A Banach Space-Valued Stochastic Integral with respect to a Jump Process

Boxtel, A.A. van

A.A. van Boxtel

A Banach Space-Valued Stochastic Integral
with respect to a Jump Process

And an application towards stochastic abstract Cauchy problems with Lévy noise

Master's thesis, defended on September 19, 2008

Thesis advisor: Dr. O. van Gaans

Mathematisch Instituut, Universiteit Leiden

Preface

“Ik keerde mij, en zag onder de zon, dat de loop niet is der snellen, noch de strijd der helden, noch ook de spijs der wijzen, noch ook de rijkdom der verstandigen, noch ook de gunst der welwetenden, maar dat tijd en toeval aan alle dezen wedervaart;”

ECCLESIASTES 9:11, Translation: Dutch Staten-Generaal

This thesis is written by Ton van Boxtel in order to obtain an M.Sc. degree in applied mathematics from Leiden University. My work has been supervised by Onno van Gaans and concerns stochastic differential equations. To speak in layman’s terms as much as possible, stochastic differential equations describe the change in time of a system in terms of the state of the system, under the influence of random fluctuations. One can of course imagine the importance of this type of models when dealing with systems influenced by many factors that can not be controlled or deterministically calculated. Examples of this are in ample supply: behaviour of weather and climate, movement of particles or large bodies through an irregular medium, behaviour of financial instruments in a volatile market, et cetera, et cetera. By rigourous study of the underlying models, one can hope to draw useful conclusions about the behaviour of these systems. To illustrate the use of the theory, I will mention a concrete example of a system in the financial markets that can be described by a stochastic differential equation, in casu the behaviour of the so-called forward rate process of a zero coupon bond.

As the rest of this thesis is of a rather theoretic nature, it is rather difficult to precisely explain its contents to the non-mathematician. I have tried to write in such a way that the text can be understood by a Master’s student in mathematical analysis, developing the relevant theory from quite an advanced level. I will assume the reader to be familiar with basic notions and results from functional analysis, measure theory and measure theoretic probability. As for functional analysis, the first few chapters of [18] will most probably suffice, and an excellent treatment of real-valued measure theory and probability theory can be found in the very concise, yet very complete book by Williams [22]. Some semigroups might be mentioned and a more than sufficient treatment of the relevant theory can be found in [15]. Also, for the reader interested in the applications of this theory, a very complete work of reference on the financial applications of stochastic calculus (as well as of other elements of probability theory) is the quite elaborate book by Shiryaev [20]. A very comprehensive, yet hard-to-read, book on the infinite dimensional theory can be found in [13].

Contents

Introduction

In recent years, stochastic calculus, with its wide range of applications, has become quite an intensively studied area of mathematics. Especially due to the increased popularity of complicated financial instruments, such as options or more exotic contingent claims, research and education in stochastic analysis have become an integral and indispensable part of financial mathematics. Also, many other applications throughout science in areas such as biology, physics and meteorology can be found.

From a purely mathematical point of view, the study of stochastic dynamical systems is a highly interesting and involved branch of mathematics combining methods from functional analysis, probability theory, semigroup theory, differential calculus, measure theory and many more areas of mathematics to form an interesting and challenging research area. Starting with the rigourous treatments of Brownian motion by the likes of Bachelier [1] and Einstein [2], through the stochastic integration defined by Itô [3] and Stratonovich, and by the Nobel prize-winning analysis leading to the famous Black-Scholes formula [4], stochastic analysis has blossomed into the active and diverse research area it is today.

This thesis will be concerned mostly with an abstract Banach space setting for stochastic analysis, and will be centered primarily around the (inhomogeneous) stochastic abstract Cauchy problem

$$dU(t) = [AU(t) + f(t, U(t))]\, dt + B(t, U(t))dM(t) \tag{1}$$

Where U is a process dependent on a time variable t, taking values in a separable Banach space E. Generally, M will be a Lévy process. For solving this equation, one can define a mild solution using a stochastic integral with respect to the process $M(t)$. By the work of Van Neerven and Weis in [7], and in [10], this integral can be defined when $M(t)$ is a Brownian motion.

In this thesis, I will adress the question as to when a stochastic integral can be defined for a process which is allowed to have discontinuities. After developing some of the theory on integration and stochastic variables in Banach spaces and very briely reviewing some of the work that has already been done on stochastic integrals with respect to Brownian motion, I will define a stochastic integral with respect to a process making finitely many jumps in a finite time interval in section 2.3. This integral turns out to be very simple and easy to work with and existence of mild solutions for equations with jump process noise will be proven in chapter 3. In chapter 4 I will present several extensions to the theory developed up to that point in the thesis. The pièce de résistance will be presented in section 4.3, were an extension of the existence result will be made towards processes with countably many jumps.

Chapter 1

Random Variables and Processes

This chapter is devoted to developing a theory on random variables and processes in a more general, Banach-space valued setting. Notions such as integrability, measurability, expectation and conditionality are introduced in a Banach-space context. Then I will be concerned with Gaussian variables and processes, Brownian motion, Lévy processes and (super- sub- or semi-) martingales in as general a setting as possible and/or relevant. A more elaborate exposition on most of the theory can be found in the booklet for the internet seminar on stochastic evolution equations [11].

