcal{B} \subset \mathbb{R}^d$ be a ball, $\alpha \in \mathbb{R}$ and $p, q \in [1, \infty]$. Suppose that (f_j) is a sequence of smooth functions such that*

$$\operatorname{supp} \mathcal{F} f_j \subset 2^j \mathcal{B} \quad \textit{and} \quad \big\| \big(2^{\alpha j} \|f_j\|_{L^p}\big)_j \big\|_{\ell^q} < \infty.$$

Then $f := \sum_j f_j$ satisfies

$$f \in B^\alpha_{p,q}(\mathbb{R}^d) \quad \textit{and} \quad \|f\|_{\alpha,p,q} \lesssim \big\| \big(2^{\alpha j} \|f_j\|_{L^p}\big)_j \big\|_{\ell^q}.$$

Lemma A.4.3. *[BCD11, Prop. 2.79] Let $p, q \in [1, \infty]$, $\alpha < 0$ and f be a tempered distribution. Then, $f \in B^\alpha_{p,q}(\mathbb{R}^d)$ if and only if*

$$\big(2^{\alpha j} \|S_j f\|_{L^p}\big)_j \in \ell^q,$$

where we recall $S_j f := \sum_{k=-1}^{j-1} \Delta_k f$. Furthermore, there exists a constant $C > 0$ such that

$$C^{-|\alpha|+1} \|f\|_{\alpha,p,q} \leq \big\| \big(2^{\alpha j} \|S_j f\|_{L^p}\big)_j \big\|_{\ell^q} \leq C \Big(1 + \frac{1}{|\alpha|}\Big) \|f\|_{\alpha,p,q}.$$

A.5. Proof of Lemma 5.4.7: Lipschitz continuity

This subsection is devoted to the proof of Lemma 5.4.7. For $j = 1, 2$ let $u_0^j \in \mathbb{R}^d$ and $\vartheta^j_{\mathcal{T}} \in C^\infty_{\mathcal{T}}$ with derviative $\xi^j_{\mathcal{T}} = \mathrm{d}\vartheta^j_{\mathcal{T}}$. Denote by u^j, $j = 1, 2$, the solutions to corresponding Cauchy problems (5.26) and $\tilde{u}^j = \psi u^j$ for a weight function ψ satisfying Assumption 2. Then Lemma 5.4.7 is proven if we can show that

$$\|\tilde{u}^1 - \tilde{u}^2\|_{\alpha,p,q} \leq C \big(\|\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}}\|_{\alpha,p,q} + \|\pi(\vartheta^1_{\mathcal{T}}, \xi^1_{\mathcal{T}}) - \pi(\vartheta^2_{\mathcal{T}}, \xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q} \big),$$

for a constant C which does not depend on $\tilde{u}$. Roughly speaking, the verification of this bound follows the pattern of the proofs of Proposition 5.4.5 and Corollary 5.4.6. However, since Lemma 5.4.7 is essential for one of our main results, we shall present it here in full length.

Taking another weight function ψ_2 fulfilling Assumption 2 and keeping Remark 5.4.3 in mind, we obtain

$$\begin{aligned}\|\tilde u^1-\tilde u^2\|_{\alpha,p,q}&\lesssim\|\psi_2(\tilde u^1-\tilde u^2)\|_{\alpha,p,q}\\&\lesssim(\mathcal T^2\vee 1)\big(|u^1(0)-u^2(0)|+\|\,\mathrm d(\tilde u^1-\tilde u^2)\|_{\alpha-1,p,q}\big)\\&\le(\mathcal T^2\vee 1)\big(|u^1(0)-u^2(0)|+\|\,\mathrm d(T_{F(\tilde u^1)}\vartheta^1_{\mathcal T}-T_{F(\tilde u^2)}\vartheta^2_{\mathcal T})\|_{\alpha-1,p,q}\\&\quad+\|\,\mathrm d(u^{\#,1}-u^{\#,2})\|_{\alpha-1,p,q}\big),\end{aligned}\tag{A.5}$$

where Lemma 5.1.2 is used in the second line and the paracontrolled ansatz $\tilde u^j=T_{F(\tilde u^j)}\vartheta^j_{\mathcal T}+u^{\#,j}$ in the third one. Let us continue by further estimating the term $\mathrm d(T_{F(\tilde u^1)}\vartheta^1_{\mathcal T}-T_{F(\tilde u^2)}\vartheta^2_{\mathcal T})$. Applying the Leibniz rule and the triangle inequality leads to

$$\begin{aligned}&\|\,\mathrm d(T_{F(\tilde u^1)}\vartheta^1_{\mathcal T}-T_{F(\tilde u^2)}\vartheta^2_{\mathcal T})\|_{\alpha-1,p,q}\\&\quad\le\|T_{\mathrm dF(\tilde u^1)}\vartheta^1_{\mathcal T}-T_{\mathrm dF(\tilde u^2)}\vartheta^2_{\mathcal T}\|_{\alpha-1,p,q}+\|T_{F(\tilde u^1)}\xi^1_{\mathcal T}-T_{F(\tilde u^2)}\xi^2_{\mathcal T}\|_{\alpha-1,p,q}\\&\quad\le\|T_{\mathrm dF(\tilde u^1)}(\vartheta^1_{\mathcal T}-\vartheta^2_{\mathcal T})\|_{\alpha-1,p,q}+\|T_{\mathrm dF(\tilde u^1)-\,\mathrm dF(\tilde u^2)}\vartheta^2_{\mathcal T}\|_{\alpha-1,p,q}\\&\qquad+\|T_{F(\tilde u^1)}(\xi^1_{\mathcal T}-\xi^2_{\mathcal T})\|_{\alpha-1,p,q}+\|T_{F(\tilde u^1)-F(\tilde u^2)}\xi^2_{\mathcal T}\|_{\alpha-1,p,q}.\end{aligned}$$

Based on Lemma 5.1.1, Besov embeddings, the lifting property of Besov spaces [Tri10, Thm. 2.3.8], (5.10) and (5.25), one has

$$\begin{aligned}&\|\,\mathrm d(T_{F(\tilde u^1)}\vartheta^1_{\mathcal T}-T_{F(\tilde u^2)}\vartheta^2_{\mathcal T})\|_{\alpha-1,p,q}\\&\quad\lesssim\|\,\mathrm dF(\tilde u^1)\|_{\alpha-1,p,q}\|\vartheta^1_{\mathcal T}-\vartheta^2_{\mathcal T}\|_{0,\infty,\infty}+\|\,\mathrm dF(\tilde u^1)-\,\mathrm dF(\tilde u^2)\|_{\alpha-1,p,q}\|\vartheta^2_{\mathcal T}\|_{0,\infty,\infty}\\&\qquad+\|F\|_\infty\|\xi^1_{\mathcal T}-\xi^2_{\mathcal T}\|_{\alpha-1,p,q}+\|F(\tilde u^1)-F(\tilde u^2)\|_\infty\|\xi^2_{\mathcal T}\|_{\alpha-1,p,q}\\&\quad\lesssim\|F(\tilde u^1)\|_{\alpha,p,q}\|\vartheta^1_{\mathcal T}-\vartheta^2_{\mathcal T}\|_{\alpha,p,q}+\|F(\tilde u^1)-F(\tilde u^2)\|_{\alpha,p,q}\|\vartheta^2_{\mathcal T}\|_{\alpha,p,q}\\&\qquad+\|F\|_\infty\|\xi^1_{\mathcal T}-\xi^2_{\mathcal T}\|_{\alpha-1,p,q}+\|F'\|_\infty\|\tilde u^1-\tilde u^2\|_\infty\|\xi^2_{\mathcal T}\|_{\alpha-1,p,q}\\&\quad\lesssim\|F\|_{C^1}\|\tilde u^1\|_{\alpha,p,q}\|\vartheta^1_{\mathcal T}-\vartheta^2_{\mathcal T}\|_{\alpha,p,q}+\|F'\|_\infty\|\vartheta^2_{\mathcal T}\|_{\alpha,p,q}\|\tilde u^1-\tilde u^2\|_{\alpha,p,q}\\&\qquad+\|F\|_\infty\|\xi^1_{\mathcal T}-\xi^2_{\mathcal T}\|_{\alpha-1,p,q}+\|F'\|_\infty\|\xi^2_{\mathcal T}\|_{\alpha-1,p,q}\|\tilde u^1-\tilde u^2\|_{\alpha,p,q}\\&\quad\lesssim\|F\|_{C^1}\big(1+\|\tilde u^1\|_{\alpha,p,q}+\|\xi^2_{\mathcal T}\|_{\alpha-1,p,q}+\|\vartheta^2_{\mathcal T}\|_{\alpha,p,q}\big)\\&\qquad\times\big(\|\xi^1_{\mathcal T}-\xi^2_{\mathcal T}\|_{\alpha-1,p,q}+\|\vartheta^1_{\mathcal T}-\vartheta^2_{\mathcal T}\|_{\alpha,p,q}+\|\tilde u^1-\tilde u^2\|_{\alpha,p,q}\big).\end{aligned}\tag{A.6}$$

It remains to consider the difference of derivatives $\mathrm d\tilde u^{\#,j}$, which can be decomposed (cf. (5.33)) into

$$\mathrm d\tilde u^{\#,j}=\pi(F(\tilde u^j),\xi^j_{\mathcal T})+T_{\xi^j_{\mathcal T}}(F(\tilde u^j))-T_{\mathrm dF(\tilde u^j)}\vartheta^j_{\mathcal T}+\frac{\psi'}{\psi}\tilde u^j\quad\text{for } j=1,2.$$

Applying Proposition 5.3.1, we can rewrite the resonant term, differently than in the proof of Proposition 5.4.5, as

$$\pi(F(\tilde u^j),\xi^j_{\mathcal T})=F'(\tilde u^j)\pi(\tilde u^j,\xi^j_{\mathcal T})+\Pi_F(\tilde u^j,\xi^j_{\mathcal T})\tag{A.7}$$

and, taking the ansatz $\tilde u^j=T_{F(\tilde u^j)}\vartheta^j_{\mathcal T}+u^{\#,j}$ into account and applying the commutator Lemma 5.3.4, we have

$$\begin{aligned}\pi(\tilde u^j,\xi^j_{\mathcal T})&=\pi(T_{F(\tilde u^j)}\vartheta^j_{\mathcal T},\xi^j_{\mathcal T})+\pi(u^{\#,j},\xi^j_{\mathcal T})\\&=F(\tilde u^j)\pi(\vartheta^j_{\mathcal T},\xi^j_{\mathcal T})+\Gamma(F(\tilde u^j),\vartheta^j_{\mathcal T},\xi^j_{\mathcal T})+\pi(u^{\#,j},\xi^j_{\mathcal T}).\end{aligned}$$

Therefore, we decompose $\mathrm{d}u^{\#,j}$ into the following seven terms

$$\begin{aligned}\mathrm{d}\tilde{u}^{\#,j} &= F'(\tilde{u}^j)F(\tilde{u}^j)\pi(\vartheta^j_{\mathcal{T}},\xi^j_{\mathcal{T}}) + F'(\tilde{u}^j)\Gamma(F(\tilde{u}^j),\vartheta^j_{\mathcal{T}},\xi^j_{\mathcal{T}}) + F'(\tilde{u}^j)\pi(u^{\#,j},\xi^j_{\mathcal{T}})\\ &\quad + \Pi_F(\tilde{u}^j,\xi^j_{\mathcal{T}}) + T_{\xi^j_{\mathcal{T}}}(F(\tilde{u}^j)) - T_{\mathrm{d}F(\tilde{u}^j)}\vartheta^j_{\mathcal{T}} + \frac{\psi'}{\psi}\tilde{u}^j\\ &=: D^j_1 + \cdots + D^j_7.\end{aligned}$$

