Charles University in Prague
Faculty of Mathematics and Physics

MASTER THESIS

Bc. Prokop Šimon

Topological Support of Solutions to Stochastic Differential Equations

Department of Probability and Mathematical Statistics

Supervisor of the master thesis: Mgr. Martin Ondreját, Ph.D.
Study programme: Mathematics
Specialization: Probability, Math. Statistics and Econometrics

Prague 2016

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

Prague, 12th May 2016 Prokop Šimon

Title: Topological Support of Solutions to Stochastic Differential Equations

Author: Bc. Prokop Šimon

Department: Department of Probability and Mathematical Statistics

Supervisor: Mgr. Martin Ondreját, Ph.D., Ústav Teorie Informace a Automatizace AV ČR

Abstract: For every probability measure one can define its support as the smallest closed set that still has measure equal to one. This thesis concentrates on stochastic differential equations and the support of the distribution of their solution. Since the solution of an ordinary differential equation is a function, we deal with support of probability measure on a function space. The field was first examined by Stroock and Varadhan (1972). Other important result is among others from Gyöngy and Pröhle (1990) which is also the main source for this thesis.
The contribution lies partly in a thorough revision of the paper Gyöngy and Pröhle and supplying missing proofs, but mainly in a new result that characterizes the support of the solution in Hölder space (actually in an intersection of all Hölder function spaces for $\alpha \in (0, \frac{1}{2})$) while keeping weak assumptions on the coefficients of the equation. We require that the diffusion function has a continuous second derivative and that the drift function is only locally Lipschitz continuous and in contrast to similar results we do not need linear growth or smoothness of coefficients. Examples and possible applications of the obtained results are included, an emphasis is put on newly covered classes of equations. The thesis is written in English.

Keywords: Stochastic Differential Equations, Support of Probability Measure, Support of a Distribution in a Function Space, Hölder Spaces

Acknowledgement This thesis would have never come into existence without my advisor, Mgr. Martin Ondreját, Ph.D., towards whom I feel very grateful not only for providing a very interesting and challenging topic, but also for his valuable comments, advice and guidance. I could not wish for a better advisor.
I would also like to thank my parents Petr and Jitka, my girlfriend Kateřina and all my closed ones for supporting me during the times of my university studies and for always standing by my side.

Contents

Used Notation

SDE	Stochastic differential equation
$\mathbb{P}_\delta$	Member of $\{\mathbb{P}_\delta :\ \delta > 0\}$
$\mathbb{E}^\delta$	Expectation with respect to $\mathbb{P}_\delta$
$\xrightarrow{\mathbb{P}_\delta}$	Convergence in probability $\mathbb{P}_\delta$
supp μ	Support of measure μ
$L^p([0,T],\lambda)$	L^p space with Lebesque measure
$AC([0,T])$	Set of all absolutely continuous functions on $[0,T]$
$C([0,T])$	Set of all continuous functions on $[0,T]$
$\|\cdot\|_{sup}$	Supremum norm
$C^k([0,T]),\ k \in \mathbb{N}$	Set of all functions on $[0,T]$ with continuous derivatives up to k-th order
$C^\alpha([0,T]),\ \alpha \in (0,\frac{1}{2})$	Set of all α-Hölder functions on $[0,T]$
$\|\cdot\|_\alpha$	α-Hölder norm
$W^{s,p}([0,T])$	Fractional Sobolev function space on $[0,T]$
$\|\cdot\|_{W^{s,p}}$	Norm in fractional Sobolev space
$\tilde{M}$	Martingale part of semimartingale M
$\bar{M}$	Bounded variation part of semimartingale M
$\langle M\rangle_t$	Quadratic variation of M
$\|\bar{M}\|(t)$	Total variation of $\bar{M}$ up to time t
$i=\overline{1,d}$	$i \in \{1,..,d\}$
$u \times v$	Cartesian product of vectors u and v

Introduction

The topic of this thesis lies in the field of stochastic analysis, in particular stochastic differential equations (in the paper abbreviated to SDE).

Chapter one is devoted to explaining the problem, motivation, results and applications and is divided into four sections. It starts with the definition of a notion of SDE and related concepts, together with the description of the main concept - characterization of the support of the solution to SDE. In the second section, we compare previous results in the field and attempt to show the way and motivation that led to the definition of our problem. Next, the contribution of this work is described, together with the methods that were used to reach the desired result. Chapter one is concluded with examples of newly covered classes of stochastic differential equations.

In chapter two, we present some Lemmas that are needed for the proof of the main Theorem and include a section about spaces of functions we will be working with. We mostly present supporting results from other sources.

In the final chapter, we first state some preliminary results that were mostly not found elsewhere and then prove the approximation Theorem 3.7, which enables us to show the main Theorem 3.10.

As described in chapter one, the solution to an SDE on an interval $[0,T]$ is a continuous random process on $[0,T]$, in other words a distribution on the set of all continuous functions $C([0,T])$. Our goal is to describe, where exactly in the continuous functions the solution lies by means of approximation of the SDE by a set of deterministic ordinary differential equations. To characterize the solution of the SDE, we introduce a notion of support of distribution, which is defined as the smallest closed set such that almost all trajectories of the solution still belong to the set (i.e. the smallest set F with $\mathbb{P}_x(F) = 1$, where $\mathbb{P}_x$ is the distribution of the solution of SDE on $C([0,T])$).

Every SDE depends on two coefficients, b and σ, called the drift and the diffusion coefficients, respectively. Note that in this paper, we restrict ourselves to Wiener process as an integrator in the diffusion term, i.e. to an SDE labeled (1.1) in the paper. The idea is to remove the randomness from the equation by replacing Wiener process with suitable deterministic functions from a certain set H. We then take a set of all solutions to those approximating, deterministic, ordinary differential equations (1.4) and denote it U. The problem was first addressed in the paper Stroock and Varadhan [1972], where it was shown under certain conditions on b and σ that the support of the solution of the SDE is equal to the closure of the set U in continuous functions. The characterization was refined in many following papers, until now the best results were reached by Gyöngy and Pröhle [1990] and Ben Arous, Gradinaru, and Ledoux [1994].

This thesis pushes the quality of the characterization forward in two ways. Firstly, it reduces the conditions on the coefficients b and σ. We only require b to be locally Lipschitz continuous and σ to have a continuous second derivative, in comparison with Ben Arous et al. [1994] (both b and σ are required to be smooth functions) and Gyöngy and Pröhle [1990] (b is globally Lipschitz continuous with linear growth and σ has a continuous second derivative and is also at most of linear growth). A typical example of newly covered class of coefficients is when the function b does not have a continuous derivative and neither is of linear growth.

Secondly, the obtained characterization is also more accurate. Instead of working in the space of all continuous functions (as in Gyöngy and Pröhle [1990]) or in the space of all Hölder functions (as in Ben Arous et al. [1994]), we work in a smaller space (X, d), which is created as an intersection of all Hölder spaces $C^\alpha([0, T])$ for $\alpha \in (0, \frac{1}{2})$.

The contributions of this thesis are mostly the proof of Theorem 3.10 (via Lemma 3.9) and the approximation Theorem 3.7. The paper Gyöngy and Pröhle [1990] was thoroughly reviewed and the missing proofs were supplied as Lemmas 2.2 and 3.8. Also some results from Mackevicius [1985] and Mackevicius [1986] had to be extended in the form of Lemma 2.5.

1. Overview

Chapter one is divided into four section. Firstly, we introduce the concept of SDE and support characterization. Secondly, we give a summary and comparison of previous results in the field. In the third section, we explain our contribution and used methods and in the last section we give examples and applications of the obtained results.
We use the following notation: All constants in the proofs will be denoted by $C > 0$, even though they can change from one line to another. Parenthesis after the constant denote possible dependence (e.g. $C(p)$). When the constant is required to be independent of certain parameter, it will be noted explicitly. Recall that we chose to abbreviate the term stochastic differential equation to SDE. When considering a function space, we usually do not explicitely state the codomain, i.e. we write $C([0,T])$ instead of $C([0,T];\mathbb{R}^d)$. Note that the codomains are mostly spaces of the form $\mathbb{R}^k$ for some $k > 1$.

1.1 Concept of Stochastic Differential Equation

The motivation for considering the notion of a stochastic differential equation is the same as for ordinary (deterministic) differential equations. We would like to model a certain process or function but we can only compute its derivative and not the function itself. And while in the case of ordinary differential equations we are certain about the values of the inputs and we can measure them accurately, the SDE is designed for the cases when we admit a random term or error entering our measurements and computations. It is not surprising that the output - the solution - of the SDE reflects the randomness and is a stochastic process as well. The derivative or change of a stochastic process is called the stochastic differential and the main tool for examining stochastic differentials is nothing less than the famous Itô formula.
A typical SDE consists of two parts - the drift part with drift function b (the deterministic part of the equation) and the stochastic part with diffusion function σ (responsible for the random fluctuations of the solution). Just like its deterministic alternative, an SDE can be easily transposed into its integral form. However, both integrals contain random parts. Therefore, before we define the equation itself, we need to address an important problem of stochastic integration.
The first case is when the integrator is a stochastic process of bounded variation almost sure, which is the drift integral. We simply define the integral pathwise by the Lebesgue-Stieltjes definition for each ω (each realization) separately

$$\left(\int_0^t b(s,.)\mathrm{d}g(s,.)\right)(\omega) = \int_0^t b(s,\omega)\mathrm{d}g(s,\omega).$$

Most common case is when $\mathrm{d}g(s,\omega) = \mathrm{d}s$.
For the diffusion process, we need to take a random process that is not of bounded variation as an integrator, the most common choice being the Wiener process.

Definition 1.1 (Wiener process). *The stochastic process $\{W(t): t \geq 0\}$ is called Wiener process, if*

(i) $W(0) = 0$ *almost sure,*

(ii) $W(t_3) - W(t_2)$ *and* $W(t_2) - W(t_1)$ *are independent for all* $0 \leq t_1 < t_2 < t_3$ *(W has independent increments),*

(iii) $W(t) - W(s) \sim N(0, |t-s|)$ *(W has Gaussian increments),*

(iv) W *is continuous (i.e. almost all paths are continuous).*

The Wiener process is a continuous process that is nowhere differentiable for almost all trajectories. It is apparent that we cannot use the Lebesgue-Stieltjes approach to work with an integral where there is the Wiener process as an integrator. There is a whole theory about the construction of stochastic integral

$$\left(\int_0^t \sigma(s,.)\mathrm{d}W(s,.)\right)(\omega).$$

For our purposes it is enough to state that such integral is defined correctly for certain classes of integrands (we will not leave this class in this paper) and that the result of stochastic integration is not only a continuous stochastic process, but even (local) martingale.
If not transposed, the vectors in this paper are column vectors. The integrals of the form $\int_0^t f(s)\mathrm{d}g(s)$, where $f = (f_1, .., f_n)$ and $g, f_1, .., f_n : \mathbb{R} \to \mathbb{R}$ will be interpreted as

$$\int_0^t f(s)\mathrm{d}g(s) = \left(\int_0^t f_i(s)\mathrm{d}g(s)\right)_{i=1}^n$$

forming $n-$dimensional column vector. The SDE we consider here is of the form

$$\mathrm{d}x(t) = b(x(t))\mathrm{d}t + \sigma(x(t)) \circ \mathrm{d}W(t),\ x(0) = x_0.$$

We work with its multidimensional alternative

$$\mathrm{d}x(t) = b(x(t))\mathrm{d}t + \sum_{i=1}^{l} \sigma_i(x(t)) \circ \mathrm{d}W^i(t),\ x(0) = x_0, \tag{1.1}$$

where $W = (W^i)$ is l-dimensional Wiener process (which is defined as a vector of l independent one-dimensional Wiener processes), $b : \mathbb{R}^d \to \mathbb{R}^d$ and
$\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^l$, where $\sigma_i(x) \in \mathbb{R}^d$ is a column vector for each $i = \overline{1,l}$, with $\sigma_i^k(x)$ being its k-th element. The sign $\circ$ means that the stochastic integral is in Stratonovich's form and the equation can be transformed into more common Itô's form

$$\mathrm{d}x(t) = b(x(t))\mathrm{d}t + \sum_{i=1}^{l} \sigma_i(x(t))\mathrm{d}W^i(t) + \frac{1}{2}\sum_{i,j=1}^{l} \sigma_{i(j)}(x(t))\mathrm{d}\langle W^i, W^j\rangle_t,$$

where

$$\sigma_{i(j)}(x) = \sum_{k=1}^{l} \left(\frac{\partial}{\partial x_k} \sigma_i(x) \right) \sigma_j^k(x). \tag{1.2}$$

The term $\langle W^i, W^j \rangle_t$ denotes quadratic variation of W^i and W^j. As can be seen from the definition of multidimensional Wiener process, it holds that $\langle W^i, W^j \rangle_t = t\delta_{ij}$ and the Itô's version can be formulated as

$$\mathrm{d}x(t) = b(x(t))\mathrm{d}t + \sum_{i=1}^{l} \sigma_i(x(t))\mathrm{d}W^i(t) + \frac{1}{2} \sum_{i=1}^{l} \sigma_{i(i)}(x(t))\mathrm{d}t, \tag{1.3}$$
$$x(0) = x_0 \in \mathbb{R}^d$$

or, in integral form,

$$x(t) = x(0) + \int_0^t b(x(s))\mathrm{d}s + \sum_{i=1}^{l} \int_0^t \sigma_i(x(s))\mathrm{d}W^i(s) + \frac{1}{2} \sum_{i=1}^{l} \int_0^t \sigma_{i(i)}(x(s))\mathrm{d}s.$$

The solution $x(t)$ is a continuous stochastic process (since it is in form of an integral), in other words it is defined by a probability measure on a space of functions. For every probability measure $\mathbb{P}$, we can define its support as the smallest closed set of probability 1 (i.e. here the smallest closed set of functions that covers almost all solutions to the equation 1.1). Since the solution is a continuous process, the biggest space in which it makes sense to consider support is the space of all continuous functions $C([0,T])$ on an interval $[0,T]$, where we solve the SDE. And it is the characterization and description of the support of the solution to (1.1) in certain function spaces that plays the main role in this paper.

1.2 Previous Results

Let us recall some previous results. The main sources for this work that we directly work with are the following papers Mackevicius [1985], Mackevicius [1986] and Gyöngy and Pröhle [1990].

The work that introduced the idea of support characterization was Stroock and Varadhan [1972]. Their original motivation was to find the characterization of the support of the solution of so called martingale problem, which they managed to reduce to support characterization of a solution of a derived SDE. The result stated that if the coefficient functions satisfy that b is a bounded function, globally Lipschitz continuous and σ is a bounded function with continuous derivatives up to the second order ($\sigma \in C_b^2(\mathbb{R}^d)$), then

$$\text{supp } (\mathbb{P}_{x_0}) \;=\; \overline{U},$$

$\mathbb{P}_{x_0}$ being the probability measure defining the solution of the SDE (1.1) with initial condition $x(0) = x_0 \in \mathbb{R}^d$ and $\overline{U}$ is the closure of the set U in $(C([0,T];\mathbb{R}^d), \|.\|_{\sup})$, where

$$\|f\|_{\sup} = \sup_{t \le T} |f(t)|.$$

The set

$$U = \{x^w \in C([0,T];\mathbb{R}^d) :\ w \in H\}$$

is the set of all solutions x^w of an approximating deterministic equation

$$\mathrm{d}x^w(t) = b(x^w(t))\mathrm{d}t + \sum_{i=1}^{l} \sigma_i(x^w(t))\mathrm{d}w^i(t), \tag{1.4}$$
$$x^w(0) = x_0,$$

where $H := \{w \in C^2([0,T];\mathbb{R}^d) : w(0) = 0\ \}$. The proof used the following convergence result about conditional probability

$$\mathbb{P}\left(\|x(t) - x^w(t)\|_{sup} \le \varepsilon | \|W^i(t) - w^i(t)\|_{sup} \le \delta\right) \xrightarrow{\delta \to 0} 0. \tag{1.5}$$

One can see that the characterization of the support is a statement of the following form.

> Under certain assumptions on the coefficient b and σ and the integrator in the diffusion term (in our case Wiener process), the support of $\mathbb{P}_{x_0}$ from (1.1) is the same as the closure of the set obtained by solving deterministic version (1.4), where we assume every $w \in H$ instead of Wiener process $W(t)$ as an integrator.

What matters is how restrictive are the assumptions on b and σ (especially boundedness), how accurate is the characterization in terms of the space in which we assume the support and closure and how big is the class of possible integrators. Let us see what refinements of the first result the succeeding articles brought, starting with Mackevicius [1986]. The assumptions on b and σ were that both coefficients remained bounded and they were required to be in $C^3(\mathbb{R}^d)$ and the closure was again assumed in $C([0,T];\mathbb{R}^d)$ with the supremum norm. What changed was that the characterization was extended to a broader class of continuous semimartingales as integrators instead of only Wiener process, $Z = M + A$, M being the martingale part and A the bounded variation part. The set $H := \{w \in C^1([0,T];\mathbb{R}^d) : w(0) = 0\}$ also changed slightly, but since we assume the closure of U, it was only a matter of notation. We now cite the characterization Theorem.

Theorem 1.2. *(Mackevicius [1986]) Let the following conditions hold. Firstly,*

$$\langle Z_i, Z_j\rangle_t = \langle M_i, M_j\rangle_t = \int_0^t c_{i,j}(s,\omega)ds,\ t \ge 0,$$

where

$$\lambda(t,\omega)|\Theta|^2 \le |(c(t,\omega)\Theta,\Theta)| \le C_T|\Theta|^2$$

for all $t \in [0,T]$, $\Theta \in \mathbb{R}^{2r}$ *and some constants* C_T *and stochastic process* λ *such that* $\mathbb{E}\exp\left(k\int_0^t \lambda^{-1}(s)ds\right) < \infty$ *for all* $k > 0$ *and* $t > 0$. *Secondly,*

$$A_i(t) = \int_0^t a_i(s,\omega)ds \ \text{ for } \ |a_i(t,\omega)| \le C_T,\ \ t \in [0,T].$$

Then $\mathrm{supp}(\mathbb{P}_x) = \overline{U}$.