1.1 Measurability and Integrability

I start by introducing a Banach space E and a measurable space $(A, \mathscr{A})$. The Banach space E will be a separable Banach space defined over a scalar field $\mathbb{K}$, which will in general be $\mathbb{R}$ and sometimes $\mathbb{C}$. The separability is needed mostly for defining integrals, where limiting procedures play an important role. I will be using some of the notions developed in this chapter later on for nonseparable metric spaces as well. The reader should take care to verify that the proper conditions are met in each case. The fact that the real case is indeed more general follows from the observation that any complex Banach space can be viewed as a real Banach space by limiting scalar multiplication to scalars in $\mathbb{R}$. Unless otherwise stated, we will endow E with the Borel sigma algebra $\mathscr{B}(E)$, and the underlying field with the Borel sigma algebra $\mathscr{B}(\mathbb{K})$. Furthermore, I will denote by E^* the dual space of E and for $x \in E, x^* \in E^*$, either $\langle x^*, x\rangle$ or $x^*(x)$ will denote the corresponding duality pairing. Also, one has to note that the norm of a vector $x \in E$ can be written as

$$\|x\| = \sup_{\substack{x^* \in E^* \\ \|x^*\| = 1}} |\langle x^*, x\rangle| \tag{1.1}$$

1.1.1 Measurability with respect to a σ-algebra

[Measurability with respect to] First we introduce the notion of $\mathscr{A}$*-measurability*, which is just the notion of Borel measurability for functions from a measure space to a topological space:

Definition 1.1. Let T be a topological space and $(A, \mathscr{A})$ a measurable space. A function $f : A \to E$ is called $\mathscr{A}$-measurable if for all $B \in \mathscr{B}(T)$, one has $f^{-1}(B) \in \mathscr{A}$.

Next we introduce a notion more specific to our context, the notion of *strong measurability*, for this we need the notion of a *simple function*:

Definition 1.2. A function $f : A \to E$ is called $\mathscr{A}$-simple if it can be written as the finite sum $\sum_{n=1}^{N} 1_{A_n} x_n$, where every A_n is an element of $\mathscr{A}$and every $x_n \in E$

Now we can say that a function is strongly $\mathscr{A}$-measurable if it is the pointwise limit of $\mathscr{A}$-simple functions. Now we have a version of the so called *Pettis measurability theorem*

Theorem 1.3. *Let $(A, \mathscr{A})$ be a measurable space, E a separable Banach space. Then a function $f : A \to E$ is strongly $\mathscr{A}$-measurable if and only if for all $x^* \in E^*$ the function $\langle f, x^* \rangle : A \to \mathbb{K}$ is $\mathscr{A}$-measurable.*

For the proof of this theorem we will need the following lemma

Lemma 1.4. *Let E be a separable Banach space. Then there exists a sequence of unit vectors $(x_n^*)_{n=1}^{\infty}$ in E^* with the property that for each $x \in E$:*

$$\|x\| = \sup_n |\langle x_n^*, x \rangle| \tag{1.2}$$

A sequence satisfying (1.2) *for every $x \in E$ is called* norming.

Proof. Let $(x_n)_{n=1}^{\infty}$ be a dense sequence in E and let $(\epsilon_n)_{n=1}^{\infty}$ be a sequence of strictly positive real numbers smaller than one such that $\lim_{n\to\infty} \epsilon_n = 0$. Then for each n, by the norm definition in 1.1, one can choose an $x_n^* \in E^*$ of unit length, such that $|\langle x_n, x_n^* \rangle| \geq (1-\epsilon_n)\|x_n\|$. Now take an arbitrary $\delta > 0$, and by the way we chose our sequences we can find n_0 such that $\|x - x_{n_0}\| < \delta$ and $\epsilon_{n_0} < \delta$. Then one has, using the triangle inequality and the fact that the x_n^* were defined to be of unit length:

$$\begin{aligned}(1-\delta)\|x\| \leq (1-\epsilon_{n_0})\|x\| &\leq (1-\epsilon_{n_0})\left(\|x_{n_0}\| + \delta\right) \\ &\leq \left|\left\langle x_{n_0}^*, x_{n_0} \right\rangle\right| + \delta \leq \left|\left\langle x_{n_0}^*, x \right\rangle\right| + 2\delta\end{aligned}$$

Since δ was chosen arbitrarily, one has $\|x\| \leq \left|\left\langle x_{n_0}^*, x \right\rangle\right|$, from which equality follows trivially with equation 1.1 QED