Let us tackle the differences of these seven terms: The first term is estimated as follows

$$\begin{aligned}\|D^1_1 - D^2_1\|_{2\alpha-1,p/2,q} &= \|F'(\tilde{u}^1)F(\tilde{u}^1)\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}}) - F'(\tilde{u}^2)F(\tilde{u}^2)\pi(\vartheta^2_{\mathcal{T}},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ &\lesssim \|F'(\tilde{u}^1)F(\tilde{u}^1) - F'(\tilde{u}^2)F(\tilde{u}^2)\|_{\alpha,p,q}\|\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ &\quad + \|F'(\tilde{u}^2)F(\tilde{u}^2)\|_{\alpha,p,q}\|\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}}) - \pi(\vartheta^2_{\mathcal{T}},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ &\lesssim \|\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\big(\|(F'(\tilde{u}^1) - F'(\tilde{u}^2))F(\tilde{u}^1)\|_{\alpha,p,q}\\ &\quad + \|F'(\tilde{u}^2)(F(\tilde{u}^1) - F(\tilde{u}^2))\|_{\alpha,p,q}\big)\\ &\quad + \|F\|^2_{C^2}\|\tilde{u}^2\|_{\alpha,p,q}\|\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}}) - \pi(\vartheta^2_{\mathcal{T}},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ &\lesssim \|F\|^2_{C^2}\Big(\|\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\big(\|\tilde{u}^1\|_{\alpha,p,q} + \|\tilde{u}^2\|_{\alpha,p,q}\big)\|\tilde{u}^1 - \tilde{u}^2\|_{\alpha,p,q}\\ &\quad + \|\tilde{u}^2\|_{\alpha,p,q}\|\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}}) - \pi(\vartheta^2_{\mathcal{T}},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\Big),\end{aligned}$$

where we refer to (5.6), (5.10), (5.25) and (5.32) for explanations to the above estimates. Applying (5.32), Lemma 5.3.4 and Besov embeddings, we see for the next term that

$$\begin{aligned}\|D^1_2 - D^2_2\|_{2\alpha-1,p/2,q} &= \|F'(\tilde{u}^1)\Gamma(F(\tilde{u}^1),\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}}) - F'(\tilde{u}^2)\Gamma(F(\tilde{u}^2),\vartheta^2_{\mathcal{T}},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ &\lesssim \|F'\|_\infty\big(\|\Gamma(F(\tilde{u}^1) - F(\tilde{u}^2),\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}})\|_{3\alpha-1,p/3,q}\\ &\quad + \|\Gamma(F(\tilde{u}^2),\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}},\xi^1_{\mathcal{T}})\|_{3\alpha-1,p/3,q} + \|\Gamma(F(\tilde{u}^2),\vartheta^2_{\mathcal{T}},\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}})\|_{3\alpha-1,p/3,q}\big)\\ &\quad + \|F'(\tilde{u}^1) - F'(\tilde{u}^2)\|_\infty\|\Gamma(F(\tilde{u}^2),\vartheta^2_{\mathcal{T}},\xi^2_{\mathcal{T}})\|_{3\alpha-1,p/3,q}\\ &\lesssim \|F\|^2_{C^1}\Big(\|\vartheta^1_{\mathcal{T}}\|_{\alpha,p,q}\|\xi^1_{\mathcal{T}}\|_{\alpha-1,p,q}\|\tilde{u}^1 - \tilde{u}^2\|_{\alpha,p,q}\\ &\quad + \|\xi^1_{\mathcal{T}}\|_{\alpha-1,p,q}\|\tilde{u}^2\|_{\alpha,p,q}\|\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}}\|_{\alpha,p,q} + \|\tilde{u}^2\|_{\alpha,p,q}\|\vartheta^2_{\mathcal{T}}\|_{\alpha,p,q}\|\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}}\|_{\alpha-1,p,q}\Big)\\ &\quad + \|F''\|_\infty\|F\|_{C^1}\|\tilde{u}^2\|_{\alpha,p,q}\|\vartheta^2_{\mathcal{T}}\|_{\alpha,p,q}\|\xi^2_{\mathcal{T}}\|_{\alpha-1,p,q}\|\tilde{u}^1 - \tilde{u}^2\|_{\alpha,p,q}.\end{aligned}$$

For the third term, again due to (5.32) as well as Lemma 5.1.1 and Besov embeddings, we obtain

$$\begin{aligned}\|D^1_3 - D^2_3\|_{2\alpha-1,p/2,q} &= \|F'(\tilde{u}^1)\pi(u^{\#,1},\xi^1_{\mathcal{T}}) - F'(\tilde{u}^2)\pi(u^{\#,2},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ &\lesssim \|F'(\tilde{u}^1)\pi(u^{\#,1} - u^{\#,2},\xi^1_{\mathcal{T}})\|_{2\alpha-1,p/2,q} + \|F'(\tilde{u}^1)\pi(u^{\#,2},\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ &\quad + \|(F'(\tilde{u}^1) - F'(\tilde{u}^2))\pi(u^{\#,2},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\\ &\lesssim \|F'(\tilde{u}^1)\|_\infty\|\pi(u^{\#,1} - u^{\#,2},\xi^1_{\mathcal{T}})\|_{3\alpha-1,p/3,q} + \|F'(\tilde{u}^1)\|_\infty\\ &\quad \times\|\pi(u^{\#,2},\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}})\|_{3\alpha-1,p/3,q} + \|\pi(u^{\#,2},\xi^2_{\mathcal{T}})\|_{3\alpha-1,p/3,q}\|F'(\tilde{u}^1) - F'(\tilde{u}^2)\|_\infty\\ &\lesssim \|F\|_{C^1}\|\xi^1_{\mathcal{T}}\|_{\alpha-1,p,q}\|u^{\#,1} - u^{\#,2}\|_{2\alpha,p/2,q} + \|F\|_{C^1}\|u^{\#,2}\|_{2\alpha,p/2,q}\|\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}}\|_{\alpha-1,p,q}\\ &\quad + \|F''\|_\infty\|u^{\#,2}\|_{2\alpha,p/2,q}\|\xi^2_{\mathcal{T}}\|_{\alpha-1,p,q}\|\tilde{u}^1 - \tilde{u}^2\|_{\alpha,p,q}.\end{aligned}$$

Proposition 5.3.1 and the embedding $B^{3\alpha-1}_{p/3,\infty} \subset B^{2\alpha-1}_{p/2,q}$ yield for the fourth term

$$\begin{aligned}\|D^1_4 - D^2_4\|_{2\alpha-1,p/2,q} &= \|\Pi_F(\tilde u^1, \xi^1_{\mathcal{T}}) - \Pi_F(\tilde u^2, \xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q} \\ &\lesssim \|\Pi_F(\tilde u^1, \xi^1_{\mathcal{T}}) - \Pi_F(\tilde u^2, \xi^2_{\mathcal{T}})\|_{3\alpha-1,p/3,\infty} \\ &\lesssim \|F\|_{C^3} C(\tilde u^1, \tilde u^2, \xi^1_{\mathcal{T}}, \xi^2_{\mathcal{T}})\Big(\|\tilde u^1 - \tilde u^2\|_{\alpha,p,q} + \|\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}}\|_{\alpha-1,p,q}\Big),\end{aligned}$$

where the constant $C(\tilde u^1, \tilde u^2, \xi^1_{\mathcal{T}}, \xi^2_{\mathcal{T}})$ is given in Proposition 5.3.1. The fifth term can be bounded by

$$\begin{aligned}&\|D^1_5 - D^2_5\|_{2\alpha-1,p/2,q} = \|T_{\xi^1_{\mathcal{T}}}(F(\tilde u^1)) - T_{\xi^2_{\mathcal{T}}}(F(\tilde u^2))\|_{2\alpha-1,p/2,q} \\ &\lesssim \|T_{\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}}}(F(\tilde u^1))\|_{2\alpha-1,p/2,q} + \|T_{\xi^2_{\mathcal{T}}}(F(\tilde u^1) - F(\tilde u^2))\|_{2\alpha-1,p/2,q} \\ &\lesssim \|F\|_{C^1}\|\tilde u^1\|_{\alpha,p,2q}\|\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}}\|_{\alpha-1,p,2q} + \|\xi^2_{\mathcal{T}}\|_{\alpha-1,p,2q}\|F(\tilde u^1) - F(\tilde u^2)\|_{\alpha,p,2q} \\ &\lesssim \|F\|_{C^1}\|\tilde u^1\|_{\alpha,p,q}\|\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}}\|_{\alpha-1,p,q} + \|F'\|_\infty\|\xi^2_{\mathcal{T}}\|_{\alpha-1,p,q}\|\tilde u^1 - \tilde u^2\|_{\alpha,p,q}\end{aligned}$$

because of Lemma 5.1.1 and (5.25). For the sixth term, the lifting property [Tri10, Thm. 2.3.8], an analog to (5.25) and (5.10) yield

$$\begin{aligned}&\|D^1_6 - D^2_6\|_{2\alpha-1,p/2,q} = \|T_{\mathrm{d}F(\tilde u^1)}\vartheta^1_{\mathcal{T}} - T_{\mathrm{d}F(\tilde u^2)}\vartheta^2_{\mathcal{T}}\|_{2\alpha-1,p/2,q} \\ &\lesssim \|T_{\mathrm{d}F(\tilde u^1) - \mathrm{d}F(\tilde u^2)}\vartheta^1_{\mathcal{T}}\|_{2\alpha-1,p/2,q} + \|T_{\mathrm{d}F(\tilde u^2)}(\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q} \\ &\lesssim \|\mathrm{d}F(\tilde u^1) - \mathrm{d}F(\tilde u^2)\|_{\alpha-1,p,q}\|\vartheta^1_{\mathcal{T}}\|_{\alpha,p,q} + \|\mathrm{d}F(\tilde u^2)\|_{\alpha-1,p,q}\|\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}}\|_{\alpha,p,q} \\ &\lesssim \|F(\tilde u^1) - F(\tilde u^2)\|_{\alpha,p,q}\|\vartheta^1_{\mathcal{T}}\|_{\alpha,p,q} + \|F(\tilde u^2)\|_{\alpha,p,q}\|\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}}\|_{\alpha,p,q} \\ &\lesssim \|F'\|_\infty\|\vartheta^1_{\mathcal{T}}\|_{\alpha,p,q}\|\tilde u^1 - \tilde u^2\|_{\alpha,p,q} + \|F\|_{C^1}\|\tilde u^2\|_{\alpha,p,q}\|\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}}\|_{\alpha,p,q}.\end{aligned}$$

Since $2\alpha - 1 < 0$, the last difference $D^1_7 - D^2_7$ can be easily estimated by

$$\begin{aligned}\|\frac{\psi'}{\psi}(\tilde u^1 - \tilde u^2)\|_{2\alpha-1,p/2,q} &\lesssim \|\frac{\psi'}{\psi}(\tilde u^1 - \tilde u^2)\|_{L^{p/2}} \\ &\le \|\frac{\psi'}{\psi}\|_\infty\|\tilde u^1 - \tilde u^2\|_{L^{p/2}} \lesssim (\mathcal{T} \vee 1)\|\frac{\psi'}{\psi}\|_\infty\|\tilde u^1 - \tilde u^2\|_{\alpha,p,q}.\end{aligned}$$

Defining the constants

$$\begin{aligned}&\tilde C_{\tilde u, u^\#} := 1 + \sum_{i=1,2}\big(\|\tilde u^j\|_{\alpha,p,q} + \|\tilde u^j\|^2_{\alpha,p,q} + \|u^{\#,j}\|_{2\alpha,p/2,q}\big), \\ &C_{\xi^j,\vartheta^j} := \|\vartheta^j_{\mathcal{T}}\|_{\alpha,p,q} + \|\vartheta^j_{\mathcal{T}}\|^2_{\alpha,p,q} + \|\pi(\vartheta^j_{\mathcal{T}}, \xi^j_{\mathcal{T}})\|_{2\alpha-1,p/2,q}, \quad j = 1,2, \\ &\tilde C_{\xi,\vartheta} := 1 + C_{\xi^1,\vartheta^1} + C_{\xi^2,\vartheta^2},\end{aligned}$$

we altogether obtain

$$\begin{aligned}&\|\mathrm{d}u^{\#,1} - \mathrm{d}u^{\#,2}\|_{2\alpha-1,p/2,q} \\ &\quad\lesssim \tilde C_{\xi,\vartheta}\tilde C_{\tilde u,u^\#}\big(\|F\|_{C^3} + \|F\|^2_{C^2}\big) \\ &\qquad\times\Big(\|\tilde u^1 - \tilde u^2\|_{\alpha,p,q} + \|u^{\#,1} - u^{\#,2}\|_{2\alpha,p/2,q} + \|\xi^1_{\mathcal{T}} - \xi^2_{\mathcal{T}}\|_{\alpha-1,p,q} \\ &\qquad\qquad + \|\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}}\|_{\alpha,p,q} + \|\pi(\vartheta^1_{\mathcal{T}}, \xi^1_{\mathcal{T}}) - \pi(\vartheta^2_{\mathcal{T}}, \xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\Big) \\ &\qquad + (\mathcal{T} \vee 1)\|\frac{\psi'}{\psi}\|_\infty\|\tilde u^1 - \tilde u^2\|_{\alpha,p,q}.\end{aligned}$$

The factor $\tilde{C}_{\tilde{u},u^\#}$ is (locally) bounded since $\|\tilde{u}^1\|_{\alpha,p,q}$ and $\|\tilde{u}^2\|_{\alpha,p,q}$ can be bounded by Corollary 5.4.6 and $\|u^{\#,j}\|_{2\alpha,p/2,q}$, for $j=1,2$, can be bounded analogously to (5.34) and (5.35) by

$$\begin{aligned}\|u^{\#,j}\|_{2\alpha,p/2,q} &\lesssim (\mathcal{T}\vee 1)\Big((\|F\|_\infty\|\vartheta^j_{\mathcal{T}}\|_{L^p}+\|\tilde{u}^j\|_{L^p})\\ &\quad + C_{\xi^j,\vartheta^j}(\|F\|_{C^2}\vee\|F\|^2_{C^2})(\|\tilde{u}^j\|_{\alpha,p,q}+\|F\|_\infty\|\vartheta^j_{\mathcal{T}}\|_{\alpha,p,q})+\|\frac{\psi'}{\psi}\|_\infty\|\tilde{u}^j\|_{\alpha,p,q}\Big)\\ &\lesssim (\mathcal{T}\vee 1)(1+(\|F\|_{C^2}\vee\|F\|^3_{C^2})(1+\|\vartheta^j_{\mathcal{T}}\|)C_{\xi^j,\vartheta^j}+\|\frac{\psi'}{\psi}\|_\infty)(1+\|\tilde{u}^j\|_{\alpha,p,q}).\end{aligned}$$