Although the conditions on M and A are quite restrictive, the paper introduced an important method. Instead of concentrating on the estimation of the property (1.5), it uses an absolutely continuous change of probability $\mathbb{P}$ to construct a family of probabilities $\{\mathbb{P}_\delta : \delta > 0\}$ with the use of Girsanov Theorem (see [Ikeda and Watanabe, 1980, Theorem IV. 4.1]). The approximation is then computed with respect to these probabilities.
The same idea was then used in the main source for this thesis - the paper Gyöngy and Pröhle [1990] - and the way it is used can therefore be found here as well, particularly in Lemma 3.8. In their paper, Gyöngy and Pröhle showed characterization of the same accuracy (i.e. in continuous functions) and for the same class of continuous semimartingales as integrators as in Mackevicius [1986], but they significantly lowered the conditions on b and σ - they removed the requirement of boundedness (although they still assume the first partial derivative of σ to be bounded, b to be globally Lipschitz continuous and b and σ to be of linear growth). It will be shown later in the work that some of the assumptions on b and σ required in the article are not necessary, either in such generality or at all.
We mention two other articles that used different approaches, but their results are important for comparison. They both consider only Wiener process instead of continuous semimartingale as an integrator in (1.1), but they pushed the accuracy of the support characterization forward by proving the equality in Hölder space $(C^\alpha([0,T]), \|.\|_\alpha)$ with the norm

$$\|f\|_\alpha := \sup_{0\le t\le T} |f(t)| + \sup_{s\neq t} \frac{|f(t)-f(s)|}{|t-s|^\alpha}$$

for every $\alpha \in (0, \frac{1}{2})$ instead of $C([0,T])$ with supremum norm. We will now cite their conclusions.
The first article is Millet and Sanz-Solé [1994]. They assumed the functions b and σ satisfy the following condition

> σ is of class C^2, bounded together with its partial derivative of order one and two, and b is globally Lipschitz continuous and bounded.

Note that the assumptions are identical to those in the original article Stroock and Varadhan [1972], only the set H was changed to

$$H = \{w \in AC([0,T];\mathbb{R}^d) : w(0) = 0 \text{ and } \dot{w} \in L^2([0,1];\mathbb{R}^d)\}.$$

Their result is then again

$$\text{supp } (\mathbb{P}_{x_0}) \;=\; \overline{U},$$

but we assume the support and the closure of U in $C^\alpha([0,T];\mathbb{R}^d)$ and for every fixed $\alpha \in (0, \frac{1}{2})$.
The last mentioned paper is Ben Arous, Gradinaru, and Ledoux [1994]. The support characterization is also in Hölder space and for Wiener process, but the assumptions on the coefficients are different - the functions b and σ are both required to be smooth, i.e. in $C^\infty(\mathbb{R}^d)$, but not bounded. Other assumptions and the result are the same as in the previous article.
Those were the main results in the field, how does this paper fit into the picture?

1.3 Our Contribution and Methods

In our work, we only require b to be locally Lipschitz continuous and $\sigma \in C^2(\mathbb{R}^d)$ (i.e. all partial derivatives up to the second order are continuous functions), which is in comparison with previous papers a very weak set of assumptions. These assumptions yield local uniqueness and local existence of the solutions of (1.1) and (1.4), but, unlike the assumptions from previous papers, do not guarantee existence of the solution on the whole interval $[0, T]$. Therefore, we need to additionally assume that we have unique solutions on the whole interval $[0, T]$. In other words, we need some assumptions for non-explosion. In the last section of this chapter, we give examples of coefficients that satisfy our assumptions but have not been covered in any of the previous papers.
The characterization result is much more accurate than before, we obtain the result in the space denoted X or (X, d) (d being a metric on X), which is an intersection of all Hölder spaces for $\alpha \in (0, \frac{1}{2})$, i.e.

$$X = \bigcap_{\alpha < \frac{1}{2}} C^\alpha([0, T]).$$

As stated in the introduction, we restrict ourselves to the case of Wiener process as an integrator. The characterization is again

$$\text{supp } (\mathbb{P}_{x_0}) = \overline{U},$$

where we assume the support and the closure of U in (X, d) with an inductive topology. The characterization is formulated in Theorem 3.10. There is a section called "Hölder and Fractional Sobolev Spaces" in chapter two that describes everything we need to know about the space (X, d) for our purposes.
Let us now shortly describe the methods used in our work. From the reasons described in the section "Hölder and Fractional Sobolev Spaces", it is convenient for us to work in fractional Sobolev spaces $W^{s,p}([0, T])$ rather than in Hölder spaces or the intersection space (X, d) itself. The general course of action is the same as in Gyöngy and Pröhle [1990], the only difference is that we work in the space $(W^{s,p}([0, T]), \|.\|_{W^{s,p}})$ instead of $(C([0, T]), \|.\|_{sup})$. We first define the needed relations in the norm $\|.\|_{W^{s,p}}$ instead of the supremum norm, then the [Gyöngy and Pröhle, 1990, Theorem 2.2] is reformulated as Theorem 3.7. New assumptions had to be added, but the obtained result was stronger. Most of the Lemmas from the section "Hölder and Fractional Sobolev Spaces" in chapter two and from chapter three are used during the proof of this theorem.
After proving Theorem 3.7, we can state the first part of the main result, Lemma 3.9, where the support characterization in the space $(W^{s,p}([0, T]), \|.\|_{W^{s,p}})$ is formulated. It is in the proof of this Lemma, where we need to draw many results from Mackevicius [1985] and Mackevicius [1986]. The needed results were directly cited in section "Smooth Approximation Lemmas" in chapter two. Two slight extensions of the Lemmas had to be made and are stated and proven in the same section labeled as Lemma 2.5. When the proof of Lemma 3.9 is finished, we use the equivalence of convergence in $W^{s,p}$ for all $s < \frac{1}{2}$, $p \in (2, \infty)$ and in (X, d) to prove the support characterization in (X, d). This assertion is stated in Theorem 3.10 and concludes chapter three and the whole work.

1.4 Examples

Let us now present a few examples and applications. As stated in the previous section, in order to use our results we need to verify that σ has a continuous second derivative, b is locally Lipschitz continuous and both the SDE (1.1) and the deterministic differential equation (1.4) have a global solution. Since the assumptions on b and σ already yield local existence and uniqueness, we only need to make sure that the solutions do not explode.

Remark. Recall that our SDE is either in the Stratonovich's form (1.1) or in the Itô's form (1.3) and these two forms can be interchanged by adding or removing a correction term (1.2) to or from the drift coefficient b. The problem is that in this thesis, we formulate our results for SDEs in Stratonovich's form, while the conditions that yield non-explosion are for SDEs in Itô's form. Note that if σ is a constant function, the correction term disappears and both equation are the same.

Example 1. *Let us start with a one-dimensional example and consider an equation*

$$dx(t) \;=\; b(x(t))\,dt + \sigma(x(t)) \circ dW(t),$$

or equivalently

$$dx(t) \;=\; \left(b(x(t)) + \frac{1}{2}\sigma'(x(t))\sigma(x(t)) \right) dt + \sigma(x(t))\,dW(t),$$

where $b, \sigma:\ \mathbb{R} \to \mathbb{R}$, b is locally Lipschitz continuous, $\sigma \in C^2(\mathbb{R})$ and W is a one-dimensional Wiener process. Define $\tilde{b} := b + \frac{1}{2}\sigma'\sigma$. The problem of non-explosion for local solutions is addressed for example in [Seidler, 2011, Theorems 5.2 and 5.4]. Sufficient conditions for non-explosion in one-dimensional case are

$$|\sigma(x)| \;\leq\; C(1+|x|) \;\; \text{and} \;\; \tilde{b}(x)x \;\leq\; C(1+|x|^2).$$

The condition on b is much weaker than linear growth. It is enough for b to satisfy so called one-sided linear growth condition, i.e.

$$\begin{aligned} \tilde{b}(x) &\leq C(1+x) && \text{for } x \;\geq\; 0,\\ \tilde{b}(x) &\geq C(x-1) && \text{for } x \;<\; 0 \end{aligned}$$

for some constant $C > 0$. Note that we do not require the function b to be differentiable.

Let us first restrict ourselves to the case $\sigma \equiv 1$. Then we do not need to consider the correction term and trivially $\sigma \in C^2(\mathbb{R})$ and $b = \tilde{b}$. Therefore any function b that is neither of linear growth nor smooth gives a SDE that was not covered in any of the previous articles but satisfies the assumptions of Theorem 3.10. An example of such equation is

$$dx(t) \;=\; (|x(t)| - x^3(t))\,dt + dW.$$

Example 2. *Consider again a one-dimensional equation*

$$dx(t) \;=\; b(x(t))\,dt + \sigma(x(t)) \circ dW(t),$$

or equivalently

$$dx(t) \;=\; \left(b(x(t)) + \frac{1}{2}\sigma'(x(t))\sigma(x(t))\right) dt + \sigma(x(t))dW(t).$$

This time it is not assumed that $\sigma \equiv 1$ and the correction term must be taken care of. We need that $\sigma \in C^2(\mathbb{R})$, b is locally Lipschitz continuous and $\tilde{b} := b + \frac{1}{2}\sigma'\sigma$ satisfies the one-sided linear growth condition. It is possible to show that σ need not satisfy linear growth condition, provided the drift term b is "strong enough" to compensate for the nonlinear effects. For example in equation

$$dx(t) \;=\; (|x(t)| - x^5(t))\,dt + x^2(t)\,dW(t)$$

we see that the correction term is x^3 and the function $\tilde{b} = |x(t)| - x^5(t) + x^3$ still satisfies the assumptions of one-sided linear growth secured by the strongest term $-x^5(t)$. Again, this is an example of an SDE that was not covered in any of the previous articles.

In both examples, Theorem 3.10 gives the equality between the support of the solution to the SDE and the closure of the set of solutions to all approximating equations in the space (X, d).

Example 3. *Our last example is multidimensional, it is the stochastic geodesic equation for the unit sphere, where we consider only space independent solutions. Assume $\mathbb{S}^2 = \{u \in \mathbb{R}^3 : \; |u| = 1\}$ to be the unit sphere in $\mathbb{R}^3$. Let $T\mathbb{S}^2$ be a restricted tangent bundle of $\mathbb{S}^2$*

$$T\mathbb{S}^2 = \{(u,v) : \; |u| = 1, \langle u, v\rangle_{\mathbb{R}^3} = 0, |v| = 1\} = \{(u,v) : \; |u| = 1, u \perp v, |v| = 1\},$$

where we added the constrain $|v| = 1$. Note that $T\mathbb{S}^2 \subseteq \mathbb{R}^6$.
A deterministic geodesic equation for the unit sphere is a second order differential equation of the form

$$u'' \;=\; -|u'|^2 u\;, \;\; |u| = 1,\; u(0) \perp u'(0),\; |u(0)| = |u'(0)| = 1. \tag{1.6}$$

We now add random noise through one-dimensional Wiener process W and transform (1.6) into a stochastic geodesic equation for unit sphere. Note that we are forced to write the equation in the form of stochastic differentials instead of derivatives.

$$du' \;=\; -|u'|^2 u\,dt + (u \times u') \circ dW\;, \;\; |u| = 1,\; u(0) \perp u'(0),\; |u(0)| = |v(0)| = 1. \tag{1.7}$$

The second order SDE (1.7) can be rewritten into two first order SDEs

$$dz = b(z)\,dt + \sigma(z) \circ dW, \;\; z \in T\mathbb{S}^2, \;\; z(0) \in T\mathbb{S}^2, \tag{1.8}$$

where

$$z = \begin{pmatrix} u \\ v \end{pmatrix}, \;\; b(z) = \begin{pmatrix} v \\ -|v|^2 u \end{pmatrix}, \;\; \sigma(z) = \begin{pmatrix} 0 \\ u \times v \end{pmatrix}.$$

It holds that b is continuous and $\sigma \in C^2$, which yields the local existence and uniqueness of solution to (1.8). It is shown in [Baňas, Brzeźniak, Neklyudov, Ondreját, and Prohl, 2015, Proposition 4.1] that the equation

$$dz = b(z)\,dt + \sigma(z) \circ dW, \;\; z(0) \in T\mathbb{S}^2$$

possesses global solution. Furthermore, $|v(t)| = |v(0)|$ *for every* $t \geq 0$ *almost sure, which yields that if our initial condition* $|v(0)| = 1$ *is satisfied then* $|v(t)| = 1$ *for every* $t \geq 0$ *almost sure and therefore the solution* z *stays in* $T\mathbb{S}^2$*, i.e. it is a solution to (1.8). The global existence for the apporixmating, deterministic equation*

$$u''_w \;=\; -|u'|^2 u + \dot{w}(t)(u \times u')\ , \;\; |u| = 1,\; u(0) \perp u'(0),\; |u(0)| = |v(0)| = 1$$

for

$$w \in H = \{w \in C([0,T];\mathbb{R}) : w(0) = 0,\; w \text{ is absolutely cont. and } \dot{w} \in L^2([0,1];\mathbb{R})\}$$

can be proven analogically.
Theorem 3.10 yields that

$$\text{supp } \mathbb{P}_{u,u'} \;=\; \overline{\{u_w :\; w \in H\}}^X ,$$

where $\mathbb{P}_{u,u'}$ *is the distribution of the solution to (1.8) for* $z = (u, u')$*, while at the same time*

$$\overline{\{u_w :\; w \in H\}}^X \subseteq \left\{ h : [0,T] \to T\mathbb{S}^2 :\; h(0) = \begin{pmatrix} u(0) \\ v(0) \end{pmatrix} \right\}$$

by Baňas et al. [2015]. The characterization result yields that the solution to the SDE (1.8) in $\mathbb{R}^6$ *is actually contained in three-dimensional set*

$$T\mathbb{S}^2 = \{(u,v) :\; |u| = 1, u \perp v, |v| = 1\}.$$

2. Preliminary Results

In this chapter, we first shortly comment on the article Gyöngy and Pröhle [1990], then cite the results from Mackevicius [1985] and Mackevicius [1986] and in the last section we talk about the spaces we work in - Hölder spaces, fractional Sobolev spaces and the space (X, d).

2.1 Revision of Gyöngy and Pröhle

This thesis is from a large part an extension of the paper published by Gyöngy and Pröhle [1990]. Therefore, a thorough revision of the article was carried out before the start of the original work itself. We now state the results of the revision (needed for our purposes).
There are a few minor typing errors in the paper, one that could be misleading occured at the lower part of page 67 - the integrator in the definition of $A_\delta^{ij}(t)$ should be

$$\int_0^t (m_\delta^i(s) - M_\delta^i(s))\mathrm{d}\bar{M}_\delta^j(s) \text{ instead of } \int_0^t (m_\delta^i(s) - M_\delta^i(s))\mathrm{d}\bar{M}_\delta^i(s).$$

Another thing to point out is that both in Gyöngy and Pröhle [1990] and in this work, when one requires certain property to be fulfilled for all $\delta > 0$ (typically tightness of distribution uniform in $\delta > 0$), it is required to hold only for $\delta \in (0, \delta_0)$ for some δ_0. This is because in the end, we are only interested in limit behavior as δ tends to 0.
There are two Lemmas in the article, both are stated without proof. The proof of Lemma 3.2 was elaborated using the given reference to Mackevicius [1986]. For our purposes, a slightly weaker version of the Lemma is needed. This version is stated and proved in chapter three as Lemma 3.8.
The biggest issue in the article is the Lemma 2.3, which is not applicable the way it is stated. Primarily, we cannot assume the integrand $F : [0, \infty) \times \mathbb{R}^p \to \mathbb{R}^{p\times q}$ to be bounded. Also other assumptions are stronger than necessary and since we work on a finite interval $[0, T]$, we are only able to verify the assumptions of a "local" rather than "global" version of the Lemma (but, on the other hand, we also need a weaker assertion than stated in the original paper). A more suitable version of Lemma 2.3 and its proof follows.
First, we formulate a definition that is needed to formulate the Lemma.

Definition 2.1. *For random processes $x_\delta(t)$, $y_\delta(t)$ and stopping times τ_δ, defined on a stochastic basis $\Theta_\delta = (\Omega_\delta, F_\delta, (F_{\delta t})_{t\geq 0}, \mathbb{P}_\delta)$ for every $\delta > 0$, we write*

$$x_\delta(t) \ \sim \ y_\delta(t) \ \ on \ \ [0, \tau_\delta] \ (w.r.t. \ \Theta_\delta),$$

if $\quad \lim_{\delta\to 0} \mathbb{P}_\delta(\sup_{t<\tau_\delta} |x_\delta(t) - y_\delta(t)| > \bar{\varepsilon}) = 0$ *for every* $\bar{\varepsilon} > 0$.