Proof of theorem 1.3. (if) Let $(x_n^*)_{n=1}^{\infty}$ be a norming sequence of unit vectors in E^* and let f be such that $\langle f, x^* \rangle : A \to (K)$ is measurable. Then by elementary measure theory the function

$$\xi \mapsto \|f(\xi) - x\| = \sup_n |f(\xi) - x, x_n^*|$$

is measurable for any $x \in E$. Now let $(x_n)_{n=1}^{\infty}$ be a dense sequence in E. Then define the function $s_n : E \to E$, taking values only in $x_1 \ldots x_n$, as follows: let $k(n, y)$ be the index of the x_i closest to $y \in E$ for $1 \leq i \leq n$. If there are several x_i with the same distance, one takes the smallest index. In other words: for $y \in E$ the number $k(n, y)$ is the smallest integer k with $1 \leq k \leq n$ satisfying:

$$\|y - x_k\| = \min_{1 \leq j \leq n} \|y - x_j\|$$

Now we define s_n to be the corresponding element: $s_n(y) := x_{k(n,y)}$. By density of the sequence $(x_n)_{n=1}^{\infty}$one has that

$$\lim_{n\to\infty} \|s_n(y) - y\| = 0$$

for $y \in E$. Now we define for each n the function $f_n : A \to E$ by $f_n(\cdot) = s_n(f(\cdot))$. We see that the f_n all take on only finitely many values and that the inverse images (for $1 \le k \le n$)

$$\begin{aligned} \{\xi \in A : \ f_n(\xi) = x_k\} \quad &= \quad \left\{\xi \in A : \ \|f(\xi) - x_k\| = \min_{1 \le j \le n} \|f(\xi) - x_j\|\right\} \\ &\quad \cap \left\{\xi \in A : \ \|f(\xi) - x_l\| > \min_{1 \le j \le n} \text{ for } l = 1, \dots, k-1\right\} \end{aligned}$$

are in $\mathscr{A}$because of the $\mathscr{A}$-measurability of $\|f(\cdot) - x\|$. So we can conclude that all of the f_n are $\mathscr{A}$-simple functions. And pointwise we have for each $\xi \in A$ that

$$\lim_{n \to \infty} \|f_n(\xi) - f(\xi)\| = \lim_{n \to \infty} \|s_n(f(\xi)) - f(\xi)\| = 0$$

This explicitely proofs that f is strongly measurable.

(only if) Let $(f_n)_{n=1}^{\infty}$be a sequence of $\mathscr{A}$-simple functions converging to f. Then because of continuity for each n the function $\langle f_n, x^* \rangle : A \to E$ is $\mathscr{A}$-measurable and then so is the pointwise limit $\langle f, x^* \rangle : A \to E$ QED

1.1.2 Measurability with respect to a measure

Now we introduce a measure space $(A, \mathscr{A}, \mu)$, which we will assume to be σ-finite and still work with a separable Banach space E. First we define a *μ-simple function* $f : A \to E$ to be a finite sum

$$f = \sum_{n=1}^{N} 1_{A_n} x_n \tag{1.3}$$

where all the $x_n \in E$ and all the $A_n \in \mathscr{A}$ with $\mu(A_n) < \infty$ We say furthermore that two functions are μ-versions of each other if they agree μ-almost everywhere on A. Anologously to the previous subsection, we arrive at the following definition:

Definition 1.5. A function $f : A \to E$ is called *μ-strongly measurable* if it is the μ-almost everywhere limit of a sequence of μ-simple functions.

The bridge between the two forms of strong measurability is made by the following proposition, which I state here without proof.

Proposition 1.6. *A function $f : A \to E$ is strongly μ-measurable if and only if f has a μ-version which is strongly $\mathscr{A}$-measurable.*

The proof is rather straightforward and can be found in [11]. Now a combination of proposition 1.6 and theorem 1.3 now implies the following version of the Pettis measurability theorem.

Theorem 1.7. *A function $f : A \to E$ is strongly μ-measurable if and only if for each $x^* \in E^*$, the function $\langle f, x^* \rangle : E \to \mathbb{K}$ is μ-measurable.*

1.1.3 The Bochner Integral

We are now ready to translate the notion of the Lebesgue integral to the Banach space-valued setting. Again we have the σ-finite measure space $(A, \mathscr{A}, \mu)$ and the separable Banach space E. I will be omitting the prefix μ if no confusion arises.

Definition 1.8. A function $f : A \to E$ is called *μ-Bochner integrable* if there exists a sequence $(f_n)_{n=1}^{\infty}$of μ-simple functions from A tot E such that:

1. $\lim_{n\to\infty} f_n = f$ almost everywhere
2. $\lim_{n\to\infty} \int_A \|f_n - f\| d\mu = 0$

Of course, f is strongly μ-measurable if definition 1.8 holds. Trivially every μ-simple function is μ-Bochner integrable. If we have a μ-simple function $f = \sum_{n=1}^{N} 1_{A_n} x_n$ we will take the natural definition:

$$\int_A f d\mu := \sum_{n=1}^{N} \mu(A_n) x_n \tag{1.4}$$

Now if f is a Bochner integrable function and $(f_n)_{n=1}^{\infty}$is a sequence as in definition 1.8 we have the very intuitive analogon to the Lebesgue integral:

$$\int_A f d\mu := \lim_{n\to\infty} f_n d\mu \tag{1.5}$$

As in the real-valued case of the Lebesgue-integral, it is routine to check that equations 1.4 and 1.5 lead to proper and unambiguous definitions. Also one sees that the Bochner integrals of functions that are versions of each other agree. By a limiting argument, one also sees that for $x^* \in E^*$ the identity

$$\left\langle \int_A f d\mu, x^* \right\rangle = \int_A \langle f, x^* \rangle \, d\mu$$

A practical necessary and sufficient condition for Bochner integrability is given by the following proposition:

Proposition 1.9. *A strongly μ-measurable function $f : A \to E$ is μ-Bochner integrable if and only if*

$$\int_A \|f\| d\mu < \infty,$$

in which case we have the inequality

$$\left\| \int_A f d\mu \right\| \leq \int_A \|f\| d\mu$$

I will omit the proof.