Relying on the lifting property of Besov spaces together with the definition of $u^\#$, $\|\tilde{u}^1-\tilde{u}^2\|_{L^{p/2}}\lesssim(\mathcal{T}\vee 1)\|\tilde{u}^1-\tilde{u}^2\|_{L^p}$ and the compact support of $\vartheta^j_{\mathcal{T}}$, we have

$$\begin{aligned}&\|u^{\#,1}-u^{\#,2}\|_{2\alpha,p/2,q}\\ &\quad\lesssim \|u^{\#,1}-u^{\#,2}\|_{L^{p/2}}+\|\,\mathrm{d}u^{\#,1}-\,\mathrm{d}u^{\#,2}\|_{2\alpha-1,p/2,q}\\ &\quad\le \|T_{F(\tilde{u}^1)}\vartheta^1_{\mathcal{T}}-T_{F(\tilde{u}^2)}\vartheta^2_{\mathcal{T}}\|_{L^{p/2}}+\|\tilde{u}^1-\tilde{u}^2\|_{L^{p/2}}+\|\,\mathrm{d}u^{\#,1}-\,\mathrm{d}u^{\#,2}\|_{2\alpha-1,p/2,q}\\ &\quad\le \|T_{F(\tilde{u}^1)-F(\tilde{u}^2)}\vartheta^1_{\mathcal{T}}\|_{L^{p/2}}+\|T_{F(\tilde{u}^2)}(\vartheta^1_{\mathcal{T}}-\vartheta^2_{\mathcal{T}})\|_{L^{p/2}}+\|\tilde{u}^1-\tilde{u}^2\|_{L^{p/2}}\\ &\qquad+\|\,\mathrm{d}u^{\#,1}-\,\mathrm{d}u^{\#,2}\|_{2\alpha-1,p/2,q}\\ &\quad\lesssim(\mathcal{T}\vee 1)(\|F(\tilde{u}^1)-F(\tilde{u}^2)\|_\infty\|\vartheta^1_{\mathcal{T}}\|_{L^p}+\|F\|_\infty\|\vartheta^1_{\mathcal{T}}-\vartheta^2_{\mathcal{T}}\|_{L^p}+\|\tilde{u}^1-\tilde{u}^2\|_{L^p})\\ &\qquad+\|\,\mathrm{d}u^{\#,1}-\,\mathrm{d}u^{\#,2}\|_{2\alpha-1,p/2,q}\\ &\quad\lesssim(\mathcal{T}\vee 1)(\|F'\|_\infty\|\vartheta^1_{\mathcal{T}}\|_{\alpha,p,q}\|\tilde{u}^1-\tilde{u}^2\|_{\alpha,p,q}+\|F\|_\infty\|\vartheta^1_{\mathcal{T}}-\vartheta^2_{\mathcal{T}}\|_{\alpha,p,q}\\ &\qquad+\|\tilde{u}^1-\tilde{u}^2\|_{\alpha,p,q})+\|\,\mathrm{d}u^{\#,1}-\,\mathrm{d}u^{\#,2}\|_{2\alpha-1,p/2,q}.\end{aligned}$$

Therefore, if $\|F\|_{C^3}+\|F\|^2_{C^2}$ is sufficiently small, depending on $\tilde{C}_{\xi,\vartheta}$, $\tilde{C}_{\tilde{u},u^\#}$ and $\mathcal{T}$, then

$$\begin{aligned}&\|\,\mathrm{d}u^{\#,1}-\,\mathrm{d}u^{\#,2}\|_{2\alpha-1,p/2,q}\\ &\quad\lesssim(1+\|\vartheta^1_{\mathcal{T}}\|_{\alpha,p,q})\tilde{C}_{\xi,\vartheta}\tilde{C}_{\tilde{u},u^\#}(\mathcal{T}\vee 1)(\|F\|_{C^3}+\|F\|^3_{C^2})\\ &\qquad\times\Big(\|\tilde{u}^1-\tilde{u}^2\|_{\alpha,p,q}+\|\xi^1_{\mathcal{T}}-\xi^2_{\mathcal{T}}\|_{\alpha-1,p,q}+\|\vartheta^1_{\mathcal{T}}-\vartheta^2_{\mathcal{T}}\|_{\alpha,p,q}\\ &\qquad+\|\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}})-\pi(\vartheta^2_{\mathcal{T}},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\Big)+(\mathcal{T}\vee 1)\|\frac{\psi'}{\psi}\|_\infty\|\tilde{u}^1-\tilde{u}^2\|_{\alpha,p,q}.\end{aligned}$$

Plugging this estimate and (A.6) into (A.5), we obtain

$$\begin{aligned}&\|\tilde{u}^1-\tilde{u}^2\|_{\alpha,p,q}\\ &\quad\lesssim(\mathcal{T}^2\vee 1)|u^1(0)-u^2(0)|+(1+\|\vartheta^1_{\mathcal{T}}\|_{\alpha,p,q})\tilde{C}_{\xi,\vartheta}\tilde{C}_{\tilde{u},u^\#}(\mathcal{T}\vee 1)(\|F\|_{C^3}+\|F\|^3_{C^2})\\ &\qquad\times\Big(\|\tilde{u}^1-\tilde{u}^2\|_{\alpha,p,q}+\|\xi^1_{\mathcal{T}}-\xi^2_{\mathcal{T}}\|_{\alpha-1,p,q}+\|\vartheta^1_{\mathcal{T}}-\vartheta^2_{\mathcal{T}}\|_{\alpha,p,q}\\ &\qquad+\|\pi(\vartheta^1_{\mathcal{T}},\xi^1_{\mathcal{T}})-\pi(\vartheta^2_{\mathcal{T}},\xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\Big)+(\mathcal{T}^2\vee 1)\|\frac{\psi'}{\psi}\|_\infty\|\tilde{u}^1-\tilde{u}^2\|_{\alpha,p,q}.\end{aligned}$$

For a possibly smaller $\|F\|_{C^3} + \|F\|_{C^2}^3$ and a sufficiently small $\|\frac{\psi'}{\psi}\|_\infty$, we conclude

$$\begin{aligned}\|\tilde{u}^1 - \tilde{u}^2\|_{\alpha,p,q} \lesssim{}& (\mathcal{T}^2 \vee 1)|u^1(0) - u^2(0)| + (1 + \|\vartheta^1_{\mathcal{T}}\|_{\alpha,p,q})\\ &\tilde{C}_{\xi,\vartheta}\tilde{C}_{\tilde{u},u^\#}(\mathcal{T} \vee 1)(\|F\|_{C^3} + \|F\|_{C^2}^3)\\ &\times \Big(\|\vartheta^1_{\mathcal{T}} - \vartheta^2_{\mathcal{T}}\|_{\alpha,p,q} + \|\pi(\vartheta^1_{\mathcal{T}}, \xi^1_{\mathcal{T}}) - \pi(\vartheta^2_{\mathcal{T}}, \xi^2_{\mathcal{T}})\|_{2\alpha-1,p/2,q}\Big).\end{aligned}$$

Finally, note again that $\tilde{C}_{\tilde{u},u^\#}$ is (locally) bounded by Corollary 5.4.6. □

A.6. BMO - Processes and their properties

In the following, let $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a complete filtered probability space such that the filtration satisfies the usual hypotheses. Assume furthermore that there exists a d-dimensional Brownian motion W on $[0,T]$, which is progressive w.r.t. $(\mathcal{F}_t)_{t\in[0,T]}$, independent of $\mathcal{F}_0$ and such that $\mathcal{F}_t = \sigma(\mathcal{F}_0, \mathcal{F}^W_t)$, where $\mathcal{F}^W$ is the natural filtration generated by W and $\mathcal{F}_0$ contains all null sets.

For a probability measure $\mathbb{Q}$ and any $q > 0$ and $m \in \mathbb{N}$ define $\mathcal{H}^q(\mathbb{R}^m, \mathbb{Q})$ as the space of all progressive processes $(Z_t)_{t\in[0,T]}$ with values in $\mathbb{R}^m$ normed by

$$\|Z\|_{\mathcal{H}^q} := \mathbb{E}_{\mathbb{Q}}\left[\left(\int_0^T |Z_s|^2 \,\mathrm{d}s\right)^{\frac{q}{2}}\right]^{\frac{1}{q}} < \infty.$$

Definition A.6.1. Let $\mathbb{Q} \sim \mathbb{P}$ be an equivalent probability measure and define

$$\begin{aligned}BMO(\mathbb{Q}) := \Bigg\{&Z\colon [0,T]\times\Omega \,:\, Z \text{ is progressively measurable and vector-valued s.t.}\\ &\exists C \geq 0\,\forall t \in [0,T] : \mathbb{E}_{\mathbb{Q}}\left[\int_t^T |Z_s|^2 \,\mathrm{d}s \middle| \mathcal{F}_t\right] \leq C \text{ a.s.}\Bigg\}.\end{aligned}$$

By vector-valued we mean that Z assumes values in some normed vector space.

The smallest constant C such that the above bound holds is denoted by $\check{C} := \|Z\|^2_{BMO(\mathbb{Q})}$. For processes $Z \notin BMO(\mathbb{Q})$ we define $\|Z\|_{BMO(\mathbb{Q})} := \infty$.

Furthermore, we call a martingale M a BMO-martingale if

$$M_t = M_0 + \int_0^t Z_s \,\mathrm{d}W_s =: M_0 + (Z \bullet W)_t$$

with some $\mathbb{R}^{1\times d}$-valued $Z \in BMO(\mathbb{P})$. Also, if a progressive process Z is only defined on a subinterval of $[0,T]$, the statement $Z \in BMO(\mathbb{Q})$ means that its natural extension to $[0,T]$, obtained by setting it equal to 0 everywhere outside its initial domain, is in $BMO(\mathbb{Q})$.

Theorem A.6.2 (Theorem 2.3. in [Kaz94])**.** *Let $\mu \in BMO(\mathbb{P})$ be $\mathbb{R}^{1\times d}$-valued, then*

$$\mathbb{Q}^\mu := \mathcal{E}(\mu \bullet W)_T \cdot \mathbb{P}$$

is a probability measure.

Lemma A.6.3. *For a probability measure $\mathbb{Q} \sim \mathbb{P}$ let $Z \in BMO(\mathbb{Q})$ be $\mathbb{R}^m$-valued. Then $Z \in \mathcal{H}^{2n}(\mathbb{R}^m, \mathbb{Q})$ for all $n \in \mathbb{N}$ and*

$$\|Z\|_{\mathcal{H}^{2n}(\mathbb{R}^m,\mathbb{Q})} \leq \sqrt[2n]{n!}\,\|Z\|_{BMO(\mathbb{Q})}.$$

Proof. Define $A_t := \int_0^t |Z_s|^2 \,\mathrm{d}s$ for all $t \in [0,T]$. A is progressive, non-decreasing, starts at 0 and satisfies $\mathbb{E}_{\mathbb{Q}}[A_T - A_t|\mathcal{F}_t] \leq \|Z\|^2_{BMO(\mathbb{Q})}$ for all $t \in [0,T]$. Therefore, using energy inequalities we have

$$\mathbb{E}_{\mathbb{Q}}[(A_T)^n] \leq n! \left(\|Z\|^2_{BMO(\mathbb{Q})}\right)^n,$$

which implies the assertion. □

Lemma A.6.4. *For all $K > 0$ there exists a constant $C > 0$, which is increasing in K, such that*

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_t^T |Z_s|\,\mathrm{d}s\right)\middle|\mathcal{F}_t\right] \leq C \quad \textit{a.s. for all } t \in [0,T],$$

all probability measures $\mathbb{Q} \sim \mathbb{P}$ and all $Z \in BMO(\mathbb{Q})$ such that $\|Z\|_{BMO(\mathbb{Q})} \leq K$.