Lemma 2.2. *Let τ_δ be an $F_{\delta t}$-stopping time, $T > 0$ and $u_\delta(t)$ and $v_\delta(t)$ be continuous $F_{\delta t}$-adapted stochastic processes on Θ_δ for every $\delta > 0$ such that*

$$u_\delta(t) - v_\delta(t) \sim \int_0^t (F(s, u_\delta(s)) - F(s, v_\delta(s)))dS_\delta(s) \ \ on \ \ [0, T \wedge \tau_\delta] \tag{2.1}$$

with respect to Θ_δ, *where* $F\colon [0,\infty] \times \mathbb{R}^p \longrightarrow \mathbb{R}^{p\times q}$ *is locally Lipschitz continuous in* $x \in \mathbb{R}^p$, *uniformly in* $t \in [0,T]$ *and* $S_\delta(t)$ *is a continuous semimartingale in* $\mathbb{R}^q$ *for every* $\delta > 0$. *Assume that the distributions of* $\sup_{t\leq T\wedge\tau_\delta} |u_\delta(t)|$, $\sup_{t\leq T\wedge\tau_\delta} |v_\delta(t)|$, $\|\overline{S}_\delta\|(T\wedge\tau_\delta)$ *and* $\langle S_\delta\rangle(T\wedge\tau_\delta)$ *are tight, uniformly in* $\delta > 0$. *Then*

$$u_\delta(t) \sim v_\delta(t) \text{ on } [0, T\wedge\tau_\delta].$$

Proof. Our aim is to prove that

$$\forall \bar{\varepsilon} > 0 : \lim_{\delta\to 0} \mathbb{P}^\delta \left[\sup_{t\leq T\wedge\tau_\delta} |u_\delta(t) - v_\delta(t)| \geq \bar{\varepsilon} \right] = 0$$

The assumption (2.1) is equivalent to showing

$$\lim_{\delta\to 0} \mathbb{P}^\delta \left[\sup_{t\leq T\wedge\tau_\delta} |u_\delta(t) - v_\delta(t) - \int_0^t (F(s, u_\delta(s)) - F(s, v_\delta(s)))\mathrm{d}S_\delta(s)| \geq \bar{\varepsilon} \right] = 0$$

for all $\bar{\varepsilon} > 0$.
We start with some useful reductions used in [Gyöngy and Pröhle, 1990, proof of Theorem 2.2]. Define an $F_{\delta t}$-stopping time ${\tau_\delta}^L$ as

$${\tau_\delta}^L = \inf\{t \geq 0 : H_\delta(t) \geq L\} \wedge T$$

where

$$H_\delta(t) = \sup_{0\leq s\leq t} |u_\delta(s)| + \sup_{0\leq s\leq t} |v_\delta(s)| + \|\overline{S}_\delta\|(t) + \langle S_\delta\rangle(t).$$

It follows the very same way from the mentioned article using the assumptions of uniform tightness that we can suppose $H_\delta(t) < L$ for some fixed $L > 0$.
The second reduction also follows the arcticle and deals with $|u_\delta(t) - v_\delta(t)|$. We use an $F_{\delta t}$-stopping time

$$\sigma_\delta^\varepsilon = \inf\{t \geq 0 : |u_\delta(t) - v_\delta(t)| \geq \varepsilon\} \wedge {\tau_\delta}^L \wedge T$$

and as a result can assume that $|u_\delta(t) - v_\delta(t)| \leq \varepsilon$ on $[0, T\wedge\tau_\delta]$ for some fixed $\varepsilon > 0$.
Let $T > 0$, $\varepsilon > 0$ and τ_δ be fixed. First we use the Markov inequality to get expected value instead of probability.

$$\mathbb{P}^\delta \left[\sup_{t\leq T\wedge\tau_\delta} |u_\delta(t) - v_\delta(t)| \geq \bar{\varepsilon} \right] \leq \frac{1}{\bar{\varepsilon}^2} \mathbb{E}^\delta \sup_{t\leq T\wedge\tau_\delta} |u_\delta(t) - v_\delta(t)|^2$$

The main tool of the proof is from [Métivier, 1982, Lemma 29.1]. Let

$$\Phi(t) = \sup_{0\leq s\leq t} |u_\delta(s\wedge\tau_\delta) - v_\delta(s\wedge\tau_\delta)|^2 \text{ and } T_\delta(t) = \|\overline{S}_\delta\|(t) + \langle S_\delta\rangle(t).$$

Then $(T_\delta(t))_{t\geq 0}$ is a continuous, nondecreasing, adapted stochastic process for each $\delta > 0$, $T_\delta(t) \geq 0$ on $[0, T\wedge\tau_\delta]$ and by the first reduction

$$\sup_{t\leq T\wedge\tau_\delta} T_\delta(t) \leq 2L.$$

The process $\Phi(t)$ is also nondecreasing, the only thing left to prove is that for all stopping times σ such that $\sigma \leq \tau = T \wedge \tau_\delta$ it holds

$$\mathbb{E}^\delta \ \Phi(\sigma) \leq A_\delta + \rho \, \mathbb{E}^\delta \left[\int_0^\sigma \Phi(s) \mathrm{d}T_\delta(s) \right]$$

for $\rho \geq 0$ constant and $A_\delta \xrightarrow{\delta \to 0} 0$.
Let us choose a stopping time σ such that $\sigma \leq T \wedge \tau_\delta$. Then

$$\begin{aligned}
\mathbb{E}^\delta \left(\Phi(\sigma) \right) &= \mathbb{E}^\delta \sup_{0 \leq s \leq \sigma \wedge T} |u_\delta(s \wedge \tau_\delta) - v_\delta(s \wedge \tau_\delta)|^2 \\
&= \mathbb{E}^\delta \sup_{0 \leq s \leq T} |u_\delta(s \wedge \tau_\delta \wedge \sigma) - v_\delta(s \wedge \tau_\delta \wedge \sigma)|^2 \\
&\leq 2 \, \mathbb{E}^\delta \sup_{0 \leq s \leq T} | \int_0^{s \wedge \sigma \wedge \tau_\delta} (F(r, u_\delta(r)) - F(r, v_\delta(r))) \mathrm{d}S_\delta(r)|^2 \\
&+ 2 \, \mathbb{E}^\delta \sup_{0 \leq s \leq T} |u_\delta(s \wedge \tau_\delta \wedge \sigma) - v_\delta(s \wedge \tau_\delta \wedge \sigma) \\
&\quad - \int_0^{s \wedge \sigma \wedge \tau_\delta} (F(r, u_\delta(r)) - F(r, v_\delta(r))) \mathrm{d}S_\delta(r)|^2 \equiv (1) + (2).
\end{aligned}$$

Using Burkholder-Davis-Gundy inequality and the fact that F is locally Lipschitz continuous (together with the first reduction) we obtain

$$\begin{aligned}
(1) = 4 \, \mathbb{E}^\delta & \sup_{0 \leq s \leq T} | \int_0^{s \wedge \sigma \wedge \tau_\delta} (r, F(u_\delta(r)) - F(r, v_\delta(r))) \mathrm{d}\overline{S}_\delta(r)|^2 \\
&+ 4 \, \mathbb{E}^\delta \sup_{0 \leq s \leq T} | \int_0^{s \wedge \sigma \wedge \tau_\delta} (F(r, u_\delta(r)) - F(r, v_\delta(r))) \mathrm{d}\widetilde{S}_\delta(r)|^2 \\
&\leq 4 \, \mathbb{E}^\delta \sup_{0 \leq s \leq T} \int_0^{s \wedge \sigma \wedge \tau_\delta} |F(r, u_\delta(r)) - F(r, v_\delta(r))|^2 \mathrm{d}\|\overline{S}_\delta\|(r) \\
&+ 4 \, \mathbb{E}^\delta \sup_{0 \leq s \leq T} \int_0^{s \wedge \sigma \wedge \tau_\delta} |F(r, u_\delta(r)) - F(r, v_\delta(r))|^2 \mathrm{d}\langle S_\delta \rangle(r) \\
&\leq C \, \mathbb{E}^\delta \int_0^{T \wedge \sigma \wedge \tau_\delta} |u_\delta(r) - v_\delta(r)|^2 \mathrm{d}T_\delta(r) \\
&= C \, \mathbb{E}^\delta \int_0^{\sigma} |u_\delta(r) - v_\delta(r)|^2 \mathrm{d}T_\delta(r) \\
&\leq C \, \mathbb{E}^\delta \int_0^{\sigma} \sup_{u \leq r} |u_\delta(u) - v_\delta(u)|^2 \mathrm{d}T_\delta(r) = C \, \mathbb{E}^\delta \int_0^{\sigma} \Phi(r) \mathrm{d}T_\delta(r).
\end{aligned}$$

For the second term we use the following property. If it holds for a process $(X_\delta)_{\delta \geq 0}$ that $\mathbb{E}^\delta \ |X_\delta|^p \leq K$ and $X_\delta \xrightarrow{\mathbb{P}^\delta} 0$ as $\delta \to 0$, then $\forall \ a \ > 0 \ \mathbb{E}^\delta \ |X_\delta|^{p-a} \xrightarrow{\delta \to 0} 0$. For our purposes let $p = 3$, $a = 1$ and denote

$$\begin{aligned}
X_\delta(T) := 2 \, \mathbb{E}^\delta \sup_{0 \leq s \leq T} \Big| & u_\delta(s \wedge \tau_\delta \wedge \sigma) - v_\delta(s \wedge \tau_\delta \wedge \sigma) \\
& - \int_0^{s \wedge \sigma \wedge \tau_\delta} (F(r, u_\delta(r)) - F(r, v_\delta(r))) \mathrm{d}S_\delta(r) \Big|^2.
\end{aligned}$$

Then using $K > 0$ as the local Lipschitz constant for $\|x\| \vee \|y\| \leq L$ (and for all times $t \in [0, T]$) we get

$$\begin{aligned}
\mathbb{E}^\delta\ |X_\delta(T)|^3 &\leq 18\,\mathbb{E}^\delta \sup_{0\leq s\leq T} |\underbrace{u_\delta(s \wedge \tau_\delta \wedge \sigma)}_{\leq L}|^3 + \sup_{0\leq s\leq T} |\underbrace{v_\delta(s \wedge \tau_\delta \wedge \sigma)}_{\leq L}|^3 \\
&\quad + \sup_{0\leq s\leq T} |\int_0^{s\wedge\sigma\wedge\tau_\delta} (F(r, u_\delta(r)) - F(r, v_\delta(r)))\mathrm{d}S_\delta(r)|^3 \\
&\leq 18L^3 + 9\int_0^{T\wedge\sigma\wedge\tau_\delta} |(F(r, u_\delta(r)) - F(r, v_\delta(r)))|^3 \mathrm{d}\|\overline{S}_\delta\|(r)| \\
&\quad + 9C_3\,\mathbb{E}^\delta \left(\int_0^{s\wedge\sigma\wedge\tau_\delta} |(F(r, u_\delta(r)) - F(r, v_\delta(r)))\mathrm{d}\langle S_\delta\rangle(r)|\right)^{\frac{3}{2}} \\
&\leq 18L^3 + 9K^3 \int_0^{T\wedge\sigma\wedge\tau_\delta} |\underbrace{u_\delta(r) - v_\delta(r)}_{\leq L}|^3 \mathrm{d}\|\overline{S}_\delta\|(r) \\
&\quad + 9K^3C_3\left(\int_0^{T\wedge\sigma\wedge\tau_\delta} |\underbrace{u_\delta(r) - v_\delta(r)}_{\leq L}|^2 \mathrm{d}\|\overline{S}_\delta\|(r)\right)^{\frac{3}{2}} \\
&\leq 18L^3 + 9K^3(C_3+1)L^3\big(\mathbb{E}^\delta\ \underbrace{\|\overline{S}_\delta\|(T\wedge\sigma\wedge\tau_\delta)}_{\leq L} + \big(\mathbb{E}^\delta\ \underbrace{\langle S_\delta\rangle(T\wedge\sigma\wedge\tau_\delta)}_{\leq L}\big)^{\frac{3}{2}}\big) \leq C
\end{aligned}$$

Using the proposition we obtain

$$(2) = A_\delta = \mathbb{E}^\delta\ |X_\delta|^2 \xrightarrow{\delta\to 0} 0.$$

This yields

$$\mathbb{E}^\delta\left(\Phi(\sigma)\right) \leq A_\delta + \rho\,\mathbb{E}^\delta \int_0^\sigma \Phi(r)\mathrm{d}T_\delta(r),$$

from which using [Métivier, 1982, Lemma 29.1] follows that

$$\forall\, \delta > 0:\ \mathbb{E}^\delta\left(\Phi(T\wedge\tau_\delta)\right) \leq 2A_\delta \sum_{j=0}^{\lfloor 2\rho l\rfloor} (2\rho l)^j = CA_\delta, \text{ i.e.}$$

$$\mathbb{E}^\delta \sup_{0\leq t\leq T\wedge\tau_\delta} |u_\delta(t) - v_\delta(t)|^2 \leq CA_\delta \xrightarrow{\delta\to 0} 0,$$

where $\sum_{j=0}^{\lfloor 2\rho l\rfloor} (2\rho l)^j = C$ does not depend on δ.

□

2.2 Smooth Approximation Lemmas

In this part, we cite all the results from Mackevicius [1985] and Mackevicius [1986] that were used in this thesis and a generalization of one of the results needed for our purposes. The generalized Lemma is denoted 2.5, all other lemmas are direct citations. It should be noted that the proof of the generalized lemma is

based mostly on the proof of Lemma 3 from Mackevicius [1986]. It is vitally important that the constants C do not depend on $\delta > 0$ (dependence on p, T does not matter).
Suppose $\varphi \in C^1[0,1]$ is a nonnegative function such that $\varphi(0) = \varphi(1) = 0$ and $\int_0^1 \varphi(s)\mathrm{d}s = 1$ and define $\varphi^\delta(s) = \frac{1}{\delta}\varphi(\frac{s}{\delta}), s \in \mathbb{R}$. Let $z \in C([0,\infty))$, $z(0) = 0$ and $\delta > 0$, then

$$f^\delta(z,t) = \int_{-\infty}^{t} \varphi^\delta(t-s)z(s)\mathrm{d}s = \int_{t-\delta}^{t} \varphi^\delta(t-s)z(s)\mathrm{d}s$$

and

$$g^\delta(z,t) = \int_{-\infty}^{t} \varphi^\delta(t-s)(z(s)\wedge\delta^{-1})\vee(-\delta^{-1})\mathrm{d}s = \int_{t-\delta}^{t} \varphi^\delta(t-s)(z(s)\wedge\delta^{-1})\vee(-\delta^{-1})\mathrm{d}s$$

for $z(s) = z(0), \quad s < 0$.
Suppose M, N are two one-dimensional martingales such that

$$\langle M\rangle_t \; + \; \langle N\rangle_t \; \leq \; C\; t$$

for some constant $C > 0$ and that we restrict ourselves to a finite interval $[0,T]$.

Lemma 2.3 (Mackevicius [1985], Lemma 1 and Corollary 1)**.**

(i) $\mathbb{E}\left(\int_0^T |M - f^\delta(M,t)|d\|f^\delta(N,.)\|(t)\right)^p \leq C(p),\ p \geq 2$

(ii) $\mathbb{E}\left\|\int_0^t (M(s) - f^\delta(M,s))df^\delta(N,s) - \frac{1}{2}\langle M,N\rangle_t\right\|_{sup} \xrightarrow{\delta\to 0} 0$

(iii) $\mathbb{E}^\delta\left[\int_{t_1}^{t_2} |M_t - f^\delta(M,t)|d\|f^\delta(N,t)\|\right]^p \leq C|t_2 - t_1|^p$

Lemma 2.4 (Mackevicius [1986], Lemma 2 and Lemma 3)**.**

(i) $\mathbb{E}\left(\|M - f^\delta(M,t)\|^2_{sup}\right) \leq C\tilde{\varepsilon}(\delta,T),$ *where* $\tilde{\varepsilon}(\delta,T) \xrightarrow{\delta\to 0} 0$

(ii) $\mathbb{E}\left(\|M - g^\delta(M,t)\|^2_{sup}\right) \leq C\tilde{\varepsilon}(\delta,T),$ *where* $\tilde{\varepsilon}(\delta,T) \xrightarrow{\delta\to 0} 0$

(iii) $\mathbb{E}\left(\int_0^T |M - g^\delta(M,t)|d\|g^\delta(N,.)\|(t)\right)^p \leq C(p),\ p \geq 2$

(iv) $\mathbb{E}\left(\|\int_0^t (M(s) - g^\delta(M,s)) \circ d(N(s) - g^\delta(N,s))\|^2_{sup}\right) \leq C\tilde{\varepsilon}(\delta,T),$
where $\tilde{\varepsilon}(\delta,T) \xrightarrow{\delta\to 0} 0$

(v) $\left|\frac{dg^\delta(M,t)}{dt}\right| \leq C\delta^{-2}$

Now for the extension of the previous results.