1.2 Random Variables

1.2.1 Random variables and expectations

From now on, the measure space $(A, \mathscr{A}, \mu)$ will be replaced by the probability space $(\Omega, \mathscr{F}, \mathbb{P})$, which is a measure space with the added property that $\mathbb{P}(\Omega) = 1$.

Definition 1.10. An E-valued *random variable* over a propability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a $\mathbb{P}$-strongly measurable function $X : (\Omega, \mathscr{F}, \mathbb{P}) \to E$.

Unless stated otherwise, the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ will be considered fixed and, in most cases, not explicitely mentioned. Also, the prefix "$\mathbb{P}$-" will be omitted from notation. Using our notion of integrals as in the previous section we can define:

Definition 1.11. Let X be a Bochner integrable random variable. The *mean value* or *expectation* of a random variable is its (Bochner) integral over all of Ω and is denoted by:

$$\mathbb{E}X := \int_\Omega X d\mathbb{P}$$

By proposition 1.6, we have that every random variable X has a strongly $\mathscr{F}$-measurable version $\tilde{X}$, and the probabilities $\mathbb{P}\{\tilde{X} \in B\}$, for $B \in \mathscr{B}(E)$ are independent of the version of $\tilde{X}$ chosen. This justifies the assumption that X is strongly $\mathscr{F}$-measurable, choosing $\tilde{X}$ for X when necessary.

Definition 1.12. The *distribution* of a random variable X is the Borel measure μ_X defined on $(E, \mathscr{B}(E))$ by

$$\mu_X(B) := \mathbb{P}\{X \in B\}$$

Two random variables, not necessararily defined on the same probability space, are said to be *identically distributed* if they have the same distribution.

The notion of distribution allows us to formulate the definition

Definition 1.13. Let $(X(i))_{i \in I}$ be a family of random variables each taking values in a Banach space E_i. We call this family independent if for each finite set of indices $i_1, \ldots, i_N \in I$ and all corresponding Borel sets $B_1, \ldots, B_N$ in $E_{i_1}, \ldots, E_{i_N}$ one has:

$$\mathbb{P}\left\{X_{i_1} \in B_1, \ldots, X_{i_N} \in B_N\right\} = \prod_{n=1}^{N} \mathbb{P}\left\{X_{i_n} \in B_n\right\} \tag{1.6}$$

With the random variables and Banach spaces as in definition 1.13, one can define the *joint distribution* on the product of the measurable spaces $(E_n, \mathscr{B}(E_n))$: $\mu_{(X_1,\ldots,X_n)}$. Completely analogous to the finite dimensional case we have the alternative definition of independence in terms of joint distributions:

$$\mu_{(X_1,\ldots,X_n)} = \mu_{X_1} \times \ldots \times \mu_{X_n} \tag{1.7}$$

The following definition will also be important in the coming sections

Definition 1.14. The *Fourier transform* of a Borel probability measure μ on E is the function

$$\begin{aligned} \hat{\mu} : E^* &\to \mathbb{C} \\ x^* &\mapsto \int_E \exp\left(-i < x, x^* >\right) d\mu(x) \end{aligned}$$

By the Fourier transform of a random variable X we mean the Fourier transform of its distribution μ_X

A result I will not prove here gives that if two random variables have the same Fourier transform, they are identically distributed.

1.2.2 Gaussian random variables

We begin with the definition of Gaussianity in the real-valued case. In this case we call γ a *Gaussian random variable* if its Fourier transform is:

$$\mathbb{E}(e^{-i\xi\gamma}) = e^{-\frac{1}{2}q\xi^2}, \quad \xi \in \mathbb{R} \tag{1.8}$$

For some $q \geq 0$. If $q = 0$ one has almost surely $\gamma = 0$ and otherwise, γ has the density function:

$$f_\gamma(t) = \frac{1}{\sqrt{2\pi q}} e^{-\frac{t^2}{2q}} \tag{1.9}$$

Quite naturally, one would like to extend this definition to the Banach space-valued setting, from now on, we will mostly assume that E is a real Banach space.

Definition 1.15. An E-valued random variable is Gaussian if for every $x^* \in E^*$ the random variable $\langle X, x^* \rangle$ is Gaussian.

A celebrated theorem by Fernique asserts that every Gaussian variable is exponentially bounded and therefore has moments of all order.

Theorem 1.16 (Fernique)**.** *Let X be an E-valued Gaussian random variable. Then there exists a constant $\beta > 0$ such that*

$$\mathbb{E}e^{\beta\|X\|^2} < \infty \tag{1.10}$$

As one can imagine, bounds like these play an essential role in integration with respect to processes with gaussian variations. This thesis will consider integrals with respect to processes with less well-behaved variations, but let's stick to the subject at hand.