Proof. We apply Lemma A.6.3 to estimate

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_t^T |Z_s|\,\mathrm{d}s\right)\middle|\mathcal{F}_t\right] &= \mathbb{E}_{\mathbb{Q}}\left[\sum_{k=0}^{\infty}\frac{1}{k!}\left(\int_t^T |Z_s|\,\mathrm{d}s\right)^k\middle|\mathcal{F}_t\right] \\
&\leq \sum_{k=0}^{\infty}\frac{1}{k!}\mathbb{E}_{\mathbb{Q}}\left[\left(\int_t^T |Z_s|\,\mathrm{d}s\right)^k\middle|\mathcal{F}_t\right] \leq \sum_{k=0}^{\infty}\frac{1}{k!}\mathbb{E}_{\mathbb{Q}}\left[\left((T-t)\int_t^T |Z_s|^2\,\mathrm{d}s\right)^{\frac{k}{2}}\middle|\mathcal{F}_t\right] \\
&\leq \sum_{k=0}^{\infty}\frac{1}{k!}\left(\mathbb{E}_{\mathbb{Q}}\left[\left(T\int_t^T |Z_s|^2\,\mathrm{d}s\right)^k\middle|\mathcal{F}_t\right]\right)^{\frac{1}{2}} \leq \sum_{k=0}^{\infty}\frac{T^{\frac{k}{2}}}{k!}\left(k!\left(\|Z\|^2_{BMO(\mathbb{Q})}\right)^k\right)^{\frac{1}{2}} \\
&\leq \sum_{k=0}^{\infty}\frac{T^{\frac{k}{2}}}{\sqrt{k!}}K^k =: C < \infty.
\end{aligned}$$

We use

$$\left(\frac{T^{\frac{k+1}{2}}}{\sqrt{(k+1)!}}K^{k+1}\right)\cdot\left(\frac{T^{\frac{k}{2}}}{\sqrt{k!}}K^k\right)^{-1} = \frac{T^{\frac{1}{2}}}{\sqrt{k+1}}K \to 0, \quad k \to \infty,$$

to see that the series converges absolutely and is monotonically increasing in K. □

Theorem A.6.5 (Theorem 3.6. in [Kaz94])**.** *Let $\mu \in BMO(\mathbb{P})$ be $\mathbb{R}^{1\times d}$-valued. Define the probability measure $\mathbb{Q}^\mu := \mathcal{E}(\mu\bullet W)_T\cdot\mathbb{P}$. Then for all progressively measurable processes Z:*

$$\|Z\|_{BMO(\mathbb{Q}^\mu)} \leq K_1\|Z\|_{BMO(\mathbb{P})} \quad \textit{and} \quad \|Z\|_{BMO(\mathbb{P})} \leq K_2\|Z\|_{BMO(\mathbb{Q}^\mu)}$$

with some real constants $K_1, K_2 > 0$ only depending on $\|\mu\|_{BMO(\mathbb{P})}$ and montonically increasing in this value.

As an application let us prove the following statement:

Lemma A.6.6. *For some $N \in \mathbb{N}$ let Y be an $\mathbb{R}^{1\times N}$-valued progressively measurable bounded process on $[0,T]$, the dynamical behavior of which is described by*

$$Y_s = Y_T - \int_s^T \mathrm{d}W_r^\top Z_r - \int_s^T \left(\alpha_r + Y_r\left(\delta_r I_N + \beta_r\right) + \sum_{i=1}^d Z_r^i \gamma_r^i + \mu_r^\top Z_r\right) \mathrm{d}r, \quad s \in [0,T], \tag{A.8}$$

where

- Y_T *is* $\mathbb{R}^{1\times N}$*-valued,* $\mathcal{F}_T$*-measurable and bounded,*
- Z *is some* $\mathbb{R}^{d\times N}$*-valued progressively measurable process s.t.* $\int_0^T |Z|_r^2 \,\mathrm{d}r < \infty$ *a.s., which can also be interpreted as a vector* $(Z^i)_{i=1,\ldots,d}$ *of* $\mathbb{R}^{1\times N}$*-valued progressively measurable processes* Z^i*,* $i = 1, \ldots, d$*,*
- α *is an* $\mathbb{R}^{1\times N}$*-valued* $BMO(\mathbb{P})$*-process,*
- δ *is some non-negative progressively measurable process with* $\int_0^T \delta_s \,\mathrm{d}s < \infty$ *a.s.,*
- $I_N \in \mathbb{R}^{N\times N}$ *is the identity matrix,*
- β *is an* $\mathbb{R}^{N\times N}$*-valued* $BMO(\mathbb{P})$*-process,*
- γ^i*,* $i = 1, \ldots, d$*, are progressively measurable and bounded* $\mathbb{R}^{N\times N}$*-valued processes,*
- μ *is an* $\mathbb{R}^d$*-valued* $BMO(\mathbb{P})$*-process.*

Then Y *is bounded by*

$$\|Y\|_\infty \le C_1 \cdot \|Y_T\|_\infty + C_2 \cdot \|\alpha\|_{BMO(\mathbb{P})},$$

with constants $C_1, C_2 \in [0,\infty)$ *which depend only on* T*,* $\|\beta\|_{BMO(\mathbb{P})}$*,* $\|\mu\|_{BMO(\mathbb{P})}$ *and* $\|\gamma^{(i)}\|_\infty$*,* $i = 1, \ldots, d$*, and are monotonically increasing in these values.*

Proof. In order to get rid of the term $\mu_r^\top Z_r$ we define a Brownian motion with drift on $[0,T]$ via

$$\tilde{W}_s := W_s + \int_0^s \mu_r \,\mathrm{d}r, \quad s \in [0,T]$$

Using a standard Girsanov measure change $\tilde{W}$ is a Brownian motion w.r.t. to some equivalent probability measure $\mathbb{Q}$. Furthermore, using (A.8) the process Y has dynamics

$$Y_s = Y_T - \int_s^T \mathrm{d}\tilde{W}_r^\top Z_r - \int_s^T \left(\alpha_r + Y_r\left(\delta_r I_N + \beta_r\right) + \sum_{i=1}^d Z_r^i \gamma_r^i\right) \mathrm{d}r, \quad s \in [0,T].$$

Now, choose a $t \in [0,T]$. We want to control Y_t. For that purpose define

$$\Gamma_s := \exp\left(-\int_t^s \left(\delta_r I_N + \beta_r\right) \mathrm{d}r - \int_t^s \sum_{i=1}^d \mathrm{d}\tilde{W}_r^i \gamma_r^i - \frac{1}{2}\int_{t_1}^s \sum_{i=1}^d \gamma_r^i \gamma_r^i \,\mathrm{d}r\right), \quad s \in [t,T].$$

According to the Itô formula Γ has dynamics

$$\Gamma_s = \Gamma_T + \int_s^T \sum_{i=1}^d \mathrm{d}\tilde{W}_r^i \gamma_r^i \Gamma_r + \int_s^T (\delta_r I_N + \beta_r)\, \Gamma_r \,\mathrm{d}r,$$

for $s \in [t, T]$. Now, apply the Itô formula to $Y_s \Gamma_s$:

$$\begin{aligned} Y_s\Gamma_s =& Y_T\Gamma_T - \int_s^T \sum_{i=1}^d \mathrm{d}\tilde{W}_r^i \left(Z_r^i \Gamma_r - Y_r \gamma_r^i \Gamma_r \right) \\ & - \int_s^T \Bigg\{ \Big(\alpha_r + Y_r(\delta_r I_N + \beta_r) + \sum_{i=1}^d Z_r^i \gamma_r^i \Big) \Gamma_r - Y_r(\delta_r I_N + \beta_r)\Gamma_r \\ & - \sum_{i=1}^d Z_r^i \gamma_r^i \Gamma_r \Bigg\} \,\mathrm{d}r. \end{aligned}$$

A few terms cancel out and we end up with

$$Y_s\Gamma_s = Y_T\Gamma_T - \int_s^T \sum_{i=1}^d \mathrm{d}\tilde{W}_r^i \left(Z_r^i \Gamma_r - Y_r \gamma_r^i \Gamma_r \right) - \int_s^T \alpha_r \Gamma_r \,\mathrm{d}r. \tag{A.9}$$

We now want to control $\sup_{s\in[t,T]} |\Gamma_s|$: Observe that due to $\delta \geq 0$ we have for all $p \geq 1$

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}} \left[\sup_{s\in[t,T]} |\Gamma_s|^p \,\middle|\, \mathcal{F}_t \right] \\ &= \mathbb{E}_{\mathbb{Q}} \left[\sup_{s\in[t,T]} \left| \exp\left(- \int_t^s \delta_r \,\mathrm{d}r - \int_t^s \beta_r \,\mathrm{d}r - \int_t^s \sum_{i=1}^d \mathrm{d}\tilde{W}_r^i \gamma_r^i - \frac{1}{2} \int_t^s \sum_{i=1}^d \gamma_r^i \gamma_r^i \,\mathrm{d}r \right) \right|^p \,\middle|\, \mathcal{F}_t \right] \\ &\leq \mathbb{E}_{\mathbb{Q}} \left[\sup_{s\in[t,T]} \left| \exp\left(\left| \int_t^s \beta_r \,\mathrm{d}r \right| + \left| \int_t^s \sum_{i=1}^d \mathrm{d}\tilde{W}_r^i \gamma_r^i \right| + \frac{1}{2} \left| \int_t^s \sum_{i=1}^d \gamma_r^i \gamma_r^i \,\mathrm{d}r \right| \right) \right|^p \,\middle|\, \mathcal{F}_t \right] \\ &\leq \mathbb{E}_{\mathbb{Q}} \left[\sup_{s\in[t,T]} \exp\left(p \int_t^s |\beta_r| \,\mathrm{d}r + \frac{p}{2} T \|\gamma\|_\infty^2 \right) \cdot \sup_{s\in[t,T]} \exp\left(p \left| \int_t^s \sum_{i=1}^d \mathrm{d}\tilde{W}_r^i \gamma_r^i \right| \right) \,\middle|\, \mathcal{F}_t \right], \end{aligned}$$

which using Cauchy-Schwarz inequality can be further controlled by

$$\begin{aligned} \Bigg(& \mathbb{E}_{\mathbb{Q}} \left[\exp\left(\int_t^T 2p|\beta_r| \,\mathrm{d}r + pT\|\gamma\|_\infty^2 \right) \,\middle|\, \mathcal{F}_t \right] \\ & \qquad \times \mathbb{E}_{\mathbb{Q}} \left[\sup_{s\in[t,T]} \exp\left(2p \left| \int_t^s \sum_{i=1}^d \mathrm{d}\tilde{W}_r^i \gamma_r^i \right| \right) \,\middle|\, \mathcal{F}_t \right] \Bigg)^{\frac{1}{2}}. \end{aligned}$$

According to Lemma A.6.4 the first of the two factors above can be controlled by a finite constant, which depends only on p, $\|\beta\|_{BMO(\mathbb{Q})}$, $\|\gamma\|_\infty$ and T and is monotonically increasing in these values. Also, note that $\|\beta\|_{BMO(\mathbb{Q})}$ can be controlled by $\|\beta\|_{BMO(\mathbb{P})}$ and $\|\mu\|_{BMO(\mathbb{P})}$ according to Theorem A.6.5.

The second factor can be estimated using Doob's martingale inequality:

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[\exp\left(2p\sup_{s\in[t,T]}\left|\int_t^s\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\gamma_r^i\right|\right)\middle|\mathcal{F}_t\right]\\
&\quad\leq\sum_{k=0}^\infty\frac{1}{k!}\mathbb{E}_{\mathbb{Q}}\left[\left(\sup_{s\in[t,T]}\left|\int_t^s\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\left(2p\gamma_r^i\right)\right|\right)^k\middle|\mathcal{F}_t\right]\\
&\quad\leq 1+\mathbb{E}_{\mathbb{Q}}\left[\sup_{s\in[t,T]}\left|\int_t^s\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\left(2p\gamma_r^i\right)\right|\middle|\mathcal{F}_t\right]\\
&\qquad+\sum_{k=2}^\infty\frac{1}{k!}\left(\frac{k}{k-1}\right)^k\mathbb{E}_{\mathbb{Q}}\left[\left|\int_t^T\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\left(2p\gamma_r^i\right)\right|^k\middle|\mathcal{F}_t\right].
\end{aligned}$$

Using Cauchy-Schwarz inequality and Doob's martingale inequality again, the above value can be controlled by

$$\begin{aligned}
&1+2\left(\mathbb{E}_{\mathbb{Q}}\left[\left|\int_t^T\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\left(2p\gamma_r^i\right)\right|^2\middle|\mathcal{F}_t\right]\right)^{\frac{1}{2}}+\sum_{k=2}^\infty\frac{1}{k!}4\mathbb{E}_{\mathbb{Q}}\left[\left|\int_t^T\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\left(2p\gamma_r^i\right)\right|^k\middle|\mathcal{F}_t\right]\\
&\quad\leq 10\mathbb{E}_{\mathbb{Q}}\left[\exp\left(2p\left|\int_t^T\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\gamma_r^i\right|\right)\middle|\mathcal{F}_t\right].
\end{aligned}$$

This value is bounded by a finite constant, which depends only on p, T and $\|\gamma\|_\infty$ and is monotonically increasing in these values: For instance use Theorem 2.1 in [Kaz94] by applying it to finitely many sufficiently small subintervals of $[t,T]$ such that $2p\|\gamma\|_\infty$ multiplied by the square root of the size of every subinterval is smaller $\frac{1}{5}$. Also, use the triangle inequality and the tower property after splitting up the stochastic integral. One implication of the above control for $\sup_{s\in[t,T]}|\Gamma_s|$ is that the stochastic integral in (A.9) represents a uniformly integrable martingale w.r.t. $\mathbb{Q}$ since

$$\int_t^s\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\left(Z_r^i\Gamma_r-Y_r\gamma_r^i\Gamma_r\right)=Y_s\Gamma_s-Y_t\Gamma_t-\int_t^s\alpha_r\Gamma_r\,\mathrm{d}r\quad\text{a.s., for all } s\in[t,T],$$

and, therefore, using triangle inequality, Cauchy-Schwarz inequality and simple estimates