Lemma 2.5. *Let M, N be two one-dimensional martingales such that $M_0 = N_0 = 0$ and $d(\langle M\rangle_t + \langle N\rangle_t) \leq C dt$ for some constant $C > 0$, where C does not depend on $\delta > 0$. Then*

(i) $\mathbb{E}^\delta\left[\sup_{t\leq T}\left|\int_0^t (M_s - g^\delta(s,M))\mathrm{d}g^\delta(s,N) - \frac{1}{2}\langle M,N\rangle_t\right|\right] \xrightarrow{\delta\to 0} 0$

(ii) $\mathbb{E}^\delta\left[\int\limits_{t_1}^{t_2} |M_t - g^\delta(M,t)| d\|g^\delta(N,t)\|\right]^p \leq C|t_2 - t_1|^p$

Proof. Let us denote $M_T^* = \sup_{t\leq T}|M_t|$ and $N_T^* = \sup_{t\leq T}|N_t|$ and A^δ the event $A^\delta := \{M_T^* \vee N_T^* > \delta^{-1}\}$. From the definitions of f^δ and g^δ it can be easily verified that the following estimates hold for every $t \in [0,T]$ and M martingale.

$$\left|\frac{\mathrm{d}f^\delta(M,t)}{\mathrm{d}t}\right| = |\dot{f}^\delta(t,M)| \leq C\delta^{-2}M_T^*,$$
$$\left|\frac{\mathrm{d}g^\delta(M,t)}{\mathrm{d}t}\right| = |\dot{g}^\delta(t,M)| \leq C\delta^{-2}M_T^*,$$
$$|M_t - f^\delta(t,M)| \leq 2M_T^*,$$
$$|M_t - g^\delta(t,M)| \leq 2M_T^*.$$

Using Hölder and Burkholder-Davis-Gundy inequalities and the fact that on the set $\{M_T^* > \delta^{-1}\}$ it holds $1 < (\delta M_T^*)^3$, we now prove an estimate that turns out to be useful for both (i) and (ii).

$$\begin{aligned}
\delta^{-2}\,\mathbb{E}^\delta\,[M_T^*N_T^*; M_T^* \vee N_T^* > \delta^{-1}]& \\
&\leq \delta^{-2}\Big(\mathbb{E}^\delta\,[M_T^*N_T^*;\, M_T^* > \delta^{-1}] + \mathbb{E}^\delta\,[M_T^*N_T^*;\, N_T^* > \delta^{-1}]\Big)\\
&\leq \delta^{-2}\Big(\mathbb{E}\,[M_T^*N_T^*(\delta M_T^*)^3] + \mathbb{E}\,[M_T^*N_T^*(\delta N_T^*)^3]\Big)\\
&\leq \delta\Big((\mathbb{E}\,[(M_T^*)^8]\,\mathbb{E}\,[(N_T^*)^2])^{\frac{1}{2}} + (\mathbb{E}\,[(M_T^*)^2]\,\mathbb{E}\,[(N_T^*)^8])^{\frac{1}{2}}\Big)\\
&\leq \delta\Big((\mathbb{E}\,[(M_T^*)^8]\,\mathbb{E}\,[(N_T^*)^2])^{\frac{1}{2}} + (\mathbb{E}\,[(M_T^*)^2]\,\mathbb{E}\,[(N_T^*)^8])^{\frac{1}{2}}\Big)\\
&\leq \delta C\Big((\mathbb{E}\,\langle M\rangle_T^4\,\mathbb{E}\,\langle N\rangle_T)^{\frac{1}{2}} + (\mathbb{E}\,\langle M\rangle_T\,\mathbb{E}\,\langle N\rangle_T^4)^{\frac{1}{2}}\Big)\\
&\leq \delta((CT)^4CT)^{\frac{1}{2}} + CT(CT)^4) \equiv C\,\widetilde{\varepsilon}(\delta,T) \xrightarrow{\delta\to 0} 0
\end{aligned} \tag{2.2}$$

Let us now prove the first assertion. We use the result from Lemma 2.3 (ii), which states

$$\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t (M_s - f^\delta(s,M))\mathrm{d}f^\delta(s,N) - \frac{1}{2}\langle M,N\rangle_t\right|\right] \xrightarrow{\delta\to 0} 0.$$

From triangle inequality and the fact, that $f^\delta(t,M) = g^\delta(t,M)$ and $f^\delta(t,N) = g^\delta(t,N)$ on the event $[M_T^* \vee N_T^* \leq \delta^{-1}]$, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\le T}\Big|\int_0^t (M_s-g^\delta(s,M))\mathrm{d}g^\delta(s,N)-\frac{1}{2}\langle M,N\rangle_t\Big|\right]\\
&\le \underbrace{\mathbb{E}\left[\sup_{t\le T}\Big|\int_0^t (M_s-g^\delta(s,M))\mathrm{d}g^\delta(s,N)-\int_0^t (M_s-f^\delta(s,M))\mathrm{d}f^\delta(s,N)\Big|\right]}_{=0 \text{ on } [M_T^*\vee N_T^*\le\delta^{-1}]}\\
&+\underbrace{\mathbb{E}\left[\sup_{t\le T}\Big|\int_0^t (M_s-f^\delta(s,M))\mathrm{d}f^\delta(s,N)-\frac{1}{2}\langle M,N\rangle_t\Big|\right]}_{\le\varepsilon_T^\delta\xrightarrow{\delta\to0}0 \text{ by Lemma 2.3}}\\
&\le \varepsilon_T^\delta+\mathbb{E}\left[\int_0^T \underbrace{|(M_s-g^\delta(s,M))|}_{\le 2M_T^*}\underbrace{|\dot g^\delta(s,N)|}_{\le C\delta^{-2}N_T^*}\,\mathrm{d}s+\int_0^T \underbrace{|(M_s-f^\delta(s,M))|}_{\le 2M_T^*}\underbrace{|\dot f^\delta(s,N)|}_{\le C\delta^{-2}N_T^*}\,\mathrm{d}s; A^\delta\right]\\
&\le \varepsilon_T^\delta+4C\delta^{-2}\,\mathbb{E}\,[M_T^*N_T^*; M_T^*\vee N_T^*>\delta^{-1}],
\end{aligned}
$$

which with the help of (2.2) yields the desired result.
The second estimate follows a similar path. We estimate the term on two events separately.

$$
\begin{aligned}
\mathbb{E}^\delta\left[\int_{t_1}^{t_2}|M_t-g^\delta(M,t)|\mathrm{d}\|g^\delta(N,t)\|\right]^p &\le C\,\mathbb{E}^\delta\left[\int_{t_1}^{t_2}|M_t-g^\delta(M,t)|\mathrm{d}\|g^\delta(N,t)\|; A^\delta\right]^p\\
&+C\,\mathbb{E}^\delta\left[\int_{t_1}^{t_2}|M_t-g^\delta(M,t)|\mathrm{d}\|g^\delta(N,t)\|; M_T^*\vee N_T^*\le\delta^{-1}\right]^p\equiv(1)+(2)
\end{aligned}
$$

The first term is straightforward, since again $f^\delta(t,M)=g^\delta(t,M)$ and $f^\delta(t,N)=g^\delta(t,N)$ on A^δ and therefore

$$
\begin{aligned}
(1)&=\mathbb{E}^\delta\left[\int_{t_1}^{t_2}|M_t-g^\delta(M,t)|\mathrm{d}\|g^\delta(N,t)\|; M_T^*\vee N_T^*\le\delta^{-1}\right]^p\\
&=\mathbb{E}^\delta\left[\int_{t_1}^{t_2}|M_t-f^\delta(M,t)|\mathrm{d}\|f^\delta(N,t)\|; M_T^*\vee N_T^*\le\delta^{-1}\right]^p\\
&\le\mathbb{E}^\delta\left[\int_{t_1}^{t_2}|M_t-f^\delta(M,t)|\mathrm{d}\|f^\delta(N,t)\|\right]^p\le C|t_2-t_1|^p,
\end{aligned}
$$

by Lemma 2.3 (iii). The second term leads to (2.2).

$$
\begin{aligned}
\mathbb{E}^\delta\left[\int_{t_1}^{t_2}|M_t-g^\delta(M,t)|\mathrm{d}\|g^\delta(N,t)\|; A^\delta\right]^p&=\mathbb{E}^\delta\left[\int_{t_1}^{t_2}\underbrace{|M_t-g^\delta(M,t)|}_{\le 2M_T^*}\underbrace{|\dot g^\delta(N,t)|}_{\le C\delta^{-2}N_T^*}\mathrm{d}t; A^\delta\right]^p\\
&\le 2C|t_2-t_1|^p\delta^{-2}\,\mathbb{E}^\delta\left[M_T^*N_T^*; M_T^*\vee N_T^*>\delta^{-1}\right]
\end{aligned}
$$

□

2.3 Hölder and Fractional Sobolev Spaces

This section sheds some light on the spaces, in which we consider both the support of the SDE solution (1.1) and the closure of the set of solutions of the approximation equation (1.4) (these two sets are then proven to be equal). As stated in chapter one, the most accurate characterization of the support was presented in the papers Millet and Sanz-Solé [1994] and Ben Arous et al. [1994], who proved the equality in Hölder space $(C^{\alpha}([0,T]), \|f\|_{\alpha})$ for any fixed $\alpha \in (0, \frac{1}{2})$. Since

$$C^{\beta}([0,T]) \subset C^{\alpha}([0,T]) \text{ for } \alpha < \beta < \frac{1}{2},$$

a natural question arises, whether it would be possible to characterize the support more accurately, namely in the space

$$X = \bigcap_{\alpha < 1/2} C^{\alpha}([0,T]).$$

It turns out that such assertion is valid, the way towards proving it is, however, not that straightforward.

2.3.1 Reasons for choosing $W^{s,p}$

Let us denote $W^{s,p}([0,T])$ a fractional Sobolev space for $s < \frac{1}{2}$ and $p \in (2, \infty)$ real, i.e.

$$W^{s,p}([0,T]) = \{f \in L^p([0,T]) : \|f\|_{W^{s,p}([0,T])} < \infty\},$$

where

$$\begin{aligned}\|f\|_{W^{s,p}([0,T])} &= \|f\|_{L^p([0,T])} + \|f\|_{I^{s,p}([0,T])} \\ &= \left(\int_0^T |f(s)|^p \mathrm{d}s\right)^{\frac{1}{p}} + \left(\int_0^T\int_0^T \frac{|f(a)-f(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a\mathrm{d}b\right)^{\frac{1}{p}},\end{aligned}$$

where we work with such (s,p) that $s = \frac{1}{p} + \varepsilon$ for some $\varepsilon > 0$. For some purposes, it will be more useful to write $1 + sp = 1 + (\frac{1}{p} + \varepsilon)p = 2 + \varepsilon p$ in the denominator. It can be shown that $\|.\|_{W^{s,p}([0,T])}$ is a norm on $W^{s,p}([0,T])$ and they together form a complete separable metric space. We give a sketch of the most direct way to come to such conclusion. For a $\sigma-$finite measure μ on $[0,T]^2$ defined as

$$\mathrm{d}\mu_{s,p} = \frac{\mathrm{d}a \ \mathrm{d}b}{|a-b|^{1+sp}},$$

and for $F(x,y) := f(x) - f(y)$ it holds that

$$\|f\|_{I^{s,p}([0,T])} = \left(\int_0^T\int_0^T |f(a)-f(b)|^p \mathrm{d}\mu\right)^{\frac{1}{p}} = \|F\|_{L^p([0,T]^2;\mu)}.$$

Clearly, $\|f\|_{W^{s,p}} = \|f\|_{L^p([0,T])} + \|F\|_{L^p([0,T]^2;\mu)}$ is a norm. Define a mapping

$$\begin{aligned}T:\ (W^{s,p}([0,T]), \|.\|_{W^{s,p}}) &\longrightarrow L^p([0,T],\lambda) \times L^p([0,T]^2, \mathrm{d}\mu_{s,p}) \\ f &\longmapsto (f, F)\end{aligned}$$

where λ denotes Lebesgue measure and we consider an additive norm on $L^p([0,T],\lambda) \times L^p([0,T]^2, \mathrm{d}\mu_{s,p})$. Then T is linear, injective and trivially isometric mapping (the norms on domain and codomain are identical), which carries the completeness and separability of

$$L^p([0,T],\lambda) \times L^p([0,T]^2, \mathrm{d}\mu_{s,p}) \text{ to } (W^{s,p}([0,T]), \|.\|_{W^{s,p}}).$$

It has already been said that we adapted methods from Mackevicius [1986] and mostly from Gyöngy and Pröhle [1990], who proved the characterization in the space of all continuous functions $(C([0,T];\mathbb{R}^d), \|.\|_{\sup})$. To prove the result in Hölder spaces and X, we chose to do most of the work in fractional Sobolev spaces $W^{s,p}([0,T])$ rather than $C^\alpha([0,T])$. This decision requires an adequate explanation. Apart from the fact that the estimations are smoother and quicker, fractional Sobolev spaces posses an important property in comparison with Hölder spaces - they are Polish, in particular separable (while Hölder spaces are not). And since the original space $(C([0,T];\mathbb{R}^d), \|.\|_{\sup})$ is separable as well, it is much more convenient to extend the results to $W^{s,p}([0,T])$ rather than $C^\alpha([0,T])$ and then return to the intersection space X.

Remark. Note that the norms in Hölder and fractional Sobolev spaces are not defined that differently. Indeed, it is possible to show that

$$\begin{aligned} \|.\|_{W^{s,p}} &\leq K\|.\|_{s-\frac{1}{p}} && \forall\, \frac{1}{p} < s < \frac{1}{2} \\ \|.\|_{\beta} &\leq K\|.\|_{W^{s,p}} && \text{for } s = \frac{1}{p} + \beta < \frac{1}{2}, \end{aligned} \tag{2.3}$$

where the second inequality is implied by the Garsia-Rodemich-Rumsey Lemma (see Garsia, Rodemich, and Rumsey [1970/1971]).

Let us now see how to obtain the characterization in the space X. Since we deal with support of some probability and closure of a set of approximating solutions, we first need to know how the closed and open sets in X are actually defined. In other words, we need a topology on X. We consider the space $(X, \mathcal{S})$, where $\mathcal{S}$ is the induced topology on X (it is the smallest topology such that the mapping from $(X, \mathcal{S})$ to $C^\beta([0,T])$ with its topology is continuous for each $\beta < \frac{1}{2}$). Define

$$\overline{X} = \bigcap_{\substack{s < \frac{1}{2} \\ p \in (2,\infty)}} W^{s,p}([0,T])$$

and consider its induced topology $\overline{\mathcal{S}}$.
With the help of (2.3) it is possible to show that the spaces X and $\overline{X}$ and their topologies $\mathcal{S}$ and $\overline{\mathcal{S}}$ coincide, in other words the spaces $(X, \mathcal{S})$ and $(\overline{X}, \overline{\mathcal{S}})$ are equal both pointwise and topologically. From now on, we will denote this one space $(X, \mathcal{S})$.
Furthermore, the inductive topology on X is metrizable. Let us choose sequences $(s_n)_{n\in\mathbb{N}}$ and $(p_n)_{n\in\mathbb{N}}$ such that $s_n \nearrow \frac{1}{2}$ and $p_n \nearrow +\infty$ (and, naturally, $s_n < \frac{1}{2}$) and define

$$d_{(s_n),(p_n)}(f,g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min\{1, \|f-g\|_{W^{s_n,p_n}}\}.$$

Since $W^{s,p} \subset W^{t,q}$ for all $t < s < \frac{1}{2}$ and $q < p$, the metrics d_{s_n,p_n} are equivalent for all choices of suitable $(s_n)_n \in \mathbb{N}$ and $(p_n)_n \in \mathbb{N}$ and they all give the same topology, which enables us to fix one choice and denote the obtained space (X, d). We do not explicitely state it, but we suppose that (X, d) is defined on the same interval as $W^{s,p}$ (usually either $[0, T]$ or $[0, \tau_\delta(\omega)]$).
It is interesting to note that even though the Hölder spaces $C^\beta([0, T])$ are not separable for any $\beta \in (0, \frac{1}{2})$, the final space (X, d) is a separable, complete, normed vector space. We now give a proof of another useful property - that the convergence in (X, d) is equal to convergence in $W^{s,p}$ for all s, p, i.e.

Lemma 2.6.

$$d(f, f_j) \xrightarrow{j\to\infty} 0 \iff \|f - f_j\|_{W^{s,p}} \xrightarrow{j\to\infty} 0 \;\; \forall s < \frac{1}{2} \; \forall p \in (2, \infty).$$

Proof.

"$\Longrightarrow$" Let $d(f, f_j) \xrightarrow{j\to\infty} 0$ and s,p be fixed. Then

$$\|f - f_j\|_{W^{s,p}} \xrightarrow{j\to\infty} 0 \iff 2^{-n} \min\{1, \|f - g\|_{W^{s,p}}\} \xrightarrow{j\to\infty} 0$$

and for all $n \in \mathbb{N}$

$$0 \le 2^{-n} \min\{1, \|f - g\|_{W^{s,p}}\} \le \sum_{n=1}^{\infty} 2^{-n} \min\{1, \|f - g\|_{W^{s,p}}\} = d(f, f_j) \xrightarrow{j\to\infty} 0$$

"$\Longleftarrow$" Let us denote $\|.\|_{W^n} := \|.\|_{W^{s_n,p_n}}$ for each pair (s_n, p_n) from the definition of the norm d. We use Lebesgue Theorem (Dominated Convergence Theorem) for counting measure ν on $\mathbb{N}$ ($\nu\{n\} = 1$ for each $n \in \mathbb{N}$, otherwise 0). Then for $f(n, j) := 2^{-n} \min\{1, \|f - g\|_{W^n}\}$ one can write

$$\sum_{n=1}^{\infty} f(n, j) = \int_{\mathbb{N}} f(n, j) \mu(\mathrm{d}n).$$

The assumptions of Lebesgue Theorem can be verified quite easily. First, $f(n, j) \xrightarrow{j\to\infty} 0$ for each $n \in \mathbb{N}$, which is implied by the assumption and

$$|f(n, j)| \le 2^{-n} =: g(n) \;\text{ and }\; \int_{\mathbb{N}} g(n, j)\mu(\mathrm{d}n) = \sum_{n=1}^{\infty} 2^{-n} = 1 < \infty.$$

By interchanging the integral and limit we get

$$0 = \sum_{n=1}^{\infty} \left[\lim_{j\to\infty} f(n, j)\right] = \lim_{j\to\infty} \left[\sum_{n=1}^{\infty} f(n, j)\right] = \lim_{j\to\infty} d(f, f_j)$$

□

2.3.2 Deterministic Estimates in $W^{s,p}$

The extension of the characterization from $W^{s,p}([0,T])$ to (X,d) is with the help of Lemma 2.6 quite straightforward, so it is clear, that most of the technical work and estimates is done in the space $W^{s,p}$. We now state some Lemmas and methods that we use in chapter three.