1.2.3 Conditionality

We will again be working with a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and we define $\mathscr{G}$ to be a sub-σ-algebra of $\mathscr{F}$. For notational convenience, we define for $1 \leq p \leq \infty$ the spaces $L^p(\Omega, \mathscr{G})$ as being the spaces of real-valued random variables in $L^p(\Omega)$ having a $\mathbb{P}$-version which is strongly $\mathscr{G}$-measurable. A simple limiting argument may show that $L^p(\Omega, \mathscr{G})$ is a closed linear subspace of $L^p(\Omega)$. If we take p to be 2, we know that $L^2(\Omega)$ is a Hilbert space and one has the orthogonal decomposition:

$$L^2(\Omega) = L^2(\Omega, \mathscr{G}) \oplus L^2(\Omega, \mathscr{G})^\perp$$

In this case we define the conditional expectation of some random variable $X \in L^2(\Omega)$ given $\mathscr{G}$ to be the random variable in $L^2(\Omega, \mathscr{G})$ obtained by orthogonally projecting X onto $L^2(\Omega, \mathscr{G})$. Letting $P_K(\cdot)$ denote orthogonal projection onto a subspace K, we can write for X:

$$\mathbb{E}(X|\mathscr{G}) = P_{L^2(\Omega,\mathscr{G})}(X) \tag{1.11}$$

The objective is of course to extend this definition to the setting of variables in $L^p(\Omega)$ for $p \neq 2$, and ultimately to variables in $L^p(\Omega, E)$ for a Banach space E. But first, we have some properties of conditional expectations in the L^2-setting.

Lemma 1.17. *Let X be a random variable in $L^2(\Omega)$ and $G \in \mathscr{G}$, then*

$$\int_G \mathbb{E}(X|\mathscr{G})d\mathbb{P} = \int_G X d\mathbb{P}$$

Proof. One has that $X - E(X|\mathscr{G}) \perp L^2(\Omega, \mathscr{G})$, now because $G \in \mathscr{G}$, $1_G \in L^2(\Omega, \mathscr{G})$, so that

$$\int_G [X - \mathbb{E}(X|\mathscr{G})]\, d\mathbb{P} = \int_\Omega 1_G [X - \mathbb{E}(X|\mathscr{G})]\, d\mathbb{P} = 0$$

QED

A simple corollary to this lemma is that if X is almost surely positive (or nonnegative, negative, etc.), then so is $\mathbb{E}(X|\mathscr{G})$.

By density of L^2 in L^1, and by inclusion of all L^p-spaces in L^2 for $p > 2$ (since $(\Omega, \mathscr{F}, \mathbb{P})$ is finite, cf. [14], section 5.5) one can uniquely extend or restrict the definition of conditional expectation to all L^p-spaces for $1 \leq p < \infty$. For any L^p-space we will call this extension or restriction the *conditional expectation operator*.

The purpose is now to extend this analysis to the spaces of random variables in E, $L^p(\Omega, E)$, i.e. the spaces of E-valued random variables X for which the integral

$$\int_\Omega \|X\|^p d\mathbb{P}$$

exists and is finite. Endowed with the norm

$$\|X\|_p = \left(\int_\Omega \|X|^p d\mathbb{P}\right)^{\frac{1}{p}}$$

these spaces, called the *Lebesgue-Bochner spaces*, form Banach spaces.

First we look at functions of the form $f \otimes x$, for $f \in L^p(\Omega)$ and $x \in E$. Clearly, these functions are in $L^p(\Omega, E)$. We denote the linear span of all these functions by $L^p(\Omega) \otimes E$. I state without proof that $L^p(\Omega) \otimes E$ is dense in $L^p(\Omega, E)$. The idea is of course to linearly extend any bounded operator T (in this case, the conditional expectation operator) on $L^p(\Omega)$ to an operator $T \otimes I$ on $L^p(\Omega) \otimes E$ by setting:

$$T \otimes I : f \otimes x \mapsto Tf \otimes x$$

Unfortunately this extension doesn't work for any bounded linear operator, but it does for positive operators:

Proposition 1.18. *Let T be a positive operator on $L^p(\Omega)$, then $T \otimes I$ extends uniquely to a bounded operator on $L^p(\Omega, E)$*

Proof. The proof will be omitted QED

Now by proposition 1.17, this extension can be made. The analogon for proposition 1.17 can also be given by the following proposition:

Proposition 1.19. *Let $X \in L^p(\Omega, E)$ be a random variable and $\mathscr{G} \subset \mathscr{F}$. Then the random variable*

$$\mathbb{E}(X|\mathscr{G}) := (\mathbb{E}(\cdot|\mathscr{G}) \otimes I)\, X$$

is the unique random variable in $L^p(\Omega, E, \mathscr{G})$ such that

$$\int_G (X|\mathscr{G}) d\mathbb{P} = \int_G X d\mathbb{P}$$

for every $G \in \mathscr{G}$

Proof. Again, the proof will be omitted. QED

1.3 Random Processes

Next we turn our attention to stochastic processes in Banach space.