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\Big[\sup_{s\in[t,T]}\Big|\int_t^s\sum_{i=1}^d \mathrm{d}\tilde{W}_r^i\left(Z_r^i\Gamma_r-Y_r\gamma_r^i|\Gamma_r|\right)\Big|\Big]\\
&\qquad\leq 2\|Y\|_\infty\mathbb{E}_{\mathbb{Q}}\left[\sup_{s\in[t,T]}|\Gamma_s|\right]+\mathbb{E}_{\mathbb{Q}}\left[\sup_{s\in[t,T]}\left|\int_t^s\alpha_r\Gamma_r\,\mathrm{d}r\right|\right]\\
&\qquad\leq 2\|Y\|_\infty\mathbb{E}_{\mathbb{Q}}\left[\sup_{s\in[t,T]}|\Gamma_s|\right]+\mathbb{E}_{\mathbb{Q}}\left[\sup_{s\in[t,T]}|\Gamma_s|\int_t^T|\alpha_r|\,\mathrm{d}r\right]\\
&\qquad\leq 2\|Y\|_\infty\mathbb{E}_{\mathbb{Q}}\left[\sup_{s\in[t,T]}|\Gamma_s|\right]+\left(\mathbb{E}_{\mathbb{Q}}\left[\sup_{s\in[t,T]}|\Gamma_s|^2\right]\mathbb{E}_{\mathbb{Q}}\left[T\int_t^T|\alpha|_r^2\,\mathrm{d}r\right]\right)^{\frac{1}{2}},
\end{aligned}$$

which is finite due to $\alpha \in BMO(\mathbb{P})$ and Theorem A.6.5. We can finally estimate using (A.9) and Cauchy-Schwarz inequality:

$$\begin{aligned}
|Y_t| = |\mathbb{E}_\mathbb{Q}[Y_t\Gamma_t|\mathcal{F}_t]| &= \left|\mathbb{E}_\mathbb{Q}[Y_T\Gamma_T|\mathcal{F}_t] - \mathbb{E}_\mathbb{Q}\left[\int_t^T \alpha_r\Gamma_r \,\mathrm{d}r\middle|\mathcal{F}_t\right]\right| \\
&\le \|Y_T\|_\infty \mathbb{E}_\mathbb{Q}[|\Gamma_T||\mathcal{F}_t] + \left(\mathbb{E}_\mathbb{Q}\left[\sup_{s\in[t,T]}|\Gamma_s|^2\middle|\mathcal{F}_t\right]\right)^{\frac{1}{2}}\left(\mathbb{E}_\mathbb{Q}\left[T\int_t^T|\alpha_r|^2\,\mathrm{d}r\middle|\mathcal{F}_t\right]\right)^{\frac{1}{2}} \\
&\le \|Y_T\|_\infty\sqrt{\mathbb{E}_\mathbb{Q}[|\Gamma_T|^2|\mathcal{F}_t]} + \sqrt{T}\|\alpha\|_{BMO(\mathbb{Q})}\left(\mathbb{E}_\mathbb{Q}\left[\sup_{s\in[t,T]}|\Gamma_s|^2\middle|\mathcal{F}_t\right]\right)^{\frac{1}{2}} \\
&\le \|Y_T\|_\infty\sqrt{\mathbb{E}_\mathbb{Q}[|\Gamma_T|^2|\mathcal{F}_t]} + K_1\|\alpha\|_{BMO(\mathbb{P})}\left(\mathbb{E}_\mathbb{Q}\left[\sup_{s\in[t,T]}|\Gamma_s|^2\middle|\mathcal{F}_t\right]\right)^{\frac{1}{2}},
\end{aligned}$$

where we again used Theorem A.6.5. K_1 depends only on $\|\mu\|_{BMO(\mathbb{P})}$ and T. □

The following theorem is an extension of a result from [BE09].

Theorem A.6.7. *Let Y, Z, X, ψ, φ be some progressively measurable processes on $[0,T]$ such that*

- *Y is real-valued and bounded,*
- *Z is $\mathbb{R}^{1\times d}$-valued and s.t. $\int_0^T |Z_s|^2 < \infty$ a.s.,*
- *ψ, φ are real-valued and in $BMO(\mathbb{P})$,*
- *X is real-valued and satisfies $X \le \psi^2 + |Z|\varphi + C|Z|^2$ with some constant $C > 0$.*

Assume furthermore

$$Y_t = Y_T + \int_t^T X_s\,\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}W_s \quad a.s., \quad t\in[0,T].$$

Then we have $\|Z\|_{BMO(\mathbb{P})} \le K < \infty$ for some constant K, which only depends on $\|Y\|_\infty$, C, $\|\varphi\|_{BMO(\mathbb{P})}$, $\|\psi\|_{BMO(\mathbb{P})}$ and is monotonically increasing in theses values.

Proof. Clearly, we see

$$X \le \psi^2 + |Z|\varphi + C|Z|^2 \le (\psi^2 + \frac{1}{2}\varphi^2) + (C + \frac{1}{2})|Z|^2.$$

Define $\tilde\psi := \sqrt{\psi^2 + \frac{1}{2}\varphi^2} \in BMO(\mathbb{P})$, $\tilde C := C + \frac{1}{2}$, and write

$$Y_t = Y_0 - \int_0^t X_s\,\mathrm{d}s + \int_0^t Z_s\,\mathrm{d}W_s.$$

Let $\beta\in\mathbb{R}$ be some constant specified later. Using Itô's formula we get

$$\begin{aligned}
\exp(\beta Y_t) = {}&\exp(\beta Y_0) - \int_0^t \beta\exp(\beta Y_s)X_s\,\mathrm{d}s + \int_0^t \beta\exp(\beta Y_s)Z_s\,\mathrm{d}W_s \\
&+ \frac{\beta^2}{2}\int_0^t \exp(\beta Y_s)|Z_s|^2\,\mathrm{d}s.
\end{aligned}$$

So for every stopping time $\tau \in [t,T]$ we can write

$$\begin{aligned}\exp(\beta Y_t) = &\exp(\beta Y_\tau) + \int_t^\tau \beta \exp(\beta Y_s) X_s \,\mathrm{d}s - \int_t^\tau \beta \exp(\beta Y_s) Z_s \,\mathrm{d}W_s \\ &- \frac{\beta^2}{2} \int_t^\tau \exp(\beta Y_s)|Z_s|^2 \,\mathrm{d}s,\end{aligned}$$

which can be rearranged to

$$\beta \int_t^\tau \exp(\beta Y_s) \left(\frac{\beta}{2} |Z_s|^2 - X_s \right) \mathrm{d}s = \exp(\beta Y_\tau) - \exp(\beta Y_t) - \int_t \tau \beta \exp(\beta Y_s) Z_s \,\mathrm{d}W_s,$$

or again to

$$\begin{aligned}&\beta \int_t^\tau \exp(\beta Y_s) \left(\frac{\beta}{2} |Z_s|^2 + \tilde{\psi}_s^2 - X_s \right) \mathrm{d}s \\ &\quad = \exp(\beta Y_\tau) - \exp(\beta Y_t) + \beta \int_t^\tau \exp(\beta Y_s) \tilde{\psi}_s^2 \,\mathrm{d}s - \int_t^\tau \beta \exp(\beta Y_s) Z_s \,\mathrm{d}W_s.\end{aligned}$$

Setting $\beta := 2\tilde{C} + 2 = 2C + 3$, we have

$$|Z_s|^2 \le \frac{\beta}{2} |Z_s|^2 + \tilde{\psi}_s^2 - X_s.$$

Now choose a localizing sequence of stopping times $\tau_n \in [t,T]$, $n \in \mathbb{N}$, such that

$$\mathbb{E}\left[\int_t^{\tau_n} |Z_s|^2 \,\mathrm{d}s \right] < \infty$$

for every $n \in \mathbb{N}$ while $\tau_n \uparrow T$ for $n \to \infty$. Applying conditional expectations we have

$$\begin{aligned}\mathbb{E}\Big[\beta \int_t^{\tau_n} \exp(\beta Y_s)|Z_s|^2 \,\mathrm{d}s \Big| \mathcal{F}_t\Big] &\le \mathbb{E}\left[\beta \int_t^{\tau_n} \exp(\beta Y_s) \left(\frac{\beta}{2} |Z_s|^2 + \tilde{\psi}_s^2 - X_s \right) \mathrm{d}s\right] \\ &\le \mathbb{E}\left[\exp(\beta Y_{\tau_n}) - \exp(\beta Y_t) + \beta \int_t^{\tau_n} \exp(\beta Y_s)(\psi^2 + \frac{1}{2}\varphi^2) \,\mathrm{d}s \middle| \mathcal{F}_t \right],\end{aligned}$$

which we can rewrite as

$$\begin{aligned}&\mathbb{E}\Big[\int_t^{\tau_n} \exp(\beta Y_s)|Z_s|^2 \,\mathrm{d}s \Big| \mathcal{F}_t\Big] \\ &\quad \le \mathbb{E}\left[\frac{\exp(\beta Y_{\tau_n}) - \exp(\beta Y_t)}{\beta Y_T - \beta Y_t} (Y_{\tau_n} - Y_t) + \int_t^{\tau_n} \exp(\beta Y_s)(\psi^2 + \frac{1}{2}\varphi^2) \,\mathrm{d}s \middle| \mathcal{F}_t \right] \\ &\quad \le \left\| \frac{\exp(\beta Y_{\tau_n}) - \exp(\beta Y_t)}{\beta Y_{\tau_n} - \beta Y_t} \right\|_\infty \cdot \|Y_{\tau_n} - Y_t\|_\infty \\ &\qquad + \exp\left(\beta \|Y\|_\infty\right) \left(\|\psi\|_{BMO(\mathbb{P})}^2 + \frac{1}{2} \|\varphi\|_{BMO(\mathbb{P})}^2 \right).\end{aligned}$$

Finally, note that the exponential function is Lipschitz continuous on any interval $[a,b]$ with $\exp(a \vee b)$ as Lipschitz constant, so

$$\left\| \frac{\exp(\beta Y_{\tau_n}) - \exp(\beta Y_t)}{\beta Y_{\tau_n} - \beta Y_t} \right\|_\infty \cdot \|Y_{\tau_n} - Y_t\|_\infty \le \exp(\beta \|Y\|_\infty) \cdot 2 \cdot \|Y\|_\infty.$$

We obtain by monotone convergence

$$\begin{aligned}
\mathbb{E}\left[\int_t^T |Z_s|^2 \,\mathrm{d}s \middle| \mathcal{F}_t\right] &= \lim_{n\to\infty} \mathbb{E}\left[\int_t^{\tau_n} |Z_s|^2 \,\mathrm{d}s \middle| \mathcal{F}_t\right] \\
&\leq \lim_{n\to\infty} \exp(\beta\|Y\|_\infty)\mathbb{E}\left[\int_t^{\tau_n} \exp(\beta Y_s)|Z_s|^2 \,\mathrm{d}s \middle| \mathcal{F}_t\right] \\
&\leq 2\exp(2\beta\|Y\|_\infty)\|Y\|_\infty + \exp\left(2\beta\|Y\|_\infty\right)\left(\|\psi\|^2_{BMO(\mathbb{P})} + \frac{1}{2}\|\varphi\|^2_{BMO(\mathbb{P})}\right) \\
&= 2\exp(2(2C+3)\|Y\|_\infty)\|Y\|_\infty \\
&\quad + \exp\left(2(2C+3)\|Y\|_\infty\right)\left(\|\psi\|^2_{BMO(\mathbb{P})} + \frac{1}{2}\|\varphi\|^2_{BMO(\mathbb{P})}\right),
\end{aligned}$$

which is finite and increasing in $\|Y\|_\infty$, C, $\|\varphi\|_{BMO(\mathbb{P})}$, $\|\psi\|_{BMO(\mathbb{P})}$. □

Miscellaneous

Lemma A.6.8. *For $N, m \in \mathbb{N}$, let $g\colon \mathbb{R}^N \to \mathbb{R}^m$ be Lipschitz continuous. Moreover, let $X\colon \mathbb{R}^n \to \mathbb{R}^N$, $n \in \mathbb{N}$ be weakly differentiable. Then*

- *$g(X)$ is also weakly differentiable,*
- *for almost every $\lambda \in \mathbb{R}^n$ the restriction $g|_{T^X_\lambda}$ of g to the affine space*

$$T^X_\lambda := \left\{x \in \mathbb{R}^N \middle| x = X(\lambda) + \frac{\mathrm{d}}{\mathrm{d}\lambda}X(\lambda)v, \text{ for some } v \in \mathbb{R}^n\right\}$$

  is differentiable at $X(\lambda)$ and
- *for almost all $\lambda \in \mathbb{R}^n$ we have*

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}g(X)(\lambda) = \frac{\mathrm{d}}{\mathrm{d}x}g|_{T^X_\lambda}(X(\lambda))\frac{\mathrm{d}}{\mathrm{d}\lambda}X(\lambda).$$

This implies in particular:

- *If $n = N$ and the matrix $\frac{\mathrm{d}}{\mathrm{d}\lambda}X(\lambda)$ is invertible for a.a. λ, then $T^X_\lambda = \mathbb{R}^N$ for a.a. λ and*

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}g(X) = \left(\frac{\mathrm{d}}{\mathrm{d}x}g\right)(X)\frac{\mathrm{d}}{\mathrm{d}\lambda}X$$

  a.e., where $\frac{\mathrm{d}}{\mathrm{d}x}g$ is a weak derivative of g.
- *If g is differentiable everywhere then $\frac{\mathrm{d}}{\mathrm{d}\lambda}g(X) = \left(\frac{\mathrm{d}}{\mathrm{d}x}g\right)(X)\frac{\mathrm{d}}{\mathrm{d}\lambda}X$ a.e.*
- *If g is only locally Lipschitz continuous rather than Lipschitz continuous, but differentiable everywhere, while X is bounded, then still $\frac{\mathrm{d}}{\mathrm{d}\lambda}g(X) = \left(\frac{\mathrm{d}}{\mathrm{d}x}g\right)(X)\frac{\mathrm{d}}{\mathrm{d}\lambda}X$ a.e.*

Proof. For the main statement consult Corollary 3.2 in [AD90]. Concerning the three implications we may state:

- Clearly, if $\frac{\mathrm{d}}{\mathrm{d}\lambda}X(\lambda)$ is invertible for some $\lambda \in \mathbb{R}^n$, then T^X_λ must be the whole $\mathbb{R}^N$ for this λ. So for almost all λ the expression $\frac{\mathrm{d}}{\mathrm{d}x}g|_{T^X_\lambda}(X(\lambda))$ coincides with the classical derivative of g at the point $X(\lambda)$.