Remark. Suppose we have s, p fixed. How can we prove that some function f lies in the space $W^{s,p}([0,T])$?
First of all, f has to be in $L^p([0,T])$, i.e. $\|f\|_{L^p([0,T])} < \infty$. The next thing to show is that $\|f\|_{I^{s,p}([0,T])} < \infty$. A sufficient condition for our purposes turns out to be

$$|f(a) - f(b)|^p \ \leq C\ |a-b|^{\frac{p}{2}} \ \ \forall a, b \in [0,T]\ (a \neq b). \tag{2.4}$$

Then

$$\|f\|_{I^{s,p}([0,T])} = \left(\int_0^T \int_0^T \frac{|f(a)-f(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a\mathrm{d}b \right)^{\frac{1}{p}} \leq C^{\frac{1}{p}} \left(\int_0^T \int_0^T |a-b|^{\frac{p}{2}-1-sp} \mathrm{d}a\mathrm{d}b \right)^{\frac{1}{p}},$$

where the integral converges if and only if

$$\begin{aligned} \frac{p}{2} - 1 - sp &> -1 \\ \frac{p}{2} &> sp \\ \frac{1}{2} &> s = \frac{1}{p} + \varepsilon, \end{aligned}$$

which is exactly what we assume to be fulfilled by (s,p).

Lemma 2.7. *For each $s \in (0, \frac{1}{2})$ and $p \in (2, \infty)$ there exist $\lambda \in (0,1)$, $q < \infty$ and $t \in (0, \frac{1}{2})$ such that*

$$\|f\|_{W^{s,p}} \leq C \|f\|_{sup}^{1-\lambda} \|f\|_{W^{t,q}}^{\lambda}.$$

Furthemore, it holds $q < p$, $\lambda = \frac{q}{p}$ and $t = s\frac{p}{q} = \frac{1}{q} + \varepsilon\frac{p}{q}$ and $C = C(p,q)$.

Proof. The norm in fractional Sobolev space $W^{s,p}([0,T])$ can be written as $\|f\|_{W^{s,p}} = \|f\|_{L^p} + \|f\|_{I^{s,p}}$. We will deal with these two parts separately.

$$\left(\int_0^T |f(x)|^p \mathrm{d}x \right)^{\frac{1}{p}} = \left(\int_0^T \underbrace{|f(x)|^{p-q}}_{\leq \|f\|_{sup}^{p-q}} |f(x)|^q \mathrm{d}x \right)^{\frac{1}{p}} \leq \|f\|_{sup}^{1-\frac{q}{p}} \left(\int_0^T |f(x)|^q \mathrm{d}x \right)^{\frac{1}{p}}$$

yields $\|f\|_{L^p} \leq \|f\|_{sup}^{1-\lambda} \|f\|_{L^q}^{\lambda}$. For the double integral part of the norm, choose q sufficiently close to p, such that $t = s\frac{p}{q}$ is still less than $\frac{1}{2}$.

$$\begin{aligned} \int_0^T \int_0^T \frac{|f(a)-f(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a\mathrm{d}b &= \int_0^T \int_0^T \frac{|f(a)-f(b)|^{p-q} |f(a)-f(b)|^q}{|a-b|^{1+tq}} \mathrm{d}a\mathrm{d}b \leq \\ &\leq 2^{p-q} \|f\|_{sup}^{p-q} \int_0^T \int_0^T \frac{|f(a)-f(b)|^q}{|a-b|^{1+tq}} \mathrm{d}a\mathrm{d}b, \end{aligned} \tag{2.5}$$

where again

$$|f(a)-f(b)|^{p-q} \leq \Big(|f(a)|+|f(b)|\Big)^{p-q} \leq 2^{p-q} \sup_{t\in[0,T]} |f(t)|^{p-q}.$$

The equation (2.5) gives $\|f\|_{I^{s,p}} \leq C(p,q)\|f\|_{sup}^{1-\lambda}\|f\|_{I^{t,q}}^{\lambda}$ and $sp = tq$. □

Lemma 2.8. *Let $f,g \in W^{s,p}([0,T])$, then $\|fg\|_{W^{s,p}} \leq C\|f\|_{W^{s,p}}\|g\|_{W^{s,p}}$, where the constant C does not depend on the functions f,g.*

Proof. The Lemma is proved in an equivalent form $\|fg\|_{W^{s,p}}^p \leq C\|f\|_{W^{s,p}}^p\|g\|_{W^{s,p}}^p$. Choose $s \in [0,T]$ arbitrarily, it follows from Garsia-Rodemich-Rumsey Lemma (see Garsia, Rodemich, and Rumsey [1970/1971]) and Hölder inequality that

$$\begin{aligned}|f(s)| \leq |f(s) - \frac{1}{T}\int_0^T f(t)\mathrm{d}t| + |\frac{1}{T}\int_0^T f(t)\mathrm{d}t| \\ \leq \frac{1}{T}\int_0^T |f(s)-f(t)|\mathrm{d}t + \frac{1}{T}\Big(\int_0^T |f(t)|^p\mathrm{d}t\Big)^{\frac{1}{p}} \\ \leq \frac{1}{T}\int_0^T C\|f\|_{I^{s,p}}\underbrace{|t-s|^\beta}_{\leq T^\beta}\mathrm{d}t + \frac{1}{T}\|f\|_{L^p} \leq C(T,\beta,p)\|f\|_{W^{s,p}}\end{aligned}$$

with $s = \frac{1}{p} + \beta$ for $\beta > 0$.
With the help of the estimate, we are now able to prove the assertion for both parts of the $W^{s,p}$ norm. We start with the L^p norm.

$$\|fg\|_{L^p}^p = \int_0^T |f(t)g(t)|^p\mathrm{d}t \leq \int_0^T C^p\|f\|_{W^{s,p}}^p\|g\|_{W^{s,p}}^p\mathrm{d}t \leq C^pT\|f\|_{W^{s,p}}^p\|g\|_{W^{s,p}}^p$$

The double integral norm will be proven as follows

$$\begin{aligned}\|fg\|_{I^{s,p}}^p = \int_0^T\int_0^T \frac{|f(a)g(a)-f(b)g(b)|^p}{|a-b|^{1+sp}}\mathrm{d}a\mathrm{d}b \\ \leq \int_0^T\int_0^T \frac{(|f(a)||g(a)-g(b)| + |g(b)|f(a)-f(b)|)^p}{|a-b|^{1+sp}}\mathrm{d}a\mathrm{d}b \\ \leq 2^{p-1}\int_0^T\int_0^T \frac{(\|f\|_{W^{s,p}}|g(a)-g(b)|)^p + (\|g\|_{W^{s,p}}|f(a)-f(b)|)^p}{|a-b|^{1+sp}}\mathrm{d}a\mathrm{d}b \\ \leq C(\|f\|_{W^{s,p}}\|g\|_{I^{s,p}} + \|g\|_{W^{s,p}}\|f\|_{I^{s,p}}) \leq 2C\|f\|_{W^{s,p}}\|g\|_{W^{s,p}}.\end{aligned}$$

To sum up,

$$\|fg\|_{W^{s,p}}^p \leq 2^{p-1}(\|fg\|_{L^p}^p + \|fg\|_{I^{s,p}}^p) \leq C\|f\|_{W^{s,p}}^p\|g\|_{W^{s,p}}^p$$

□

3. Characterization of Support

In this chapter, the final results of the thesis are presented. After a few definitions, we prove some Lemmas for convergence in $W^{s,p}([0,T])$. Then we show the approximation theorem, which is a crucial tool for proving the main Theorem 3.10 that characterizes the support of the distribution of the solution to (1.1). First, recall the definition 2.1 from chapter two.

Definition (2.1). *For random processes $x_\delta(t)$, $y_\delta(t)$ and stopping times τ_δ, defined on a stochastic basis $\Theta_\delta = (\Omega_\delta, F_\delta, (F_{\delta t})_{t\geq 0}, \mathbb{P}_\delta)$ for every $\delta > 0$, we write*

$$x_\delta(t) \;\sim\; y_\delta(t) \;\; on \;\; [0,\tau_\delta) \; (w.r.t. \; \Theta_\delta),$$

if $\quad \lim_{\delta\to 0} \mathbb{P}_\delta(\sup_{t<\tau_\delta} |x_\delta(t) - y_\delta(t)| > \bar{\varepsilon}) = 0$ *for every* $\bar{\varepsilon} > 0$.

Definition 3.1. *For random processes $x_\delta(t)$, $y_\delta(t)$ and stopping times τ_δ, defined on a stochastic basis $\Theta_\delta = (\Omega_\delta, F_\delta, (F_{\delta t})_{t\geq 0}, \mathbb{P}_\delta)$ for every $\delta > 0$, we write*

$$x_\delta(t) \;\approx_{s,p}\; y_\delta(t) \;\; on \;\; [0,\tau_\delta) \; (w.r.t. \; \Theta_\delta),$$

if $\quad \lim_{\delta\to 0} \mathbb{P}_\delta(\|x_\delta - y_\delta\|_{W^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}) = 0$ *for every* $\bar{\varepsilon} > 0$.

Definition 3.2. *For random processes $x_\delta(t)$, $y_\delta(t)$ and stopping times τ_δ, defined on a stochastic basis $\Theta_\delta = (\Omega_\delta, F_\delta, (F_{\delta t})_{t\geq 0}, \mathbb{P}_\delta)$ for every $\delta > 0$, we write*

$$x_\delta(t) \;\approx_X\; y_\delta(t) \;\; on \;\; [0,\tau_\delta) \; (w.r.t. \; \Theta_\delta),$$

if $\quad \lim_{\delta\to 0} \mathbb{P}_\delta(d(x_\delta, y_\delta)_{([0,\tau_\delta])} > \bar{\varepsilon}) = 0$ *for every* $\bar{\varepsilon} > 0$.

3.1 Convergence in Probability in $W^{s,p}$

Lemma 3.3. *Let $n \in \mathbb{N}$, $p \in \mathbb{N}$ and $a_i \geq 0$ for $i = \overline{1,n}$. Then*

$$\left(\sum_{i=1}^{n} a_i\right)^p \leq n^{p-1} \sum_{i=1}^{n} a_i^p.$$

If we already know that $x_\delta \sim y_\delta$ on $[0,T]$, we do not need to prove the convergence in $W^{s,p}([0,T])$ to get $x_\delta \approx_{s,p} y_\delta$ on $[0,T]$. Instead, it is enough to show boundedness, as Lemma 3.4 points out.

Lemma 3.4. *Suppose that $x_\delta(t) \sim y_\delta(t)$ on $[0,\tau_\delta]$ and let us choose $s < \frac{1}{2}$ and $q < p < \infty$, q close enough to p in the sense of Lemma 2.7. Define $\lambda = \frac{q}{p}$ and $u = s\frac{p}{q}$. Then*

$$\sup_{\delta\in(0,\delta_0)} \mathbb{E}^\delta \left[\| x_\delta - y_\delta \|^q_{W^{u,q}([0,\tau_\delta])}\right] < \infty$$

for some $\delta_0 > 0$ implies

$$\mathbb{P}_\delta \left[\| x_\delta - y_\delta \|_{W^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}\right] \xrightarrow{\delta\to 0} 0 \;\; \forall \bar{\varepsilon} > 0, \quad i.e. \;\; x_\delta(t) \approx y_\delta(t) \;\; on \;\; [0,\tau_\delta].$$

Proof. The proof is based on Lemma 2.7, we prove the assertion from the definition of a limit. Let $\bar{\varepsilon} > 0$, we need to prove

$$\forall \theta > 0 \, \exists \delta_0 > 0 \, \forall \delta \in (0, \delta_0) : \; \mathbb{P}_\delta \left[\| x_\delta - y_\delta \|_{W^{s,p}} > \bar{\varepsilon} \right] \le \theta$$

Choose $\theta > 0$ arbitrarily, we need to find suitable $\delta_0 > 0$. Assume further $R > 0$, then using Chebyshev's inequality

$$\begin{aligned}
&\mathbb{P}_\delta \left[\| x_\delta - y_\delta \|_{W^{s,p}} > \bar{\varepsilon} \right] \\
&\qquad \le \mathbb{P}_\delta \left[\bar{\varepsilon} < K \| x_\delta - y_\delta \|_{sup}^{1-\lambda} \| x_\delta - y_\delta \|_{W^{u,q}}^{\lambda} ; \| x_\delta - y_\delta \|_{sup}^{1-\lambda} \le R \right] \\
&\qquad + \mathbb{P}_\delta \left[\bar{\varepsilon} < K \| x_\delta - y_\delta \|_{sup}^{1-\lambda} \| x_\delta - y_\delta \|_{W^{u,q}}^{\lambda} ; \| x_\delta - y_\delta \|_{sup}^{1-\lambda} > R \right] \\
&\qquad \le \mathbb{P}_\delta \left[\left(\frac{\bar{\varepsilon}}{RK} \right)^{\frac{1}{\lambda}} < \| x_\delta - y_\delta \|_{W^{u,q}} \right] + \underbrace{\mathbb{P}_\delta \left[\| x_\delta - y_\delta \|_{sup}^{1-\lambda} > R \right]}_{=: A_R^\delta} \\
&\qquad \le \left(\frac{RK}{\bar{\varepsilon}} \right)^{\frac{1}{\lambda}} \underbrace{\mathbb{E}^\delta \, \| x_\delta - y_\delta \|_{W^{u,q}}^{q}}_{=\text{const.}} + A_R^\delta .
\end{aligned}$$

Find R small enough such that the first term is smaller than $\frac{\theta}{2}$. Since $A_R^\delta \xrightarrow{\delta \to 0} 0$ for each $R > 0$, there is $\delta_0 > 0$ such that for all $\delta \in (0, \delta_0)$ it holds $A_R^\delta \le \frac{\theta}{2}$. □

Remark. By making use of Lemma 3.4, we never actually need to prove that $\mathbb{E}^\delta \, \| f^\delta \|_{I^{s,p}([0,\tau_\delta])}^p \xrightarrow{\delta \to 0} 0$, provided that $f^\delta \sim 0$ on $[0, \tau_\delta]$. A sufficient result is $\mathbb{E}^\delta \, \| f^\delta \|_{I^{s,p}([0,\tau_\delta])}^p \le C$, where $C > 0$ does not depend on $\delta > 0$.
Further, let us denote $a^\delta = a \wedge \tau_\delta$ and $b^\delta = b \wedge \tau_\delta$, then

$$\begin{aligned}
\mathbb{E}^\delta \, \| f^\delta \|_{I^{s,p}([0,\tau_\delta])}^p &= \mathbb{E}^\delta \int_0^{\tau_\delta} \int_0^{\tau_\delta} \frac{|f^\delta(a) - f^\delta(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a \mathrm{d}b \\
&= \int_0^T \int_0^T \frac{\mathbb{E}^\delta \, |f^\delta(a^\delta) - f^\delta(b^\delta)|^p}{|a-b|^{1+sp}} \mathrm{d}a \mathrm{d}b.
\end{aligned}$$

By (2.4), it is enough to show that

$$\mathbb{E}^\delta \, |f^\delta(a^\delta) - f^\delta(b^\delta)|^p \le K |a-b|^{\frac{p}{2}}. \tag{3.1}$$

The following Lemma is in combination with Lemma 3.4 a powerful tool for proving convergence (through boundedness) for most of the integral terms in the proof of Theorem 3.7.

Lemma 3.5. *Let τ_δ be a stopping time and $Z(t) = \int_0^t Y_s dX_s$ for Y_s continuous process such that $|Y_s| \le C$ on $[0, \tau_\delta]$ and semimartingale $X_s = \tilde{X}_s + \bar{X}_s$ such that $\langle X \rangle_a - \langle X \rangle_b \le C |a-b|^{\frac{1}{2}}$ and $\|\bar{X}\|(a) - \|\bar{X}\|(b) \le C|a-b|^{\frac{1}{2}}$ on $[0, \tau_\delta]$ defined on a stochastic basis Θ_δ, where the constant $C > 0$ does not depend on $\delta > 0$. Then*

$$\mathbb{E}^\delta \, \|Z\|_{W^{s,p}([0,\tau_\delta])} \le C < \infty \quad \textit{for each} \;\; \delta > 0.$$

Proof. As discussed in the previous remark (3.1), we only need to focus on

$$\mathbb{E}^\delta \, |Z(a^\delta) - Z(b^\delta)|^p = \mathbb{E}^\delta \left| \int_{b^\delta}^{a^\delta} Y_s \mathrm{d}X_s \right|^p$$

$$\leq 2^{p-1}\, \mathbb{E}^\delta \left| \int_{b^\delta}^{a^\delta} Y_s \mathrm{d}\tilde{X}_s \right|^p + 2^{p-1}\, \mathbb{E}^\delta \left| \int_{b^\delta}^{a^\delta} Y_s \mathrm{d}\bar{X}_s \right|^p = 2^{p-1}\left((1) + (2)\right)$$

For (1) we employ Burkholder-Davis-Gundy inequality

$$(1) \leq C_p\, \mathbb{E}^\delta \left[\int_{b^\delta}^{a^\delta} \underbrace{|Y_s|}_{\leq K} \mathrm{d}\langle \tilde{X} \rangle_s \right]^{\frac{p}{2}} \leq C_p K\, \mathbb{E}^\delta \left[\langle \tilde{X} \rangle_{a^\delta} - \langle \tilde{X} \rangle_{b^\delta} \right]^{\frac{p}{2}}$$

$$\leq C_p K^{\frac{p}{2}+1}\, \mathbb{E}^\delta \, |a^\delta - b^\delta|^{\frac{p}{2}} \leq C_p K^{\frac{p}{2}+1} |a-b|^{\frac{p}{2}}.$$

The term (2) is easier.

$$(2) \leq \mathbb{E}^\delta \left[\int_{b^\delta}^{a^\delta} \underbrace{|Y_s|}_{\leq K} \mathrm{d}\|\bar{X}\|(s) \right]^p \leq K\, \mathbb{E}^\delta \, \left[\|\bar{X}\|(a^\delta) - \|\bar{X}\|(b^\delta) \right]^p$$

$$\leq K^{p+1}\, \mathbb{E}^\delta \, |a^\delta - b^\delta|^p \leq K^{p+1} |a-b|^p.$$

□

We will mostly use Lemma 3.5 in the following form.