Definition 1.20. Let I be some set of indices, then an E-valued *stochastic process*, *random process* or, for short, *process* indexed by I is a family $(X(i))_{i\in I}$ of stochastic variables in E, all definied on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Just like we did for random variables, we will generally not explicitly mention the underlying probability space for a process. In most cases I will be either the nonnegative real numbers or the interval $[0, T]$ for some positive T, justifying the notion of a *continuous path* for a certain $\omega \in \Omega$, being the intuitive notion that the function $i \mapsto X(i)(\omega)$ is a continuous function. These choices also justify the following much used concept:

Definition 1.21. Let U be a (closed connected subset of) the real numbers. An E-valued process indexed by U, $(X(t))_{t\in U}$ is called *càdlàg* (short for the french words *continue à droite, limites à gauche*) for a certain $\omega \in \Omega$ if for every $t \in U$:

1. $\lim_{s\nearrow t} X(s)(\omega)$ exists, and
2. $\lim_{s\searrow t} X(s)(\omega) = X(t)(\omega)$

In honour of some of the main early contributors to this area of mathematics, we will use the common French acronym.

The measure theoretic foundation of a stochastic processes $(X(i))_{i\in I}$ over a partially ordered index set I starts with the notion of a *filtration*, indexed by the same set:

Definition 1.22. A filtration on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a family of σ-algebra's on Ω, $(\mathscr{F}_t)_t$, such that whenever $s \leq t$, $\mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{F}$. A probability space endowed with a filtration $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \mathbb{P})$ will be called a *filtered space*.

Definition 1.23. A process $X(i)_{i\in I}$ defined on a filtered space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \mathbb{P})$ is called *adapted* if for every $i \in I$, one has $X(i) \in \mathscr{F}_i$

1.3.1 Gaussian processes and Brownian motion

Definition 1.24. A *Gaussian process* is an E-valued process if for any finite set of indices $\{i_1, \dots, i_N\}$ the E^N-valued random variable $(X(i_1), \dots, X(i_N))$ is a Gaussian variable.

First, we can define Brownian motion in the real valued case.

Definition 1.25. A *Brownian motion*, or *Wiener Process*, is a real valued process $(W(t))_{t\in[0,T]}$ satisfying:

1. $W(0) = 0$ almost surely,
2. $W(t) - W(s)$ is a Gaussian variable with variance $t - s$ whenever $0 \leq s \leq t \leq T$ and
3. Whenever $0 \leq s \leq t \leq T$, the variable $W(t) - W(s)$ is independent of $\{W(r) : 0 \leq r \leq s\}$

A process satisfying the third property in definition 1.25 is said to have *independent increments*. There are various proofs for the existence of Brownian motion. I will follow the quite elegant line found among others in [11]. But first, a quick look at the definition and the fact that the distribution of a Gaussian variable is determined by its variance give us that Brownian motion, if it exists, is unique in the sense that if there are two Brownian motions $(W(t))_{t\in[0,T]}$ and $(\tilde{W}(t))_{t\in[0,T]}$, then for any $t \in [0,T]$, the random variables $W(t)$ and $\tilde{W}(t)$ are identically distributed.

A well known result in probability theory ([11], [21] for example) asserts that every Brownian motion (if it exists) is a Gaussian process, moreover there exists a sort of "converse" to this statement.

Proposition 1.26. *A real valued process $(W(t))_{t\in[0,T]}$ is a Brownian motion if and only if it is a Gaussian process with*

$$\mathbb{E}(W(s)W(t)) = \min\{s,t\} \tag{1.12}$$

whenever $0 \le s,t \le T$

Proof. I will proof the 'if' part, which is the least trivial, as well as the most important for our proof for existence of Brownian motions.

First we have that $\mathbb{E}(W(0)^2) = 0$, giving the first property of definition 1.25.

Now let $0 \le s \le t \le T$, then $\mathbb{E}(W(t) - W(s))^2 = \mathbb{E}[W(t)^2 - 2W(s)W(t) + W(s)^2] = t - 2min\{s,t\} + s = t - s$, and by the fact that $(W(t))_{t\in[0,T]}$ is a Gaussian process, the second property holds too.

QED

Now we know that we only have to prove the existence of a Gaussian process satisfying 1.12 in order to prove the existence of Brownian motion. To this end, we first define a Hilbert space H with inner product $[\cdot,\cdot]$ and let $(\Omega, \mathscr{F}, \mathbb{P})$ be a measure space.

Definition 1.27. An H-isonormal process on Ω is a mapping $W : H \to L^2(\Omega)$ satisfying:

1. For any $h \in H$, the variable Wh is Gaussian and
2. Whenever $h_1, h_2 \in H$, one has $\mathbb{E}\left(Wh_1 \cdot Wh_2\right) = [h_1, h_2]$.

It is routine to check that H-isonormal processes are linear mappings. This linearity implies that $\sum_{n=1}^N c_n Wh_n = W(\sum_{n=1}^N c_n h_n)$ (with all the $h_n \in H$), such that the $\mathbb{R}^N$-valued vector $(Wh_1, \ldots, Wh_N)$ is Gaussian, making $(Wh)_{h\in H}$ a Gaussian process.