  Furthermore, if we choose the identity on $\mathbb{R}^n$ for X, the main statement of the lemma implies that

  - g is differentiable almost everywhere,
  - g is weakly differentiable and
  - any weak derivative of g coincides with the classical derivative up to a null set.

  So, if we define a function on $\mathbb{R}^n$ by setting it to the classical derivative of g at all points for which the classical derivative exists and to 0 for all those points in which it does not, we obtain a weak derivative.

- If g is differentiable everywhere, then $\frac{\mathrm{d}}{\mathrm{d}x}g|_{T^X_\lambda}(X(\lambda))$ is just the classical derivative of g at $X(\lambda)$.

- If X is bounded, we can assume without loss of generality that g is Lipschitz continuous by restricting its domain or using a removable inner cutoff.

$\square$

Notation

Basic notation and conventions

$\mathbf{1}_A$	indicator function of a set A
$\lfloor \alpha \rfloor$	smallest integer greater or equal $\alpha \in \mathbb{R}$
$\lceil \alpha \rceil$	smallest integer larger or equal than $\alpha > 0$
xy	usual inner product for $x, y \in \mathbb{R}^d$, i.e. $xy := \sum_{i=1}^d x_i y_i$
$\langle x, y \rangle$	usual inner product for $x, y \in \mathbb{R}^d$, i.e. $\langle x, y \rangle := \sum_{i=1}^d x_i y_i$
$\lvert \cdot \rvert$	Euclidean norm on $\mathbb{R}^d$, i.e. $\lvert x \rvert := \sqrt{\langle x, x \rangle}$ for $x \in \mathbb{R}^d$
$\exists$	there exists
$\nexists$	there exists no
$\forall$	for all
$\emptyset$	empty set
$\lesssim$	there exists a constant $c > 0$, independent of the variables a, b such that $a \leq c \cdot b$
$\sim$	$a \lesssim b$ and $b \lesssim a$
$0/0$	0
$0 \cdot \infty$	0
$1 \cdot \infty$	∞
$\inf \emptyset$	∞
$(\cdot - u)^-$	$\max\{0, u - \cdot\}$
$a \wedge b$	$\min\{a, b\}$
$a \vee b$	$\max\{a, b\}$
Δ_T	$\{(s,t) \in [0,T]^2 \,:\, 0 \leq s \leq t \leq T\}$
$\operatorname{ess\,sup}$	essential supremum
$\max I$	maximum of a set I
$\min I$	minimum of a set I
$\inf_{t \in I}$	infimum of a set I
$\sup_{t \in I}$	supremum of a set I
$\mathbb{N}$	set of non-negative integers
$\mathbb{Q}$	set of rational numbers
$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of non-negative real numbers
$\mathbb{Z}$	set of integers

Definitions

$\langle f\rangle_t$	quadratic variation of the function f, cf. Lemma 2.3.17
$\langle f,g\rangle_t$	quadratic covariation of the functions f,g, cf. Lemma 2.3.17
$[f]_t$	quadratic variation in the sense of Föllmer of the function $f\colon [0,T]\to\mathbb{R}$, cf. Definition2.3.21
$[f,g]_t$	quadratic covariation in the sense of Föllmer of the functions $f,g\colon [0,T]\to\mathbb{R}$, cf. Definition2.3.21
$(\!|u,v]\!]$	see equation (3.3)
$\|F\|_\infty$	essential supremum of $\|F\|$ for a random variable F
$\mathbb{A}_\gamma\langle X,Y\rangle$	antisymmetric part of the γ-$\int_0^T Y_t\,\mathrm{d}X_t$, cf. (4.4)
$\mathbb{A}\langle X,Y\rangle$	antisymmetric part of the γ-$\int_0^T Y_t\,\mathrm{d}X_t$, see Theorem 4.2.4
$\mathcal{B}([0,\infty))$	Borel σ-algebra on $[0,\infty)$
$\Gamma(f,g,h)$	commutator of the functions f,g,h, cf. (5.23)
d_∞	distance for processes with respect to $\overline{P}$, see (2.8)
d_c	d_∞ restricted to set of path with quadratic variation less than c, see (2.9)
d_{loc}	localized d_∞ distance, see (2.10)
$\Delta_j f$	j-th Littlewood-Paley block of a function f, cf. (5.4)
$\overline{E}$	expectation operator with respect to $\overline{P}$, see (2.6)
$\mathbb{E}_t[F]$	conditional expectation with respect to the filtration $(\mathcal{F}_t)$
$\mathbb{E}[F\|\mathcal{F}_t]$	conditional expectation with respect to the filtration $(\mathcal{F}_t)$
$\mathbb{E}_{\hat{t},\infty}[F]$	abbreviation for $\operatorname{ess\,sup}\mathbb{E}[F\|\mathcal{F}_t]$
$\mathbb{E}$	expectation with respect to a probability measure $\mathbb{P}$
$\mathcal{F}f$	Fourier transform of the distribution or function f
$\mathcal{H}_\lambda$	set of λ-admissible simple strategies for $\lambda>0$
I^M_{max}	maximal interval for the existence of a weakly regular Markovian decoupling, see Definition 6.2.26
$L_t(S,x)$	local time of the process S, see Definition 3.1.3 and 3.1.5
$L_{u,x}$	Lipschitz constant of a function u in the variable x, cf. (6.5)
$\overline{L}_0$	space of equivalence classes with respect to $\overline{P}$
$\overline{P}$	outer measure defined via the minimal superhedging price, see Definition 2.1.1
$\pi(f,g)$	resonant term of the Littlewood-Paley decomposition of the functions f,g, cf. (5.7)
$\Pi_F(u,\xi)$	for the definition see Proposition 5.3.1
$\overline{Q}$	Vovk's outer measure, see (2.5)
$R_F(g)$	for the definition see Lemma 5.3.2
$\mathbb{S}_\gamma\langle X,Y\rangle$	symmetric part of the integral γ-$\int_0^T Y_t\,\mathrm{d}X_t$, see (4.4)
$\mathcal{S}_\lambda$	set of Vovk's λ-admissible simple strategies for $\lambda>0$, cf. (2.4)
Λ_λ	scaling operator given by $\Lambda_\lambda f(\cdot):=f(\lambda\cdot)$ for any $\lambda>0$
S	Itô map, see (5.17)
$\hat{S}$	weighted version of the Itô-Lyons map, see (5.27)
$T_f g$	part of the Littlewood-Paley decomposition of the product fg for two distributions f,g, cf. (5.7)

Functions spaces and norms

BV	space of right-continuous functions with bounded variation
$B^{\alpha}_{p,q}(\mathbb{R}^d)$	Besov space, cf. (5.5)
$\mathcal{B}^{0,\alpha}_{p,q}$	space of geometric Besov rough path, cf. Definition 5.4.1
$BMO(\mathbb{Q})$	space of BMO process with respect to a probability measure $\mathbb{Q}$, see Definition A.6.1
$C([0,\infty),\mathbb{R})$	space of continuous functions $f\colon [0,\infty)\to\mathbb{R}$
C_b	space of continuous and bounded functions $f\colon\mathbb{R}^d\to\mathbb{R}^m$
C^k	space of k times continuously differentiable functions
C^k_b	space of functions in C^k that are bounded with bounded derivatives
C^{α}	space of $\lfloor\alpha\rfloor$-times continuously differentiable functions with $(\alpha-\lfloor\alpha\rfloor)$-Hölder continuous partial derivatives of order $\lfloor\alpha\rfloor$ or with continuous partial derivatives of order $\lfloor\alpha\rfloor$ in the case $\alpha=\lfloor\alpha\rfloor$
$C^{\infty}_{\mathcal{T}}$	space of smooth functions $\vartheta_{\mathcal{T}}\colon\mathbb{R}\to\mathbb{R}^n$ with support $\operatorname{supp}\vartheta_{\mathcal{T}}\subset[-2\mathcal{T},2\mathcal{T}]$
$\mathscr{C}^q_{\mathbb{S}}$	space of controlled paths with respect to $\mathbb{S}$, cf. Definition 2.3.6
$\mathbb{H}^2(\mathbb{R})$	space of progressively measurable processes $Z\colon\Omega\times[0,1]\to\mathbb{R}$ satisfying $\mathbb{E}[\int_0^1 \lvert Z_t\rvert^2\,\mathrm{d}t]<\infty$
$L^p(\mathbb{R}^d,\mathbb{R}^m)$	space of Lebesgue p-integrable functions $f\colon\mathbb{R}^d\to\mathbb{R}^m$ for $p\in(0,\infty)$
$L^{\infty}(\mathbb{R}^d,\mathbb{R}^m)$	space of bounded functions $f\colon\mathbb{R}^d\to\mathbb{R}^m$
$\overline{L}_0([0,T],\mathbb{R}^m)$	space of equivalence classes with respect to d_{∞}
$\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$	space of bounded linear operator mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$
$\mathcal{S}(\mathbb{R}^d)$	space of Schwartz functions on $\mathbb{R}^d$
$\mathcal{S}'(\mathbb{R}^d)$	space of tempered distributions on $\mathbb{R}^d$
$\mathscr{S}^q_X$	set of functions similar to X, see Definition 4.2.2
$\mathbb{S}^2(\mathbb{R})$	space of progressively measurable processes $Y\colon\Omega\times[0,1]\to\mathbb{R}$ satisfying $\sup_{t\in[0,1]}\mathbb{E}[\lvert Y_t\rvert^2]<\infty$
$\mathcal{V}^p([0,T],\mathbb{R}^d)$	space of continuous functions with finite p-variation
$\omega_p(f,\delta)$	modulus of continuity of a function f, cf. (5.2)
$\lVert\cdot\rVert_{\omega:\alpha,p,q}$	norm on the Besov space $B^{\alpha}_{p,q}$ via the modulus of continuity, cf. (5.3)
$\lVert\cdot\rVert_{\alpha,p,q}$	norm on the Besov space $B^{\alpha}_{p,q}$, cf. (5.5)
$\lVert\cdot\rVert_{p\text{-var},[s,t]}$	p-variation seminorm on the interval $[s,t]\subset\mathbb{R}$, see 2.1
$\lVert\cdot\rVert_p$	p-variation seminorm $\lVert\cdot\rVert_{p\text{-var},[0,T]}$
$\lVert\cdot\rVert_{C^k_b}$	norm on C^k_b, see 2.2
$\lVert\cdot\rVert_{L^p}$	L^p-norm with respect to the Lebesgue measure, $p\in(0,\infty)$
$\lVert\cdot\rVert_{\infty}$	supremum norm
$\lVert\cdot\rVert_{\alpha}$	Hölder norm for $\alpha\in(0,1)$
$\lVert\cdot\rVert_{\mathscr{C}^q_{\mathbb{S}}}$	norm on $\mathscr{C}^q_{\mathbb{S}}$, cf. Definition 2.3.6

Bibliography

[ABPS13] Beatrice Acciaio, Mathias Beiglböck, Friedrich Penkner, and Walter Schachermayer, *A MODEL-FREE VERSION OF THE FUNDAMENTAL THEOREM OF ASSET PRICING AND THE SUPER-REPLICATION THEOREM*, Mathematical Finance (2013), n/a–n/a.

[AD90] L. Ambrosio and G. Dal Maso, *A general chain rule for distributional derivatives*, Proc. Am. Math. Soc. **108** (1990), no. 3, 691–702.

[AHI08] Stefan Ankirchner, Gregor Heyne, and Peter Imkeller, *A BSDE approach to the Skorokhod embedding problem for the Brownian motion with drift*, Stoch. Dyn. **8** (2008), no. 1, 35–46.

[Ahn97] Hyungsok Ahn, *Semimartingale integral representation*, Ann. Probab. **25** (1997), no. 2, 997–1010.

[AHS15] Stefan Ankirchner, David Hobson, and Philipp Strack, *Finite, integrable and bounded time embeddings for diffusions*, Bernoulli **21** (2015), no. 2, 1067–1088.

[ALP95] Marco Avellaneda, Arnon Levy, and Antonio Parás, *Pricing and hedging derivative securities in markets with uncertain volatilities*, Appl. Math. Finance **2** (1995), no. 2, 73–88.