Lemma 3.6. *Let τ_δ be a stopping time and $Z(t) = \int_0^t Y_s dX_s$ for Y_s continuous process such that $|Y_s| \leq K$ on $[0, \tau_\delta]$ and semimartingale $X_s = \tilde{X}_s + \bar{X}_s$ such that $d\langle X \rangle_t \leq C dt$ and $d\|\bar{X}\| \leq C dt$ on $[0, \tau_\delta]$ defined on a stochastic basis Θ_δ. Then*

$$\mathbb{E}^\delta \, \|Z\|_{W^{s,p}([0,\tau_\delta])} \leq C < \infty \quad \textit{for each} \;\; \delta > 0.$$

3.2 Approximation Theorem

This is a crucial section, in which we state and prove an extension of an approximation Theorem labeled Theorem 2.2 in Gyöngy and Pröhle [1990]. We first state the assumptions, keeping the notation from the original article.
Just like in the definition 2.1, suppose we have a stochastic basis Θ_δ for each $\delta > 0$. Let $M_\delta^i(t)$ and $m_\delta^i(t)$ be continuous semimartingales and consider stochastic differential equations

$$\mathrm{d}\, x_\delta(t) = b(x_\delta(t))\mathrm{d}t + \sum_{i=1}^{l} \sigma_i(x_\delta(t)) \circ \mathrm{d}M_\delta^i(t) \;, \quad x_\delta(0) = x_0 \tag{3.2}$$

$$\mathrm{d}\, y_\delta(t) = b(y_\delta(t))\mathrm{d}t + \sum_{i=1}^{l} \sigma_i(y_\delta(t)) \circ \mathrm{d}m_\delta^i(t) \;, \quad y_\delta(0) = x_0, \tag{3.3}$$

where b and σ_i are the same as in (1.1), i.e. $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^l$. Let τ_δ be an $F_{\delta t}$-stopping time for every $\delta > 0$ such that $\tau_\delta \le T$ for some $T > 0$ and assume

(A1)

$$M_\delta(t) \approx m_\delta(t) \text{ on } [0, \tau_\delta)$$

$$R_\delta^{i,j}(t) = \int_0^t (m_\delta^i - M_\delta^i) \mathrm{d}\bar{M}_\delta^j(s) + \langle m_\delta^i - M_\delta^i, M_\delta^j \rangle_t$$

$$+ \frac{1}{2} \left(\langle M_\delta^i, M_\delta^j \rangle_t - \langle m_\delta^i, m_\delta^j \rangle_t \right) \sim 0$$

$$\|\dot{R}_\delta^{i,j}\|(t) \in L^2([0, \tau_\delta], \lambda)$$

(A2) The distributions of the random variables

$$\int_0^{\tau_\delta} |m_\delta^i - M_\delta^i| \mathrm{d}\|\bar{M}_\delta^j\|(t) \ , \ \ \langle M_\delta^i \rangle_{\tau_\delta} \ , \ \ \langle m_\delta^i \rangle_{\tau_\delta} \ , \ \ \|\bar{m}_\delta^i\|(\tau_\delta)$$

are tight, uniformly in $\delta > 0$ for every $i, j = \overline{1, l}$.

(B1) The function b is locally Lipschitz continuous.

(B2) The derivatives $\frac{\partial}{\partial x_k}\sigma_i$ and $\frac{\partial^2}{\partial x_k \partial x_j}\sigma_i$ are continuous functions on $\mathbb{R}^d$ for every $k, j = \overline{1, d}$ and $i = \overline{1, l}$.

(B3) The distributions of the stochastic process $z_\delta(t) := y_\delta(t \wedge \tau_\delta \wedge T)$ is tight in $C([0, T]; \mathbb{R}^d)$, uniformly in $\delta > 0$.

Remark (comments on the definition). The assumptions that differ from those in Gyöngy and Pröhle [1990] are $(A1)$ and $(B1)$.
The way we are going to use the assumption $\|\dot{R}_\delta^{ij}\|(t) \in L^2([0, \tau_\delta], \lambda)$ is that $\|R_\delta^{ij}\|(t) - \|R_\delta^{ij}\|(s) \le C|t - s|^{\frac{1}{2}}$ on $[0, \tau_\delta)$.
The assumption $(B1)$ implies that b is locally bounded.
The assumption $(B3)$ will not be hard to verify. It turns out that the solution y_δ of (3.3) in our cases does not depend on $\delta > 0$, which yields the tightness by itself using [Štěpán, 1987, I.7.4] (every Borel probability on a Polish space is tight). Note that this is one of the cases when we benefit from choosing the space $W^{s,p}([0, T])$ instead of $C^\alpha([0, T])$.

Theorem 3.7. *Choose $s < \frac{1}{2}$ and $p \in (2, \infty)$. Suppose the assumptions (A1), (A2) and (B1)-(B3) hold and that both equations (3.2) and (3.3) have a global solution on $[0, T]$. Then $x_\delta(t) \approx_{s,p} y_\delta(t)$ on $[0, \tau_\delta \wedge T]$ with respect to Θ_δ.*

Remark. Although we consider a stopping time τ_δ in the Theorem, we will only use it for fixed $T > 0$ (on a fixed interval $[0, T]$).
Under the assumptions, we are entitled to use the result of Theorem 2.2 in Gyöngy

and Pröhle [1990], which yields that $x_\delta \sim y_\delta$ on $[0,\tau_\delta]$. By Lemma 3.4, it is then enough to prove that

$$\sup_{\delta\in(0,\delta_0)} \mathbb{E}^\delta \left[\| x_\delta - y_\delta \|^q_{W^{u,q}([0,\tau_\delta])}\right] < \infty,$$

for certain values of (u,q) derived from (s,p). Because the conditions on (u,q) are the same as those on (s,p), we will keep the notation (s,p) instead of (u,q) and only bear in mind the transformation.

Proof. In this proof, let us denote $z_\delta(t) = x_\delta(t) - y_\delta(t)$ and assume a random space with random norm

$$\|z_\delta\|_{W^{s,p}([0,\tau_\delta])} = \Big(\int_0^{\tau_\delta} |z_\delta(s)|^p \mathrm{d}s\Big)^{\frac{1}{p}} + \Big(\int_0^{\tau_\delta}\int_0^{\tau_\delta} \frac{|z_\delta(a)-z_\delta(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a\mathrm{d}b\Big)^{\frac{1}{p}},$$

where $\tau_\delta = \tau_\delta(\omega)$ is a $F_{\delta t}$-stopping time.
As stated in chapter two, the norm in $W^{s,p}$ consists of two parts $\|f\|_{W^{s,p}} = \|f\|_{L^p} + \|f\|_{I^{s,p}}$. Choose $\bar{\varepsilon} > 0$ arbitrarily, then

$$\begin{aligned}
\mathbb{P}_\delta\left[\|z_\delta\|_{W^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}\right] &\le \mathbb{P}_\delta\left[\|z_\delta\|_{L^p([0,\tau_\delta])} > \bar{\varepsilon}\right] + \mathbb{P}_\delta\left[\|z_\delta\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}\right] \\
&= \mathbb{P}_\delta\left[\Big(\int_0^{\tau_\delta} |z_\delta(s)|^p \mathrm{d}s\Big)^{\frac{1}{p}} > \bar{\varepsilon}\right] + \mathbb{P}_\delta\left[\|z_\delta\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}\right] \\
&\le \underbrace{\mathbb{P}_\delta\left[T^{\frac{1}{p}} \sup_{t\le\tau_\delta} |z_\delta(t)| > \bar{\varepsilon}\right]}_{\xrightarrow{\delta\to 0} 0} + \mathbb{P}_\delta\left[\|z_\delta\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}\right],
\end{aligned}$$

so we only need to prove $\mathbb{P}_\delta\left[\|z_\delta\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}\right] \xrightarrow{\delta\to 0} 0$.
To begin with, we make our work more convenient by introducing certain reductions. The first two reductions are the same as in Gyöngy and Pröhle [1990], the other two are based on similar ideas. Define

$$\sigma_\delta^L = \inf\{t \ge 0\} : H_\delta(t) > L\} \wedge \tau_\delta \wedge T.$$

The article yiedls that we can suppose that $H_\delta(t) \le L$ for $t \in (0, T\wedge\tau_\delta]$ where $\tau_\delta = \tau_\delta \wedge \sigma_\delta^L$ and

$$\begin{aligned}
H_\delta(t) = |\xi_\delta| + |y_\delta(t)| + \sum_{i,j}\int_0^t |M_\delta^i - m_\delta^i| \mathrm{d}\|\bar{M}_\delta^i\|(s) \\
+ \langle M_\delta\rangle(t) + \langle m_\delta\rangle(t) + \sum_i \|\bar{m}_\delta^i\|(t).
\end{aligned}$$

The second reduction leads to the same result as in Gyöngy and Pröhle [1990], although the way to the result is different. We make use of the result of Theorem 2.2 from Gyöngy and Pröhle [1990] that gives $x_\delta(t) \sim y_\delta(t)$ on $[0,\tau_\delta)$. Define

$\sigma_\delta^{\varepsilon_1} = \inf\{t \geq 0 : |\, x_\delta - y_\delta \,| \geq \varepsilon_1\}$. Then $\forall \bar{\varepsilon} > 0$

$$\begin{aligned}
\mathbb{P}_\delta \left[\|z_\delta(t)\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon} \right] &\leq \mathbb{P}_\delta \left[\|z_\delta(t)\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}; \sup_{t \leq \tau_\delta} |x_\delta(t) - y_\delta(t)| \leq \varepsilon_1 \right] \\
&\quad + \mathbb{P}_\delta \left[\|z_\delta(t)\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}; \sup_{t \leq \tau_\delta} |x_\delta(t) - y_\delta(t)| > \varepsilon_1 \right] \\
&\leq \mathbb{P}_\delta \left[\|z_\delta(t)\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon}; \sigma_\delta^{\varepsilon_1} \geq \tau_\delta \right] + \underbrace{\mathbb{P}_\delta \left[\sup |x_\delta(t) - y_\delta(t)| > \varepsilon_1 \right]}_{=:A^\delta \xrightarrow{\delta \to 0} 0} \\
&\leq \mathbb{P}_\delta \left[\|z_\delta(t)\|_{I^{s,p}([0,\tau_\delta \wedge \sigma_\delta^{\varepsilon_1}])} > \bar{\varepsilon} \right] + A^\delta
\end{aligned}$$

Denote $\tau_\delta = \tau_\delta \wedge \sigma_\delta^{\varepsilon_1}$, we may assume that $|x_\delta(t) - y_\delta(t)| \leq \varepsilon_1$ on $[0, \tau_\delta)$.
The third reduction stems from the assumption $M_\delta(t) \sim m_\delta(t)$ on $(0, \tau_\delta \wedge T]$. The very same path as for the second reduction through the definition

$$\tau_\delta^{\varepsilon_2} = \inf\{t \geq 0\} : |M_\delta(t) - m_\delta(t)| \geq \varepsilon_2\}$$

leads us to the assumption

$$|M_\delta(t) - m_\delta(t)| \leq \varepsilon_2 \ \text{ on } \ [0, \tau_\delta),$$

where $\tau_\delta = \tau_\delta \wedge \tau_\delta^{\varepsilon_2}$.
For the last reduction, let us denote $z_\delta(t) = x_\delta(t) - y_\delta(t)$ and define $\eta_\delta^{\varepsilon_3} = \inf\{t \geq 0\} : \|z_\delta(t)\|_{I^{s,p}} > \varepsilon_3\} \wedge \tau_\delta$. Then

$$\begin{aligned}
&\mathbb{P}_\delta \left[\left(\int_0^{\tau_\delta} \int_0^{\tau_\delta} \frac{|z_\delta(a) - z_\delta(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a \mathrm{d}b \right)^{\frac{1}{p}} \geq 2\varepsilon_3 \right] \\
&\qquad \leq \mathbb{P}_\delta \left[\left(\int_0^{\eta_\delta^{\varepsilon_3}} \int_0^{\eta_\delta^{\varepsilon_3}} \frac{|z_\delta(a) - z_\delta(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a \mathrm{d}b \right)^{\frac{1}{p}} \geq \varepsilon_3 \right].
\end{aligned}$$

It is therefore sufficient to prove the theorem on $(0, \eta_\delta^{\varepsilon_3}]$ for every $\varepsilon_3 > 0$, which allows us to suppose, that

$$\|z_\delta(t)\|_{I^{s,p}} = \left(\int_0^t \int_0^t \frac{|z_\delta(a) - z_\delta(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a \mathrm{d}b \right)^{\frac{1}{p}} \leq \varepsilon_3 \text{ on } (0, \tau_\delta],$$

where $\tau_\delta := \tau_\delta \wedge \eta_\delta^{\varepsilon_3}$. Finally, we denote $\varepsilon := \varepsilon_1 \vee \varepsilon_2 \vee \varepsilon_3$.
To sum up, we added all reductions in the form of stopping times (dependent on δ and fixed constants) and denoted the final stopping time again τ_δ. We then prove that $\mathbb{P}_\delta \left[\|z_\delta(t)\|_{I^{s,p}([0,\tau_\delta])} > \bar{\varepsilon} \right] \xrightarrow{\delta \to 0} 0$ for every $\bar{\varepsilon} > 0$ and for every choice of L,ε_1,ε_2 and ε_3 positive.
We now have enough tools to deal with $\mathbb{E}^\delta \parallel x_\delta - y_\delta \parallel_{I^{s,p}}^p$ directly. As stated above, it is enough to prove that $\mathbb{E}^\delta \parallel x_\delta - y_\delta \parallel_{I^{s,p}}^p \leq C$, where C in not a function of $\delta > 0$.

Using the formulas (3.2) and (3.3), we obtain

$$\mathbb{E}^\delta \parallel x_\delta - y_\delta \parallel^p_{I^s,p} = \mathbb{E}^\delta \left\| \int_0^{\cdot} b(x_\delta(s)) - b(y_\delta(s))\mathrm{d}s \right.$$

$$+ \sum_{i=1}^{l} \int_0^{\cdot} \sigma_i(x_\delta(s))\mathrm{d}(M^i_\delta(s) - m^i_\delta(s)) + \sum_{i=1}^{l} \int_0^{\cdot} (\sigma_i(x_\delta(s)) - \sigma_i(y_\delta(s)))\mathrm{d}m^i_\delta(s)$$

$$\left. + \frac{1}{2}\sum_{i,j=1}^{l} \int_0^{\cdot} \sigma_{i(j)}(x_\delta(s))\mathrm{d}\langle M^i_\delta, M^j_\delta\rangle(s) - \frac{1}{2}\sum_{i,j=1}^{l} \int_0^{\cdot} \sigma_{i(j)}(y_\delta(s))\mathrm{d}\langle m^i_\delta, m^j_\delta\rangle(s) \right\|^p_{I^s,p}$$

As we have seen, there are assumptions on almost every term in the above expression but on $\|\bar{M}_\delta\|$, so it is not surprising, that the problematic term turns out to be $\int_0^{\cdot} \sigma_i(x_\delta(s))\mathrm{d}(M^i_\delta(s) - m^i_\delta(s))$. However, if we adapt the same idea as in Gyöngy and Pröhle [1990] - starting with stochastic integration by parts, followed by Itô's Lemma for multidimensional semimartingales, we obtain many different terms, but we will eventually be able to bound all of them independently of $\delta > 0$.
The integration by parts procedure yields the following result

$$\int_0^t \sigma_i(x_\delta(s)\mathrm{d}(M^i_\delta(s) - m^i_\delta(s)) = (M^i_\delta(t) - m^i_\delta(t))\sigma_i(x_\delta(t))$$

$$- \int_0^t (M^i_\delta(s) - m^i_\delta(s))\mathrm{d}\sigma_i(x_\delta(s)) - \langle M^i_\delta - m^i_\delta, \sigma_i(x_\delta)\rangle(t).$$

In the next step, we further expand the integrator from the integral term on the right-hand side using Itô's Lemma for multidimensional semimartingales. The next, rather technical part shows how to get rid of the term $\|\bar{M}_\delta\|$.
First, let us recall what $\mathrm{d}\, {x_\delta}^j(s)$ and $\mathrm{d}\langle {x_\delta}^r, {x_\delta}^p\rangle(s)$ are equal to

$$\mathrm{d}\, {x_\delta}^j(s) = b^j(x_\delta(s))\mathrm{d}(s) + \sum_{\alpha=1}^{l} \sigma^j_\alpha(x_\delta(s))\mathrm{d}M^\alpha_\delta(s)$$

$$+ \frac{1}{2}\sum_{\alpha,\beta=1}^{l} \underbrace{\sum_{k=1}^{d} \sigma_\beta(x_\delta(s))\frac{\partial}{\partial x_k}\sigma^j_\alpha(x_\delta(s))}_{=\sigma^j_{\alpha(\beta)}(x_\delta(s))} \mathrm{d}\langle M^\alpha_\delta, M^\beta_\delta\rangle(s)$$

$$\mathrm{d}\langle {x_\delta}^r, {x_\delta}^p\rangle(s) = \left\langle \sum_{\alpha=1}^{d} \sigma^p_\alpha(x_\delta(s))\mathrm{d}M^\alpha_\delta, \sum_{\beta=1}^{d} \sigma^r_\beta(x_\delta(s))\mathrm{d}M^\beta_\delta \right\rangle_s$$

$$= \sum_{\alpha,\beta=1}^{l} \sigma^r_\alpha(x_\delta(s))\sigma^p_\beta(x_\delta(s))\mathrm{d}\langle M^\alpha_\delta, M^\beta_\delta\rangle_s$$