An explicit isonormal process is given for example by taking a sequence of independent standard Gaussian variables $(\gamma_n)_{n=1}^\infty$ and a separable Hilbert space H with an orthonormal basis $(h_n)_{n=1}^\infty$ and defining:

$$\begin{array}{rcl} W : H & \to & L^2(\Omega) \\ h & \mapsto & \sum_{n=1}^\infty \gamma_n [h, h_n] \end{array} \tag{1.13}$$

Having established the existence of H-isonormal processes, the next theorem asserts the existence of Brownian motions.

Theorem 1.28. *Let $W : H \to L^2(0,T)$ be an isonormal process. The process $(W(t))_{t\in[0,T]}$, defined by $W(t) := W1_{[0,t]}$ is a Brownian motion.*

Proof. We already observed that $(W(t))_{t\in[0,T]}$ is a Gaussian process, so the observation that

$$\begin{aligned}\mathbb{E}(W(s)W(t)) &= [1_{[0,s]}, 1_{[0,t]}] \\ &= \int_0^T 1_{\min\{s,t\}} \\ &= \min\{s,t\}\end{aligned}$$

concludes our proof. QED

Although this proof for existence of Brownian motion may seem very nonconstructive, applying theorem 1.28 to the process defined in equation 1.13, one explicitely gets

$$W(t) = \sum_{n=0}^{\infty} \gamma_n \int_0^t h_n(s)ds \tag{1.14}$$

As a Brownian motion, where $(h_n)_n$ is an orthonormal basis in $L^2(0,T)$, where one can take for example a trigonometric basis, a wavelet basis or a polynomial basis. For simulating Brownian motion one could choose some orthonormal basis and explicitely substitute this into equation (1.14).

Cylindrical Brownian Motion

Having established the existence of Brownian motion in the real-valued case, I will now try and extend this definition to a more general setting. When working in $\mathbb{R}^N$ for some N, one usually defines Brownian motion as the product of N independent Brownian motions on $\mathbb{R}$. For a Hilbert space H, the concept of a *cylindrical Brownian motion* is often used.

Definition 1.29. Let H be a Hilbert space and let $(W_H(t))_{t\in[0,T]}$ be a family of mappings from H to $L^2(\Omega)$, then $(W_H(t))_{t\in[0,T]}$ is called a H-cylindrical Brownian motion if

1. $(W_H(t)h)_{t\in[0,T]}$ is a Brownian motion for all $h \in H$
2. Let $0 \leq s \leq t \leq T$ and $h, j \in H$, then

$$\mathbb{E}(W_H(s)h \cdot W_H(t)j)) = s < h, j >$$

There are several explicit constructions of cylindrical Brownian motions, but perhaps the most intuitive one is given by taking an orthonormal basis for H: $(h_n)_{n=1}^{\infty}$ and a sequence of independent Brownian motions $(W^{(n)})_n$ and defining

$$W_H(t) := \sum_{n=1}^{\infty} W^{(n)}(t) < \cdot, h_n >$$

1.3.2 Martingales

As the character two-face in the 2008 Christopher Nolan movie "The Dark Knight" says at the moment he decides by a coin flip whether or not to kill his next victim:

"You thought we could be decent men in an indecent world. But you were wrong; the world is cruel, and the only morality in a cruel world is chance. Unbiased... Unprejudiced... Fair...,"

we can see that the notion of fair games is ubiquitous. Although I do not share his pessimism, nor condone his ways of playing with chance, this gives us an illustration of the concept of a fair game. The mathematical framework for this concept is given by the notion of a *martingale.*

For the rest of this section $(M_t)_t$ will be an adapted process on the filtered space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \mathbb{P})$.

Definition 1.30. $(M_t)_t$ is called a martingale if for $j \geq i$

$$\mathbb{E}(M_j|\mathscr{F}_i) = M_i$$

As in the real case, a quite canonical example is given by:

Example 1.31*:* Let I be $0, 1, 2, \ldots$ and let $(X_i)_{i\in I}$ be a sequence of independent random variables, such that $\mathbb{E}X_i = 0$ for all i and let

$$M_n = \sum_{i=1}^{n} X_i,$$

then M_n is a martingale with respect to the natural filtration generated by $(X_i)_i$.

A very important concept in the theory of Martingales is that of a *stopping time*:

Definition 1.32. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \mathbb{P})$ be a filtered space. The positive random variable τ is called a stopping time if the event $\{\tau \leq t\} \in \mathscr{F}_t$ for all t and $\mathbb{P}(\tau \leq \infty) = 1$.

For a process $(M(t))$ and a stopping time τ we define the *stopped process* $M^\tau(t) := M(\min(t, \tau))$. I state without proof that stopped martingales are martingales.

Definition 1.33. Let $M := (M(t))_t$ be a càdlàg process and let $(\tau_n)_n$ be a sequence of stopping times increasing to infinity. M is called a *local martingale* if $1_{\{\tau_n>0\}}M^{\tau_n}$ is a martingale for each n. A process $L := (L(t))_t$ is called a *semimartingale* if it can be written as $L(t) = M(t) + A(t)$, where $M(t)$ is a local martingale and $A(t)$ is of bounded variation, i.e. for each $T > 0$ the supremum of the sums $\sum_{j=1}^{m} |A(t_j) - A(t_{j-1})|$ over all partitions $0 =: t_0 < t_1 < t_2 < \ldots < t_m := T$ is finite, almost surely.