[Ank05] Stefan Ankirchner, *Information and Semimartingales*, Ph.D. thesis, Humboldt-Universität zu Berlin, 2005.

[Ant93] Fabio Antonelli, *Backward-forward stochastic differential equations*, Ann. Appl. Probab. **3** (1993), no. 3, 777–793.

[AS11] Stefan Ankirchner and Philipp Strack, *Skorokhod embeddings in bounded time*, Stoch. Dyn. **11** (2011), no. 2-3, 215–226.

[ASS98] F. Abramovich, T. Sapatinas, and B. W. Silverman, *Wavelet thresholding via a Bayesian approach*, J. R. Stat. Soc. Ser. B Stat. Methodol. **60** (1998), no. 4, 725–749.

[AY79] Jacques Azema and Marc Yor, *Une solution simple au problème de Skorokhod*, Seminaire de probabilites XIII, Univ. Strasbourg 1977/78, Lect. Notes Math. 721, 90-115, 1979.

[Bac00] Louis Bachelier, *Théorie de la spéculation*, Ann. Sci. École Norm. Sup. (3) **17** (1900), 21–86.

[Bas83] Richard F. Bass, *Skorokhod imbedding via stochastic integrals*, Seminar on probability, XVII, Lecture Notes in Math., vol. 986, Springer, Berlin, 1983, pp. 221–224.

[BCD11] Hajer Bahouri, Jean-Yves Chemin, and Raphael Danchin, *Fourier analysis and nonlinear partial differential equations*, Springer, 2011.

[BCH$^+$15] Mathias Beiglböck, Alexander M. G. Cox, Martin Huesmann, Nicolas Perkowski, and David J. Prömel, *Pathwise super-replication via Vovk's outer measure*, Preprint arXiv:1504.03644 (2015).

[BE09] Pauline Barrieu and Nicole El Karoui, *Pricing, hedging, and designing derivatives with risk measures*, Indifference pricing: Theory and applications, Princeton, NJ: Princeton University Press, 2009, pp. 77–146.

[Ber87] Jean Bertoin, *Temps locaux et intégration stochastique pour les processus de Dirichlet*, Séminaire de Probabilités, XXI, Lecture Notes in Math., vol. 1247, Springer, Berlin, 1987, pp. 191–205.

[BHLP13] Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner, *Model-independent bounds for option prices—a mass transport approach*, Finance Stoch. **17** (2013), no. 3, 477–501.

[Bic81] Klaus Bichteler, *Stochastic Integration and L^p-Theory of Semimartingales*, Ann. Probab. (1981), 49–89.

[Bis73] Jean-Michel Bismut, *Conjugate Convex Functions in Optimal Stochastic Control*, Journal of Mathematical Analysis and Applications **44** (1973), 384–404.

[Bis78] ______, *Contrôle des Systémes Lineáres Quadratiques: Applications de l'Intégrale Stochastique*, Lecture Notes in Mathematics **649** (1978), 384–404.

[BLS06] Gérard Bourdaud, Massimo Lanza de Cristoforis, and Winfried Sickel, *Superposition operators and functions of bounded p-variation*, Rev. Mat. Iberoam. **22** (2006), no. 2, 455–487.

[Boc13] Natalia Bochkina, *Besov regularity of functions with sparse random wavelet coefficients*, Preprint arXiv:1310.3720 (2013).

[Bon81] Jean-Michel Bony, *Calcul symbolique et propagation des singularites pour les équations aux dérivées partielles non linéaires*, Ann. Sci. Éc. Norm. Supér. (4) **14** (1981), 209–246.

[BS73] Fischer Black and Myron Scholes, *The pricing of options and corporate liabilities*, The Journal of Political Economy (1973), 637–654.

[BY81] Nicolas Bouleau and Marc Yor, *Sur la variation des temps locaux de certaines semimartingales*, C. R. Acad. Sci. Paris Sér. I Math (1981), no. 292, 491–494.

[CC13] Rèmi Catellier and Khalil Chouk, *Paracontrolled Distributions and the 3-dimensional Stochastic Quantization Equation*, Preprint arXiv:1310.6869 (2013).

[CDD+12] Petru A. Cioica, Stephan Dahlke, Nicolas Döhring, Stefan Kinzel, Felix Lindner, Thorsten Raasch, Klaus Ritter, and René L. Schilling, *Adaptive wavelet methods for the stochastic Poisson equation*, BIT **52** (2012), no. 3, 589–614.

[CDR11] Rosanna Coviello, Cristina Di Girolami, and Francesco Russo, *On stochastic calculus related to financial assets without semimartingales*, Bull. Sci. Math. **135** (2011), no. 6, 733–774.

[CF10a] Rama Cont and David-Antoine Fournié, *A functional extension of the Ito formula*, C. R. Math. Acad. Sci. Paris **348** (2010), no. 1-2, 57–61.

[CF10b] ———, *Change of variable formulas for non-anticipative functionals on path space*, J. Funct. Anal. **259** (2010), no. 4, 1043–1072.

[CF13] ———, *Functional Itôcalculus and stochastic integral representation of martingales*, Ann. Probab. **41** (2013), no. 1, 109–133.

[CJ90] Peter P. Carr and Robert A. Jarrow, *The Stop-Loss Start-Gain Paradox and Option Valuation: A New Decomposition into Intrinsic and Time Value*, The Review of Financal Studies **3** (1990), no. 3, 469–492.

[CKR93] Zbigniew Ciesielski, Gérard Kerkyacharian, and Bernard Roynette, *Quelques espaces fonctionnels associés à des processus gaussiens*, Studia Math. **107** (1993), no. 2, 171–204.

[CL05] Laure Coutin and Antoine Lejay, *Semi-martingales and rough paths theory*, Electron. J. Probab. **10** (2005), no. 23, 761–785.

[CL10] Peter Carr and Roger Lee, *Hedging variance options on continuous semimartingales*, Finance Stoch. **14** (2010), no. 2, 179–207.

[CL14] Laure Coutin and Antoine Lejay, *Perturbed linear rough differential equations*, Annales mathématiques Blaise Pascal **21** (2014), no. 1, 103–150, Centre IntraFacultaire Bernoulli (CIB).

[CLPT81] R.V. Chacon, Y. Le Jan, E. Perkins, and S.J. Taylor, *Generalised arc length for Brownian motion and Levy processes*, Z. Wahrscheinlichkeitstheor. Verw. Geb. **57** (1981), 197–211.

[Cou07] Laure Coutin, *An introduction to (stochastic) calculus with respect to fractional Brownian motion*, Séminaire de Probabilités XL, Lecture Notes in Math., vol. 1899, Springer, Berlin, 2007, pp. 3–65.

[CQ02] Laure Coutin and Zhongmin Qian, *Stochastic analysis, rough path analysis and fractional Brownian motions*, Probab. Theory Relat. Fields **122** (2002), no. 1, 108–140.

[CR14] Andrea Cosso and Francesco Russo, *A regularization approach to functional Itôcalculus and strong-viscosity solutions to path-dependent PDEs*, Preprint arXiv:1401.5034 (2014).

[CW13] Alexander M.G. Cox and Jiajie Wang, *Root's barrier: construction, optimality and applications to variance options*, Ann. Appl. Probab. **23** (2013), no. 3, 859–894.

[Dav07] A. M. Davie, *Differential equations driven by rough paths: an approach via discrete approximation*, Appl. Math. Res. Express. AMRX (2007), no. 2, Art. ID abm009, 40.

[DDM69] Catherine Doléans-Dade and Paul-André Meyer, *Intégrales stochastiques par rapport aux martingales locales*, C. R. Acad. Sci. Paris Sér. A-B **269** (1969), A144–A147.

[Del02] François Delarue, *On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case*, Stochastic Process. Appl. **99** (2002), no. 2, 209–286.

[Del12] Lukasz Delong, *Applications of time-delayed backward stochastic differential equations to pricing, hedging and portfolio management in insurance and finance*, Appl. Math. **39** (2012), no. 4, 463–488.

[DH07] Mark H. A. Davis and David G. Hobson, *The range of traded option prices*, Math. Finance **17** (2007), no. 1, 1–14.

[DI10] Lukasz Delong and Peter Imkeller, *Backward stochastic differential equations with time delayed generators – results and counterexamples*, Ann. Appl. Probab. **20** (2010), no. 4, 1512–1536.

[DM82] Claude Dellacherie and Paul-André Meyer, *Probabilities and potential. B*, North-Holland Mathematics Studies, vol. 72, North-Holland Publishing Co., Amsterdam, 1982, Theory of martingales, Translated from the French by J. P. Wilson.

[DM06] Laurent Denis and Claude Martini, *A theoretical framework for the pricing of contingent claims in the presence of model uncertainty*, Ann. Appl. Probab. **16** (2006), no. 2, 827–852.

[DOR14] Mark Davis, Jan Obłój, and Vimal Raval, *ARBITRAGE BOUNDS FOR PRICES OF WEIGHTED VARIANCE SWAPS*, Math. Finance **24** (2014), no. 4, 821–854.

[DS94] Freddy Delbaen and Walter Schachermayer, *A general version of the fundamental theorem of asset pricing*, Math. Ann. **300** (1994), no. 1, 463–520.

[DS95] ———, *Arbitrage possibilities in Bessel processes and their relations to local martingales*, Probab. Theory Related Fields **102** (1995), no. 3, 357–366.

[DS01] ———, *Applications to mathematical finance*, Handbook of the geometry of Banach spaces, Vol. I, North-Holland, Amsterdam, 2001, pp. 367–391.

[DS14] Yan Dolinsky and H. Mete Soner, *Robust Hedging and Martingale Optimal Transport in Continuous Time*, Probab. Theory Related Fields **160** (2014), no. 1–2, 391–427.

[Dup09] Bruno Dupire, *Functional Itô Calculus*, Bloomberg Portfolio Research Paper (2009), no. 2009-04-FRONTIERS.

[EPQ97] Nicole El Karoui, Shige Peng, and Marie C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7** (1997), no. 1, 1–71.

[Fö81] Hans Föllmer, *Calcul d'Itô sans probabilités*, Séminaire de Probabilités XV **80** (1981), 143–150.

[FD06] Denis Feyel and Arnaud De La Pradelle, *Curvilinear integrals along enriched paths*, Electron. J. Probab. **11** (2006), 860–892.

[FDM08] Denis Feyel, Arnaud De La Pradelle, and Gabriel Mokobodzki, *A non-commutative sewing lemma*, Electron. Commun. Probab. **13** (2008), 24–34.

[FH14] Peter Friz and Martin Hairer, *A Course on Rough Paths: With an Introduction to Regularity Structures*, Springer, 2014.

[FI13] Alexander Fromm and Peter Imkeller, *Existence, Uniqueness and Regularity of Decoupling Fields to Multidimensional Fully Coupled FBSDEs*, Preprint arXiv:1310.0499 (2013).

[FIP14] Alexander Fromm, Peter Imkeller, and David J. Prömel, *An FBSDE approach to the Skorokhod embedding problem for Gaussian processes with non-linear drift*, Preprint arXiv:1408.6390 (2014).

[Fol99] Gerald B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication.

[Fre83] David Freedman, *Brownian motion and diffusion*, second ed., Springer-Verlag, New York, 1983.

[Fri05] Peter K. Friz, *Continuity of the Itô-map for Hölder rough paths with applications to the support theorem in Hölder norm*, Probability and partial differential equations in modern applied mathematics. Selected papers presented at the 2003 IMA summer program, Minneapolis, MN, USA, July 21 – August 1, 2003, New York, NY: Springer, 2005, pp. 117–135.

[FS13] Hans Föllmer and Alexander Schied, *Probabilistic aspects of finance*, Bernoulli **19** (2013), no. 4, 1306–1326.

[FS14] Peter K. Friz and Atul Shekhar, *General Rough integration, Lévy Rough paths and a Lévy–Kintchine type formula*, Preprint arXiv:1212.5888 (2014).

[FV10a] Peter Friz and Nicolas Victoir, *Differential equations driven by Gaussian signals*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 2, 369–413.

[FV10b] Peter K. Friz and Nicolas B. Victoir, *Multidimensional Stochastic Processes as Rough Paths. Theory and Applications*, Cambridge University Press, 2010.

[FZ06] Chunrong Feng and Huaizhong Zhao, *Two-parameter p,q-variation Paths and Integrations of Local Times*, Potential Analysis (2006), no. 25, 165–204.

[GF00] Peter Grandits and Neil Falkner, *Embedding in Brownian motion with drift and the Azéma-Yor construction*, Stochastic Processes Appl. **85** (2000), no. 2, 249–254.

[GIP12] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski, *Paracontrolled distributions and singular PDEs*, to appear in Forum of Mathematics, Pi, Preprint arXiv:1210.2684 (2012).

[GIP14] ______, *A Fourier approach to pathwise stochastic integration*, Preprint arXiv:1410.4006 (2014).

[GP15] Massimiliano Gubinelli and Nicolas Perkowski, *KPZ equation reloaded*, Preprint arXiv:1508.03877 (2015).

[Gub04] Massimiliano Gubinelli, *Controlling rough paths*, J. Funct. Anal. **216** (2004), no. 1, 86–140.