Itô's Lemma then yields

$$\mathrm{d}\sigma_i(x_\delta(s)) = \sum_{j=1}^{d} \frac{\partial}{\partial x_j}\sigma_i(x_\delta(s))\mathrm{d}\, {x_\delta}^j(s) + \frac{1}{2}\sum_{r,p=1}^{d} \frac{\partial^2}{\partial x_r \partial x_p}\sigma_i(x_\delta(s))\mathrm{d}\langle {x_\delta}^r, {x_\delta}^p\rangle(s)$$

$$= \sum_{j=1}^{d}\left(\frac{\partial}{\partial x_j}\sigma_i(x_\delta(s))\right) b^j(x_\delta(s))\mathrm{d}s + \sum_{\alpha=1}^{l}\underbrace{\sum_{j=1}^{d}\left(\frac{\partial}{\partial x_j}\sigma_i(x_\delta(s))\right)\sigma_\alpha^j(x_\delta(s))}_{=\sigma_{i(\alpha)}(x_\delta(s))}\mathrm{d}M_\delta^\alpha(s)$$

$$+\frac{1}{2}\sum_{\alpha,\beta=1}^{l}\left(\sum_{j=1}^{d}\left(\frac{\partial}{\partial x_j}\sigma_i(x_\delta(s))\right)\sigma^j_{\alpha(\beta)}(x_\delta(s))\right)\mathrm{d}\langle M_\delta^\alpha, M_\delta^\beta\rangle(s)$$

$$+\frac{1}{2}\sum_{\alpha,\beta=1}^{l}\left(\sum_{r,p=1}^{d}\left(\frac{\partial^2}{\partial x_r\partial x_p}\sigma_i(x_\delta(s))\right)\sigma^r_\alpha(x_\delta(s))\sigma^p_\beta(x_\delta(s))\right)\mathrm{d}\langle M_\delta^\alpha, M_\delta^\beta\rangle_s.$$

Itô's Lemma has another useful implication. Note, that if we restrict ourselves to the martingale part of $\sigma_i(x_\delta(t))$, it follows immediately that it is equal to

$$\sigma_i(x_\delta(0)) + \sum_{j=1}^{l}\sigma_{i(j)}(x_\delta(s))\mathrm{d}\tilde{M}_\delta^j(s).$$

Recall the well-known characterization of stochastic integral. For every continuous local martingales X,M and suitable process ψ it holds that

$$\left\langle X, \int_0^{\cdot}\psi(s)\mathrm{d}M(s)\right\rangle_t = \int_0^t \psi(s)\mathrm{d}\left\langle X, M\right\rangle_s.$$

This implies (note that the minus signed changed the order of M_δ^i and m_δ^i in the quadratic variation)

$$-\left\langle \sigma_i(x_\delta(.)), M_\delta^i(.) - m_\delta^i(.)\right\rangle_t = -\left\langle \sum_{j=1}^{l}\int_0^{\cdot}\sigma_{i(j)}(x_\delta(s))\mathrm{d}\tilde{M}_\delta^j(s), M_\delta^i(.) - m_\delta^i(.)\right\rangle_t$$

$$= \sum_{j=1}^{l}\int_0^t \sigma_{i(j)}(x_\delta(s))\mathrm{d}\left\langle m_\delta^i(.) - M_\delta^i(.), M_\delta^j(.)\right\rangle_s.$$

We have expanded the term $\mathbb{E}^\delta \parallel x_\delta - y_\delta \parallel^p_{I^{s,p}}$ into many, mostly integral, terms. In the second part of the proof, we need to bound all these terms in the norm $\|.\|_{I^{s,p}}([0,\tau_\delta])$.
By the reductions and assumptions of Theorem 3.7, we can suppose, that for $t \in [0,\tau_\delta]$ it holds

$$|M_\delta^i(t) - m_\delta^i(t)| \le \varepsilon,\ \ |x_\delta(t) - y_\delta(t)| \le \varepsilon,\ \ H_\delta(t) \le L, \quad \text{furthermore}$$

$$\mathrm{d}\left\langle M_\delta\right\rangle_t \le K\mathrm{d}t,\ \ \mathrm{d}\left\langle m_\delta\right\rangle_t \le K\mathrm{d}t \ \text{ and } \sum_{i=1}^{l}\mathrm{d}\|\bar{M}_\delta^i\| \le K\mathrm{d}t.$$

The first reduction in particular implies $|\, y_\delta(t)| \leq L$, which gives

$$|\, x_\delta(t)| \leq |\, x_\delta(t) - y_\delta(t)| + |\, y_\delta(t)| \leq \varepsilon + L =: L^\varepsilon < \infty.$$

Using assumption (B2) about continuity of $\sigma_i(x)$ and its derivatives up to the second order and local boundedness of $b^j(x)$, we can conclude that there is a positive constant (denoted again C) independent of δ (since the estimates above are as well) such that

$$|\sigma_i(x_\delta(t))| \vee |\sigma_i(y_\delta(t))| \vee \left| \frac{\partial}{\partial x_k} \sigma_i(x_\delta(t)) \right| \vee \left| \frac{\partial^2}{\partial x_k \partial x_j} \sigma_i(x_\delta(t)) \right| \leq C \quad \forall i \; \forall j, k$$

on $[0, \tau_\delta]$. Also $\|\, x_\delta - y_\delta \,\|_{I^{s,p}([0,\tau_\delta])} \leq \varepsilon$. Define

$$A^{ij}_\delta(t) = \int\limits_0^t (M^i_\delta - m^i_\delta) \mathrm{d}\bar{M}^j_\delta(s) + \langle m^i_\delta - M^i_\delta, M^j_\delta \rangle_t + \frac{1}{2} \langle M^i_\delta, M^j_\delta \rangle_t,$$

i.e. $R^{i,j}_\delta(t) + \frac{1}{2} \langle m^i_\delta, m^j_\delta \rangle_t = A^{i,j}_\delta(t)$.
We can use Lemma 3.3 to split the norm into single terms. Because of the imposed reductions, we can directly apply Lemma 3.6 for most of the integral terms. The remaining integral terms are bounded as follows

$$\begin{aligned}
\mathbb{E}^\delta \, \Bigg\| \sum_{i,j=1}^{l} \int\limits_0^{\cdot} \sigma_{i(j)}(x_\delta(s)) \underbrace{(m^i_\delta - M^i_\delta) \mathrm{d}\bar{M}^j_\delta + \mathrm{d}\langle m^i_\delta - M^i_\delta, M^j_\delta \rangle_s + \frac{1}{2} \mathrm{d}\langle M^i_\delta, M^j_\delta \rangle_s}_{= \mathrm{d}A^{ij}_\delta(s)} \\
- \frac{1}{2} \sum_{i,j=1}^{l} \int\limits_0^{\cdot} \sigma_{i(j)}(y_\delta(s)) \mathrm{d}\langle m^i_\delta, m^j_\delta \rangle_s \Bigg\|^p_{W^{s,p}} \\
\leq C(p,l) \sum_{i,j=1}^{l} \mathbb{E}^\delta \, \Bigg\| \int\limits_0^{\cdot} \sigma_{i(j)}(x_\delta(s)) \mathrm{d}R^{ij}_\delta(s) \Bigg\|^p_{W^{s,p}} \\
+ C(p,l) \sum_{i,j=1}^{l} \mathbb{E}^\delta \, \Bigg\| \int\limits_0^{\cdot} \sigma_{i(j)}(x_\delta(s)) - \sigma_{i(j)}(y_\delta(s)) \mathrm{d}\langle m^i_\delta, m^j_\delta \rangle_s \Bigg\|^p_{W^{s,p}}.
\end{aligned}$$

For both of these terms, it is again possible to use Lemma 3.6. For the integrator in the first term we use the assumption $\|\dot{R}^{ij}_\delta\|(t) \in L^2([0,\tau_\delta], \mathrm{d}x)$, for the integrator in the second term, we use Kunita-Watanabe inequality to bound the cross-variation term with quadratic variation terms.
Finally, we show how to get the bound for the only non-integral term (obtained through the integration-by-parts procedure)

$$\sum_{i=1}^{l} \|(M^i_\delta(t) - m^i_\delta(t)) \sigma_i(x_\delta(t))\|^p_{W^{s,p}([0,\tau_\delta])}.$$

According to Lemma 2.8 and Hölder inequality

$$\begin{aligned}
\mathbb{E}^\delta \, \left\|(M^i_\delta(t) - m^i_\delta(t)) \sigma_i(x_\delta(t))\right\|^{\frac{p}{2}}_{I^{s,p}} &\leq C \, \mathbb{E}^\delta \, \left\|M^i_\delta(t) - m^i_\delta(t)\right\|^{\frac{p}{2}}_{I^{s,p}} \|\sigma_i(x_\delta(t))\|^{\frac{p}{2}}_{I^{s,p}} \\
&\leq C \left(\mathbb{E}^\delta \, \left\|M^i_\delta(t) - m^i_\delta(t)\right\|^p_{I^{s,p}} \right)^{\frac{1}{2}} \left(\mathbb{E}^\delta \, \|\sigma_i(x_\delta(t))\|^p_{I^{s,p}} \right)^{\frac{1}{2}}
\end{aligned}$$

Since the first term tends to 0 as $\delta \to 0$, it is certainly bounded by some constant $C > 0$, which is the same for all δ from $(0, \delta_0)$ for some $\delta_0 > 0$.
We suppose $|\, x_\delta(t)| \leq L^\varepsilon$ on $[0, \tau_\delta]$, which enables us to first make use of the fact that $\sigma_i(x)$ is locally Lipschitz continuous.

$$\mathbb{E}^\delta \ \|\sigma_i(x_\delta(t))\|^p_{I^{s,p}([0,\tau_\delta])} = \mathbb{E}^\delta \int_0^{\tau_\delta}\int_0^{\tau_\delta} \frac{|\sigma_i(x_\delta(a)) - \sigma_i(x_\delta(b))|^p}{|a-b|^{1+sp}} \,\mathrm{d}a\mathrm{d}b$$

$$\leq C(L^\varepsilon, p)\, \mathbb{E}^\delta \int_0^{\tau_\delta}\int_0^{\tau_\delta} \frac{|\, x_\delta(a) - x_\delta(b)|^p}{|a-b|^{1+sp}} \mathrm{d}a\mathrm{d}b = C(L^\varepsilon, p)\, \mathbb{E}^\delta \ \|\, x_\delta(t)\|^p_{I^{s,p}([0,\tau_\delta])}$$

It follows that

$$\mathbb{E}^\delta \ \|\, x_\delta(t)\|^p_{I^{s,p}([0,\tau_\delta])} \leq 2^{p-1}\, \mathbb{E}^\delta \ \underbrace{\|\, x_\delta(t) - y_\delta(t)\|^p_{I^{s,p}([0,\tau_\delta])}}_{\leq\ \varepsilon} + 2^{p-1}\, \mathbb{E}^\delta \ \|\, y_\delta(t)\|^p_{I^{s,p}([0,\tau_\delta])}$$

$$\leq 2^{p-1}\varepsilon + 2^{p-1} \int_0^T\int_0^T \frac{\mathbb{E}^\delta \ |\, y_\delta(a \wedge \tau_\delta) - y_\delta(b \wedge \tau_\delta)|^p}{|a-b|^{1+sp}} \,\mathrm{d}a\mathrm{d}b.$$

and

$$\mathbb{E}^\delta \ |\, y_\delta(a^\delta) - y_\delta(b^\delta)|^p = \mathbb{E}^\delta \ \Big| \int_{b^\delta}^{a^\delta} b(y_\delta(r))\mathrm{d}r + \sum_i \int_{b^\delta}^{a^\delta} \sigma_i(y_\delta(r))\mathrm{d}m^i_\delta(r)$$

$$+ \frac{1}{2}\sum_{i,j} \int_{b^\delta}^{a^\delta} \sigma_{i(j)}(y_\delta(r))\mathrm{d}\left\langle m^i_\delta, m^j_\delta \right\rangle(r) \Big| \leq C_1|a-b|^p + C_2|a-b|^{\frac{p}{2}}$$

by employing triangle inequality and Lemma 3.6. As a result, we obtain a bound $\mathbb{E}^\delta \ \|\sigma_i(x_\delta(t))\|^p_{I^{s,p}([0,\tau_\delta])} \leq \tilde{K}$. The proof of the Theorem is finished. □

3.3 Support Characterization Theorem

Having proved Theorem 3.7, we are now ready to formulate the main result we have been aspiring to. Basically, we need to verify the assumptions of Theorem 3.7 for two different pairs of semimartingales $M_\delta(t)$, $m_\delta(t)$ - once for every of the two inclusions needed for the characterization. For the more difficult inclusion, a result taken from Mackevicius [1986] (denoted Lemma 3.2 in Gyöngy and Pröhle [1990], used without proof) is emplyed. As promised in chapter one, we now state and prove a slightly different version of the Lemma that better suits our needs.

3.3.1 The Family of Approximating Probabilities

Lemma 3.8. *Let w be an absolutely continuous function such that $w(0) = 0$ and its derivative $\dot{w} \in L^2([0,T];\mathbb{R}^d)$ and $W(t)$ be a $\mathbb{P}$-Brownian motion in $\mathbb{R}^l$. Then there exists a family of probability measures $\{\mathbb{P}_\delta :\ \delta > 0\}$ on (Ω, F) such that* *(i) the measure $\mathbb{P}_\delta$ is absolutely continuous with respect to $\mathbb{P}$ and $W(t)$ is a $\mathbb{P}_\delta$-semimartingale for every $\delta > 0$,*

(ii)

$$W(t) \sim w(t) \quad \text{on} \quad [0,T] \;\; (\text{ w.r.t. } \Theta_\delta),$$
$$\int_0^t (w^i(s) - W^i(s)) d\bar{W}^j(s) \sim \frac{1}{2}\langle \tilde{W}^i_\delta, \tilde{W}^j_\delta \rangle_t \quad \text{on} \quad [0,T]$$

(iii) $\int_0^T |w^i - W^i|(s) d\|\bar{W}^j_\delta\|(s)$ is tight, uniformly in $\delta > 0$ for every $i, j = \overline{1, l}$,

where $W_\delta(t) = \tilde{W}_\delta(t) + \bar{W}_\delta(t)$ is the decomposition of $\mathbb{P}_\delta$-semimartingale $W(t)$ into martingale and bounded variation part.

Proof. The first part of the proof directly follows the second part of the article Mackevicius [1985] on pages 60 and 61. After it is shown that the process $W_\delta(t)$ (in original article $Z_\delta(t)$) is $\mathbb{P}_\delta$-semimartingale, we use Lemma 2.5 to get the assertion. Just as in the original article, we define $W_\delta(t)$ as a unique solution of the equation

$$W^i_\delta = W^i - w^i + g^\delta(W^i_\delta), \quad i = \overline{1, l}.$$

Define

$$\alpha^j_\delta(s, \omega) := \frac{\mathrm{d}w^i(s)}{\mathrm{d}s} - \frac{\mathrm{d}g^\delta(W^i_\delta, s)}{\mathrm{d}s}(\omega)$$

It follows from Lemma 2.4 that $|\frac{\mathrm{d}g^\delta(W^i_\delta,s)}{\mathrm{d}s}(\omega)| \leq C\delta^{-2}$, therefore

$$|\alpha^j_\delta(s, \omega)|^2 \leq 2\left(C^2(\delta) + |\dot{w}^i(s)|^2\right). \tag{3.4}$$

Now define a local $\mathbb{P}$-martingale

$$X_\delta(t) = \sum_{j=1}^{l} \int_0^t \alpha^j_\delta(s, \omega)\mathrm{d}W_j(s), \quad t \in [0,T] \qquad \text{and a process}$$
$$\xi_\delta(t) = \exp\left(X_\delta(t) - \tfrac{1}{2}\langle X_\delta \rangle_t\right), \quad t \in [0,T].$$

From the definition of X_δ and (3.4) it follows that

$$\mathbb{E} \exp\left(\frac{1}{2}\langle X_\delta \rangle_T\right) = \mathbb{E} \exp\left(\frac{1}{2}\sum_{i,j=1}^{l} \int_0^T \alpha^i_\delta \alpha^j_\delta \mathrm{d}\langle W_i, W_j \rangle_s\right)$$
$$= \mathbb{E} \exp\left(\frac{1}{2}\sum_{i=1}^{l} \int_0^T |\alpha^i_\delta|^2 \mathrm{d}s\right) \leq \mathbb{E} \exp\left(\sum_{i=1}^{l} \int_0^T C^2(\delta)\mathrm{d}s + \int_0^T |\dot{w}^i(s)|^2 \mathrm{d}s\right) < +\infty.$$

We have just verified the Novikov's criterion from Novikov [1973], which gives $\mathbb{E}\ \xi_\delta(T) = 1$ and enables us to use Girsanov's Theorem (see [Ikeda and Watanabe, 1980, Theorem IV. 4.1]), which lets us for each $\delta > 0$ construct a probability $\mathrm{d}\,\mathbb{P}_\delta = \xi_\delta(T)\mathrm{d}\,\mathbb{P}$ such that for all $i = \overline{1,l}$ the process

$$B^i(t) := W^i(t) - \int_0^t \alpha_\delta^i(s)\mathrm{d}s$$

is $\mathbb{P}_\delta$-Wiener process. Moreover

$$\begin{aligned}B(t) = W(t) - \left(\int_0^t \frac{\mathrm{d}w^i(s)}{\mathrm{d}s} - \frac{\mathrm{d}g^\delta(W_\delta^i, s)}{\mathrm{d}s}(\omega)\mathrm{d}s \right)_{i=1}^l \\ = W(t) - w(t) + g^\delta(W_\delta, t) = W_\delta(t).\end{aligned}$$