[Hai13] Martin Hairer, *Solving the KPZ equation*, Ann. of Math. (2) **178** (2013), no. 2, 559–664.

[Hai14] ______, *A theory of regularity structures*, Inventiones mathematicae (2014), 1–236.

[Hal68] W.J. Hall, *On the Skorokhod Embedding Theorem*, Stanford University. Dept. of Statistics and National Science Foundation (U.S.), 1968, Technical Report 33.

[HL15] Martin Hairer and Cyril Labbé, *Multiplicative stochastic heat equations on the whole space*, Preprint arXiv:1504.07162 (2015).

[HN07] Yaozhong Hu and David Nualart, *Differential equations driven by Hölder continuous of order greater than 1/2*, Stochastic analysis and applications, Abel Symp., vol. 2, Springer, Berlin, 2007, pp. 399–413.

[Hob98] David G. Hobson, *Robust hedging of the lookback option*, Finance Stoch. **2** (1998), no. 4, 329–347.

[Hob11] David Hobson, *The Skorokhod embedding problem and model-independent bounds for option prices*, Paris-Princeton Lectures on Mathematical Finance 2010, Lecture Notes in Math., vol. 2003, Springer, Berlin, 2011, pp. 267–318.

[HP95] Y. Hu and S. Peng, *Solution of forward-backward stochastic differential equations*, Probab. Theory Relat. Fields **103** (1995), no. 2, 273–283.

[IP11] Peter Imkeller and Nicolas Perkowski, *The existence of dominating local martingale measures*, to appear in Finance Stoch., Preprint arXiv:1111.3885 (2011).

[IP15] Peter Imkeller and David J. Prömel, *Existence of Lévy's area and pathwise integration*, Commun. Stoch. Anal. **Vol.9** (2015), no. No.1, 93–111.

[Itô44] Kiyosi Itô, *Stochastic integral*, Proc. Imp. Acad. Tokyo **20** (1944), 519–524.

[Itô51] ______, *On stochastic differential equations*, Mem. Amer. Math. Soc. **1951** (1951), no. 4, 51.

[Jac79] Jean Jacod, *Calcul stochastique et problemes de martingales*, vol. 714, Springer, 1979.

[Kar95] Rajeeva L. Karandikar, *On pathwise stochastic integration*, Stochastic Process. Appl. **57** (1995), no. 1, 11–18.

[Kaz94] Norihiko Kazamaki, *Continuous exponential martingales and BMO*, Berlin: Springer, 1994.

[Kel14] David TB Kelly, *Rough path recursions and diffusion approximations*, Preprint arXiv:1402.3188 (2014).

[Kho94] Davar Khoshnevisan, *Exact rates of convergence to Brownian local time*, Ann. Probab. **22** (1994), no. 3, 1295–1330.

[KK07] Ioannis Karatzas and Constantinos Kardaras, *The numéraire portfolio in semimartingale financial models*, Finance Stoch. **11** (2007), no. 4, 447–493.

[Kni21] Frank H Knight, *Risk, uncertainty and profit*, New York: Hart, Schaffner and Marx (1921).

[KP11] Constantinos Kardaras and Eckhard Platen, *On the semimartingale property of discounted asset-price processes*, Stochastic Process. Appl. **121** (2011), no. 11, 2678–2691.

[KS88] Ioannis Karatzas and Steven E. Shreve, *Brownian motion and stochastic calculus*, Springer, 1988.

[KW67] Hiroshi Kunita and Shinzo Watanabe, *On square integrable martingales*, Nagoya Math. J **30** (1967), no. 209-245.

[Lé40] Paul Lévy, *Le mouvement brownien plan*, Amer. J. Math. **62** (1940), 487–550.

[LCL07] Terry J. Lyons, Michael Caruana, and Thierry Lévy, *Differential equations driven by rough paths*, Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007.

[Lej09] Antoine Lejay, *Yet another introduction to rough paths*, Séminaire de Probabilités XLII, Lecture Notes in Math., vol. 1979, Springer, Berlin, 2009, pp. 1–101.

[Lej10] ______, *Controlled differential equations as Young integrals: a simple approach*, J. Differential Equations **249** (2010), no. 8, 1777–1798.

[Lej12] ______, *Global solutions to rough differential equations with unbounded vector fields*, Séminaire de Probabilités XLIV, Lecture Notes in Math., vol. 2046, Springer, Heidelberg, 2012, pp. 215–246.

[Lem83] Marc Lemieux, *On the quadratic variation of semi-martingales*, Master's thesis, University of British Columbia, 1983, supervised by Edwin A. Perkins, available at `https://circle.ubc.ca/handle/2429/23964`.

[LQ02] Terry Lyons and Zhongmin Qian, *System control and rough paths*, Oxford University Press, 2002.

[Lyo94] Terry Lyons, *Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young*, Math. Res. Lett. **1** (1994), no. 4, 451–464.

[Lyo95a] Terry J. Lyons, *The interpretation and solution of ordinary differential equations driven by rough signals*, Stochastic analysis (Ithaca, NY, 1993), Proc. Sympos. Pure Math., vol. 57, Amer. Math. Soc., Providence, RI, 1995, pp. 115–128.

[Lyo95b] ———, *Uncertain volatility and the risk-free synthesis of derivatives*, Appl. Math. Finance **2** (1995), no. 2, 117–133.

[Lyo98] ———, *Differential equations driven by rough signals*, Rev. Mat. Iberoam. **14** (1998), no. 2, 215–310.

[Mer73] Robert C. Merton, *Theory of rational option pricing*, Bell J. Econom. and Management Sci. **4** (1973), 141–183.

[Mey67] P.-A. Meyer, *Intégrales stochastiques. I, II, III, IV*, Séminaire de Probabilités (Univ. Strasbourg, Strasbourg, 1966/67), Vol. I, Springer, Berlin, 1967, pp. 72–94, 95–117, 118–141, 142–162.

[Mis08] Yuliya S. Mishura, *Stochastic calculus for fractional Brownian motion and related processes*, Lecture Notes in Mathematics, vol. 1929, Springer-Verlag, Berlin, 2008.

[MP10] Peter Mörters and Yuval Peres, *Brownian motion*, Cambridge University Press, 2010.

[MPY94] Jin Ma, Philip Protter, and Jiongmin Yong, *Solving forward-backward stochastic differential equations explicitly – a four step scheme*, Probab. Theory Relat. Fields **98** (1994), no. 3, 339–359.

[MWZZ15] Jin Ma, Zhen Wu, Detao Zhang, and Jianfeng Zhang, *On Wellposedness of Forward-Backward SDEs — A Unified Approach*, Ann. Appl. Probab. **25** (2015), no. 4.

[Nut12] Marcel Nutz, *Pathwise construction of stochastic integrals*, Electron. Commun. Probab **17** (2012), no. 24, 1–7.

[Obł04] Jan Obłój, *The Skorokhod embedding problem and its offspring*, Probab. Surv. **1** (2004), 321–390.

[OdR13] Harald Oberhauser and Goncalo dos Reis, *Root's barrier, viscosity solutions of obstacle problems and reflected FBSDEs*, Preprint arXiv:1301.3798v1 (2013).

[Per14] Nicolas Simon Perkowski, *Studies of robustness in stochastic analysis and mathematical finance*, Ph.D. thesis, Humbodt-Universität zu Berlin, 2014.

[Pes00] Goran Peskir, *The Azéma-Yor embedding in Brownian motion with drift*, High dimensional probability II. 2nd international conference, Univ. of Washington, DC, USA, August 1–6, 1999, Boston, MA: Birkhäuser, 2000, pp. 207–221.

[PP90] Etienne Pardoux and Shige Peng, *Adapted solution of a backward stochastic differential equation*, System Control Letters **14** (1990), 55–61.

[PP01] J.L. Pedersen and G. Peskir, *The Azéma-Yor embedding in non-singular diffusions*, Stochastic Processes Appl. **96** (2001), no. 2, 305–312.

[PP13] Nicolas Perkowski and David J. Prömel, *Pathwise stochastic integrals for model free finance*, to appear in Bernoulli, Preprint arXiv:1311.6187 (2013).

[PP15] ———, *Local times for typical price paths and pathwise Tanaka formulas*, Electron. J. Probab. (2015), no. 46, 1–15.

[PT99] Etienne Pardoux and Shanjian Tang, *Forward-backward stochastic differential equations and quasilinear parabolic PDEs*, Probab. Theory Relat. Fields **114** (1999), no. 2, 123–150.

[PT15] David J. Prömel and Mathias Trabs, *Rough differential driven by signals in Besov spaces*, Preprint arXiv:1506.03252 (2015).

[PW99] Shige Peng and Zhen Wu, *Fully coupled forward-backward stochastic differential equations and applications to optimal control*, SIAM J. Control Optim. **37** (1999), no. 3, 825–843.

[Roo69] D.H. Root, *The existence of certain stopping times on Brownian motion*, Ann. Math. Stat. **40** (1969), 715–718.

[Ros71] Hermann Rost, *The stopping distributions of a Markov process*, Invent. Math. **14** (1971), 1–16.

[Ros09] Mathieu Rosenbaum, *First order p-variations and Besov spaces*, Statist. Probab. Lett. **79** (2009), no. 1, 55–62.

[Roy93] Bernard Roynette, *Mouvement Brownien et espaces de Besov*, Stochastics Stochastics Rep. (1993), 221–260.

[Ruf13] Johannes Ruf, *Hedging under arbitrage*, Math. Finance **23** (2013), no. 2, 297–317.

[Ruz00] A.A. Ruzmaikina, *Stieltjes Integrals of Hölder Continuous Functions with Applications to Fractional Brownian Motion*, Journal of Statistical Physics **100** (2000), no. 5-6, 1049–1069.

[RV93] Francesco Russo and Pierre Vallois, *Forward, backward and symmetric stochastic integration*, Probab. Theory Related Fields **97** (1993), no. 3, 403–421.

[Sko61] A.V. Skorohod, *Issledovanija po teorii slučajnyh processov: (stohastičaskie differencial'nye uravnenija i predel'nye teoremy dlja processov Markova)*, Izdatel'stvo Kievskogo universiteta, 1961 (Russian).

[Sko65] A. V. Skorokhod, *Studies in the theory of random processes*, Translated from the Russian by Scripta Technica, Inc, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965.

[Son06] Dieter Sondermann, *Introduction to stochastic calculus for finance: a new didactic approach*, vol. 579, Springer, 2006.

[SS09] Christian Seel and Philipp Strack, *Gambling in Dynamic Contests*, Journal of Economic Theory (2009), Available at SSRN: http://ssrn.com/abstract=1472078.

[STZ11] H. Mete Soner, Nizar Touzi, and Jianfeng Zhang, *Quasi-sure stochastic analysis through aggregation*, Electron. J. Probab. **16** (2011), no. 67, 1844–1879.

[SV14] Tommi Sottinen and Lauri Viitasaari, *Pathwise Integrals and Itô-Tanaka Formula for Gaussian Processes*, Journal of Theoretical Probability (2014), 1–27.

[Tao11] Terence Tao, *An introduction to measure theory*, Graduate Studies in Mathematics, vol. 126, American Mathematical Society, Providence, RI, 2011.

[Tri10] Hans Triebel, *Theory of Function Spaces*, Birkhäuser Verlag, Basel, 2010, Reprint of the 1983 Edition.

[Ver09] Mark Veraar, *Correlation Inequalities and Applications to Vector-Valued Gaussian Random Variables and Fractional Brownian Motion*, Potential Analysis **30** (2009), no. 4, 341–370.

[Vov08] Vladimir Vovk, *Continuous-time trading and the emergence of volatility*, Electron. Commun. Probab. **13** (2008), 319–324.

[Vov11a] ———, *Ito calculus without probability in idealized financial markets*, Preprint arXiv:1108.0799 (2011).

[Vov11b] ———, *Rough paths in idealized financial markets*, Lith. Math. J. **51** (2011), no. 2, 274–285.

[Vov12] ———, *Continuous-time trading and the emergence of probability*, Finance Stoch. **16** (2012), no. 4, 561–609.

[Wil01] David R. E. Williams, *Path-wise solutions of stochastic differential equations driven by Lévy processes*, Rev. Mat. Iberoamericana **17** (2001), no. 2, 295–329.

[Wue80] M. Wuermli, *Lokalzeiten für Martingale*, Master's thesis, Universität Bonn, 1980, supervised by Hans Föllmer.

[Yon97] Jiongmin Yong, *Finding adapted solutions of forward-backward stochastic differential equations: Method of continuation*, Probab. Theory Relat. Fields **107** (1997), no. 4, 537–572.

[You36] Laurence C. Young, *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Math. **67** (1936), no. 1, 251–282.

[Zä98] Martina Zähle, *Integration with respect to fractal functions and stochastic calculus. I*, Probab. Theory Related Fields **111** (1998), no. 3, 333–374.

[Zä01] ———, *Integration with respect to fractal functions and stochastic calculus. II*, Math. Nachr. **225** (2001), 145–183.

[Zä05] ———, *Stochastic differential equations with fractal noise*, Math. Nachr. **278** (2005), no. 9, 1097–1106.

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 28.07.2015 David Johannes Prömel