To conclude the first part, we showed that $W_\delta(t)$ is $\mathbb{P}_\delta$-Wiener process and $W(t)$ is $\mathbb{P}_\delta$-semimartingale that admits a decomposition

$$W(t) = \underbrace{W_\delta(t)}_{\text{martingale part}} + \underbrace{w(t) - g^\delta(W_\delta, t)}_{\text{bounded variation part}}. \tag{3.5}$$

We see that $W(t) - w(t) = W_\delta(t) - g^\delta(W_\delta, t)$. Using this property we can estimate $W(t) - w(t)$ with $W_\delta(t) - g^\delta(W_\delta, t)$, for which we can use Lemmas 2.3 and 2.4. The fact that $g^\delta(M, t)$ is a bounded variation process for any martingale M is made clear in [Mackevicius, 1985, page 345]. It is worth mentioning that since $\xi_\delta(T) > 0$, the probabilities $\mathbb{P}$ and $\mathbb{P}_\delta$ are equivalent. First we use Lemma 2.4 (ii), which yields

$$\mathbb{E}^\delta \left[\|W - w\|_{sup}^2 \right] = \mathbb{E}^\delta \left[\|W_\delta - g^\delta(W_\delta)\|_{sup}^2 \right] \xrightarrow{\delta \to 0} 0,$$

i.e. $W(t) \sim w(t)$ on $[0, T]$.
Let us now turn our attention to the second part of (ii). Note that $\mathrm{d}\bar{W}_\delta^j(s) = w - g^\delta(W_\delta)$ by (3.5). Then

$$\begin{aligned}\int_0^t (w^i(s) - W^i(s))\mathrm{d}\bar{W}_\delta^j(s) = \int_0^t -(W_\delta^i(s) - g^\delta(W_\delta^i, s))\mathrm{d}(w^j(s) - g^\delta(W_\delta^j, s)) \\ = \int_0^t (W_\delta^i(s) - g^\delta(W_\delta^i, s))\mathrm{d}g^\delta(W_\delta^j, s) - \int_0^t (W_\delta^i(s) - g^\delta(W_\delta^i, s))\mathrm{d}w^j(s) \\ = \int_0^t (W_\delta^i(s) - g^\delta(W_\delta^i, s))\mathrm{d}g^\delta(W_\delta^j, s) - \int_0^t (W_\delta^i(s) - g^\delta(W_\delta^i, s))\mathrm{d}(w^j(s) - g^\delta(w^j, s)) \\ - \int_0^t (W_\delta^i(s) - g^\delta(W_\delta^i, s))\mathrm{d}g^\delta(w^j, s) \equiv (1) + (2) + (3).\end{aligned}$$

It follows that

$$\begin{aligned}&(1) \sim \frac{1}{2}\langle W_\delta^i, W_\delta^j \rangle_t && \text{by Lemma 2.5 (i)} \\ &(2) \sim 0 && 2.4 \text{ (iv)} \\ &(3) \sim \frac{1}{2}\langle W_\delta^i, w^j \rangle_t = 0 && \text{by Lemma 2.5 (i),}\end{aligned}$$

which implies the desired result.

The only assertion left to prove is tightness of $\int_0^T |w^i - W^i|(s)\mathrm{d}\|\bar{W}_\delta^j\|(s)$, where again $\mathrm{d}\bar{W}_\delta^j(s) = w - g^\delta(W_\delta)$. Choose $K > 0$ and $0 < \delta < 1$, then

$$\begin{aligned}
&\mathbb{P}_\delta\left[\int_0^T |w^i - W^i|(s)\mathrm{d}\|\bar{W}_\delta^j\|(s) > K\right] \\
&\le \underbrace{\frac{1}{K}\mathbb{E}^\delta\left[\int_0^T |w^i - W^i|(s)\mathrm{d}\|g^\delta(W_\delta^j)\|(s)\right]}_{\le C \text{ by Lemma 2.5 (ii)}} + \frac{1}{K}\mathbb{E}^\delta\left[\int_0^T |w^i - W^i|(s)\mathrm{d}\|w^j\|(s)\right] \\
&\qquad \le \frac{C}{K} + \frac{1}{K}\mathbb{E}^\delta\left[\sup_{t\in[0,T]} |W_\delta(t) - g^\delta(W_\delta, t)|\|w^j\|(T)\right] \le \frac{1}{K}(C + L\varepsilon(1,T))
\end{aligned}$$

by Lemma 2.4 (ii) for all $\delta \in (0,1]$. For any $\bar{\varepsilon} > 0$ and every $\delta \in (0,1]$ choose $K^{\bar{\varepsilon}}$ such that

$$K^{\bar{\varepsilon}} > \frac{C + L\varepsilon(1,T)}{\bar{\varepsilon}},$$

which implies

$$\mathbb{P}_\delta\left[\int_0^T |w^i - W^i|(s)\mathrm{d}\|\bar{W}_\delta^j\|(s) > K^{\bar{\varepsilon}}\right] \le \bar{\varepsilon}.$$

□

3.3.2 Main Theorem

The environment in which we work is quite similar to the one described in the last part of Gyöngy and Pröhle [1990]. The only change we need to make is that the set H, from which the approximating functions for Wiener process arise, is a Cameron-Martin space

$$H = \{w \in AC([0,T];\mathbb{R}^l) \text{ such that } w(0) = 0 \text{ and } \dot{w} \in L^2([0,T];\mathbb{R}^l)\}.$$

For $w \in H$ consider the approximating ordinary differential equation (1.4)

$$\begin{aligned}
\mathrm{d}x^w(t) &= b(x^w(t))\mathrm{d}t + \sum_{i=1}^l \sigma_i(x^w(t))\mathrm{d}w^i(t), \\
x^w(0) &= x_0 \in \mathbb{R}^d.
\end{aligned}$$

The additional condition $\dot{w} \in L^2([0,T];\mathbb{R}^l)$ ensures that $x^w \in W^{s,p}([0,T])$ for $s < \frac{1}{2}$ and $p > 2$. It was for example also assumed in Millet and Sanz-Solé [1994], where the support characterization in Hölder space was proven.

Lemma 3.9. *Assume* $(B1)$, $(B2)$ *and that the equations (1.1) and (1.4) have global solutions on* $[0,T]$, *choose* $s \in (0,\frac{1}{2})$ *and* $p > 2$ *such that* $s = \frac{1}{p} + \varepsilon$ *for some* $\varepsilon > 0$. *Then*

$$\mathit{supp}\ \mu = \bar{U},$$

where supp μ is the topological support of the distribution of the solution (1.1) in $W^{s,p}([0,T])$ and $\bar{U}$ is the closure in $W^{s,p}([0,T])$ of the set

$$U := \{x^w \in W^{s,p}([0,T]) : w \in H\}.$$

Proof. The proof follows a similar path as in the proof of Theorem 3.1 in Gyöngy and Pröhle [1990].
Let us start with supp $\mu \subseteq \bar{U}$. Define $W_\delta(t) := f^\delta(W,t)$ for $\delta > 0$ as an approximation for Wiener process W. We would like to use the Theorem 3.7 for $M_\delta(t) := W_\delta(t)$ and $m_\delta(t) := W(t)$. We need to show that

$$W_\delta(t) \approx_{s,p} W(t) \quad \text{on} \quad [0,T].$$

First, Lemma 2.4 (ii) yields

$$W_\delta(t) \sim W(t) \quad \text{on} \quad [0,T].$$

Now set $q < p$, $u := s\frac{p}{q}$ and $\lambda = \frac{q}{p}$, Lemma 3.4 then implies that it is enough to show

$$\mathbb{E}\ \|W_\delta - W\|_{W^{u,q}([0,T])} \le \mathbb{E}\ \|W_\delta\|_{W^{u,q}([0,T])} + \mathbb{E}\ \|W\|_{W^{u,q}([0,T])} \le C < \infty. \quad (3.6)$$

For Wiener process $W(t)$, the inequality is trivial by remark 3.1, since

$$\mathbb{E}\ |W(a) - W(b)|^r \le K|a-b|^{\frac{r}{2}}.$$

for all $r > 2$. To show (3.6) for W_δ, we need to be a little more careful. We have $u = \frac{1}{q} + \eta < \frac{1}{2}$, choose $\tilde{\eta} > \eta$ such that $\tilde{u} := \frac{1}{q} + \tilde{\eta} < \frac{1}{2}$. Define $W(t) = 0$ for $t < 0$, then with the help of Garsia-Rodemich-Rumsey Lemma (see Garsia, Rodemich, and Rumsey [1970/1971])

$$|f^\delta(W,a) - f^\delta(W,b)|^q \le \left[\int_0^\delta |W(a-r) - W(b-r)|\varphi^\delta(r)\mathrm{d}r\right]^q$$
$$\le \sup_{\substack{r\in(0,\delta)\\ a>r,b>r}} |W(a-r) - W(b-r)|^q \le C\left(\int_0^T\int_0^T \frac{|W(t)-W(s)|^q}{|s-t|^{1+\tilde{u}q}}\mathrm{d}s\mathrm{d}t\right)|a-b|^{\tilde{\eta}q}. \quad (3.7)$$

Therefore $\mathbb{E}\ |W_\delta(a) - W_\delta(b)|^q \le C|a-b|^{\tilde{\eta}q}\,\mathbb{E}\ \|W\|_{\tilde{\eta},q}$, so we need $\tilde{\eta}q - 1 - \eta q > -1$, i.e. $\tilde{\eta} > \eta$, which is exactly how we chose $\tilde{\eta}$.
Regarding the other part of (A1), we can use Lemma 2.3 (iii).

$$\mathbb{E}^\delta\left[\int_a^b |W^i(s) - f^\delta(W^i,s)|\mathrm{d}\|f^\delta(W^j)\|(s)\right]^p \le C|a-b|^p$$

This estimate then implies that $\|\dot{R}^{i,j}_\delta\|(t) \in L^2([0,T],\lambda)$.
The remaining assumptions can be verified fairly easily. The fact that $R^{ij}_\delta(t) \sim 0$ follows from Lemma 2.3 (ii) and the tightness of the integral in (A2) is implied by Lemma 2.3 (i). Since we suppose (B1) and (B2) hold in this Lemma, only

the assumption (B3) is left to show. But since the equation (3.3) is the same for all $\delta > 0$, we gain the tightness of the solution with no extra effort from [Štěpán, 1987, I.7.3], as discussed in the remark after definition of the assumptions of Theorem 3.7.
Theorem 3.7 then yields that

$$x_\delta(t) \approx_{s,p} x(t) \quad \text{on} \quad [0,T].$$

When we use characterization of convergence in distribution (and consequently in probability) from [Štěpán, 1987, Theorem III.4.1 (2)] and the fact that $W_\delta \in H$, we obtain

$$\mathbb{P}\left[x(.) \in \bar{U}\right] \geq \limsup_{\delta \to 0} \mathbb{P}\left[x_\delta(.) \in \bar{U}\right] = 1,$$

which makes the first part of the proof complete.
To prove supp $\mu \supseteq \bar{U}$, we use the Theorem 3.7 with

$$M_\delta(t) := W(t) \quad \text{and} \quad m_\delta(t) := w(t)$$

(i.e. for the equations (1.1) and (1.4)) and with changing stochastic basis Θ_δ that we, together with the verification of the assumptions, obtain from Lemma 3.8, which was designed especially for this use.
As can be seen from the proof of Lemma 3.8, it holds that $W - w = W_\delta - g^\delta(W_\delta)$, where W_δ is $\mathbb{P}_\delta$-Wiener process. To show $W(t) \approx_{s,p} w(t)$, it is actually enough to show $W(t) \approx_{s,p} g^\delta(W_\delta, t)$, which enables us to proceed the same way as above. We only need to deal with g^δ instead of f^δ, but the extension is straigtforward. Define $R^\delta(x) := (x \wedge \delta^{-1}) \vee (-\delta^{-1})$, $R^\delta(x)$ is lipchitz continuous with constant 1. Then

$$|g^\delta(W,a) - g^\delta(W,b)|^q \leq \left[\int_0^\delta |R^\delta(W(a-r)) - R^\delta(W(b-r))|\varphi^\delta(r)\mathrm{d}r\right]^q$$
$$\leq C\left[\int_0^\delta |W(a-r) - W(b-r)|\varphi^\delta(r)\mathrm{d}r\right]^q$$

and the proof follows the same path as (3.7).
The relation $R_\delta^{ij} \sim 0$ follows directly from (ii) of Lemma 3.8. We need to show that

$$\mathbb{E}^\delta\left[\int_a^b |W^i(s) - g^\delta(W^i,s)|\mathrm{d}\|g^\delta(W^j)\|(s)\right]^p \leq C|a-b|^p,$$

again this time with g^δ instead of f^δ. This extension is covered by Lemma 2.5 (ii). The other assumptions are the same as above.
The crucial part of the proof is done, Theorem 3.7 yields

$$x(t) \approx_{s,p} x^w(t) \text{ on } [0,T] \text{ w.r.t. } \Theta_\delta.$$

That means $\lim_{\delta\to 0} \mathbb{P}_\delta\left(\|x - x^w\|_{W^{s,p}([0,T])} \geq \bar{\varepsilon}\right) = 0$ for each $\bar{\varepsilon} > 0$.
Hence for every $\bar{\varepsilon} > 0$: $\mathbb{P}_\delta\left(\|x - x^w\|_{W^{s,p}([0,T])} < \bar{\varepsilon}\right) > 0$ for $\delta > 0$ sufficiently small.
Since $\mathbb{P}_\delta \ll \mathbb{P}$, we have $\mathbb{P}\left(\|x - x^w\|_{W^{s,p}([0,T])} < \bar{\varepsilon}\right) > 0$.
Hence $x^w \in$ supp μ, i.e. supp $\mu \supseteq \bar{U}$, which concludes the proof. □

There is only a small step from the characterization in $W^{s,p}([0,T])$ for each (s,p) to the same result in the space (X,d).

Theorem 3.10. *Assume* $(B1)$ *and* $(B2)$ *and that the equations (1.1) and (1.4) have unique solutions on* $[0,T]$*. Then*

$$supp\ \mu = \bar{U},$$

where supp μ *is the topological support of the distribution of the solution (1.1) in* (X,d) *and* $\bar{U}$ *is the closure in* (X,d) *of the set*

$$U \ := \ \{x^w \in X : \ w \in H\}.$$

Proof. The proof of both inclusions is based on the approximation Theorem 3.7. In both cases, the Lemma 3.9 gives the convergence in $W^{s,p}([0,T])$ for each $s \in (0, \frac{1}{2})$ and $p > 2$, which is according to Lemma 2.6 equivalent to convergence in (X,d). We can then finish both parts in the same manner, using $\approx_X$ instead of $\approx_{s,p}$ and $d(.,.)$ instead of $\|.\|_{W^{s,p}([0,T])}$. □

Conclusion

The most challenging parts of the process of creating this thesis were studying the paper Gyöngy and Pröhle [1990] and adapting their method for the space (X, d) instead of $C([0, T])$. In order to obtain a stronger assertion, the assumptions of the approximating Theorem 3.7 had to be more restrictive than in the original article. But since it was possible to verify all the additional assumptions of the approximating Theorem "for free" - i.e. using only the assumptions from the original article - there were no effects on the assumptions of the main Theorem 3.10.

We even showed that the assumptions in Gyöngy and Pröhle [1990] were too restrictive by separating the assumptions needed for the support characterization itself from those that guarantee the existence and uniqueness of the solution of the examined SDE and the approximating differential equations. It is therefore possible to combine our findings with the results about existence of unique solutions to SDEs. As we argue in the section "Examples", our assumptions already give us local existence and uniqueness and so we only need to make sure the solutions of both equations do not explode in finite time.

It would certainly be possible to look for refinements of the presented results. We restricted ourselves to the case of Wiener process as an integrator, where the characterization in (X, d) is in some sense probably the best possible. However, we could assume different integrators - martingales, or even general continuous stochastic processes such as fractional Wiener process - and try to characterize the support of the distribution of the solution to the corresponding SDEs. Perhaps some interesting results could be obtained, maybe under different sets of conditions on the coefficients or in different function spaces.

Bibliography

L. Baňas, Z. Brzeźniak, M. Neklyudov, M. Ondreját, and A. Prohl. Ergodicity for a stochastic geodesic equation in the tangent bundle of the 2d sphere. *Czechoslovak Math. Journal*, 65(3):617–657, 2015.

G. Ben Arous, M. Gradinaru, and M. Ledoux. Hölder norms and the support theorem for diffusions. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 30(3):415–436, 1994.

A. M. Garsia, E. Rodemich, and H. Rumsey, Jr. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Univ. Math. J.*, 20: 565–578, 1970/1971.

I. Gyöngy and T. Pröhle. On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem. *Computers Math. Applic.*, 19(1): 65–70, 1990.

N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* First Edition. North-Holland / Kodansha, Amsterdam-Tokyo, 1980.

V. Mackevicius. Sp-Stability of solutions of symmetric stochastic differential equations. *Litovskii Matematicheskii Sbornik*, 25(4):72–84, 1985.

V. Mackevicius. Support of the solution of a stochastic differential equation. *Litovskii Matematicheskii Sbornik*, 26(1):91–98, 1986.

M. Métivier. *Semimartingales: A course on stochastic processes*, volume 2 of *de Gruyter Studies in Mathematics.* Walter de Gruyter & Co., Berlin-New York, 1982.

A. Millet and M. Sanz-Solé. A simple proof of the support theorem for diffusion processes. *Séminaire de probabilités de Strasbourg*, 28:36–48, 1994.

A. A. Novikov. On moment inequalities and identities for stochastic integrals. *Lecture Notes in Math.*, 330:333–339, 1973.

J. Seidler. *Vybrané kapitoly ze stochastické analysy.* První vydání. Matfyzpress, Praha, 2011.

D.W. Stroock and S.R.S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob.*, 3:333–359, 1972.

J. Štěpán. *Teorie pravděpodobnosti.* První vydání. Academia, Praha, 1987.