Now we assume that the differential $dM(t)$ makes sense in one way or another and restate the martingale property for a continuous process $(M(t))_t$ with independent increments as follows

$$\mathbb{E}dM(t) = 0$$

The fact that this is equivalent to the martingale property is given by stating first that for $b > a$

$$\begin{aligned}\mathbb{E}\left(M(b)|\mathscr{F}_a\right) &= \mathbb{E}\left(M(a) + \int_a^b dM(t)|\mathscr{F}_a\right) \\ &= M(a) + \int_a^b \mathbb{E}dM(t) = M(a)\end{aligned}$$

And, of course that $\mathbb{E}\left(\int_a^b dM(t)\right) = \mathbb{E}(M(b) - M(a)) = 0$, so by Fubini and the fact that a and b were arbitrary, we see that $\mathbb{E}dM(t) = 0$

Now the notion of arbitrage is a very important one in financial mathematics and this notion is very closely related to the theory of martingales. A process is said to *admit arbitrage* if staking an amount on its behaviour garantees a sure profit. In an ideal economy in equilibrium, arbitrage is supposed to be impossible, so the pricing of derivatives and the study of the behaviour of financial instruments is based heavily on no-arbitrage arguments. This theorem for processes based on Brownian motion was stated and proved in 1981 by Harrison and Pliska in [5].

A process M on a filtered space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \mathbb{P})$ is said to allow a *martingale measure* if there exists a measure $\tilde{\mathbb{P}}$, absolutely continuous with respect to $\mathbb{P}$ such that M is a martingale on the filtered space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_t, \tilde{\mathbb{P}})$. A theorem in economic theory is that M admits no arbitrage if and only if M allows a martingale measure. We will not go into the depth of this theorem, but we will give a very simple example[1].

Example 1.34*:* We will consider a one-time balanced coin flip. Our index set will be $\{0, 1\}$, where 0 corresponds to the situation before the coin flip and 1 to the situation after the coin flip. We consider three payout schemes, all starting with having no money at time zero:

1. The player gets a payout of one unit stake when heads is thrown and has to pay one unit stake when tails comes up.
2. The player gets a payout of two units for heads and has to pay one unit for tails.
3. The player gets one unit stake, whichever side comes up.

We can model the probability space as $\Omega := \{H, T\}$, with its power set as the σ-algebra and $\mathbb{P}(\{H\}) = \mathbb{P}(\{T\}) = \frac{1}{2}$. One can model each of these games as a stochastic process $(M_t^{(1)})_{t=0,1}, (M_t^{(2)})_{t=0,1}, (M_t^{(3)})_{t=0,1}$ where $M_t^{(i)}$ represents the number of units stake won at time t in game i. One sees that $M^{(1)}$ is already a martingale. The second game, although it is preferential for the player, admits no strategy that ensures a riskless profit. And indeed "remeasuring" the probabilty space by $\tilde{\mathbb{P}}(\{T\}) = 2\tilde{\mathbb{P}}(\{H\}) = \frac{2}{3}$ makes $M^{(2)}$ into a martingale. For the third game, it is easy to see that there exists no measure that ensures $\mathbb{E}M_1^{(3)} = M_0^{(3)} = 0$ and indeed, staking any positive amount will result in a riskless positive payout.

[1]Although I devised this example myself, I can imagine by its rather obvious nature that it features in some way in any textbook on this matter. If this is the case, I apologise beforehand for the unintended infringement.

1.3.3 Lévy processes

Lévy processes form a large and important class of stochastic processes that can be defined on (a subinterval of) the real numbers. They appear throughout applied mathematics to model all kinds of stochastic phenomena.

Definition 1.35. Let $L := (L(t))_t$ be a stochastic process. $L(t)$ is said to have *stationary increments* if for arbitrary $s \leq t \leq \sigma \leq \tau$ the random variables $L(\tau)-L(\sigma)$ and $L(t)-L(s)$ are identically distributed.

Definition 1.36. Let $L := (L(t))_t$ be a stochastic process. $L(t)$ is said to have *independent increments* if for arbitrary $s \leq t \leq \sigma \leq \tau$ the random variables $L(\tau)-L(\sigma)$ and $L(t)-L(s)$ are independent.

Now we can define Lévy processes as follows:

Definition 1.37. A stochastic process L is called a Lévy process if

1. L admits a càdlàg version
2. L has stationary increments
3. L has independent increments

Probably the most evident example of a Lévy process is Brownian Motion.

Chapter 2

Stochastic integration (I)

2.1 Real Brownian Motion

Integration with respect to Brownian motion was developed in two different ways by Itô [3] and Stratonovich. However, in applied mathematics the Itô-definition is the most common. Excellent surveys of the Itô integral and its application in stochastic differential equations, filtering and financial mathematics can be found in [16] and [21]. I will only be concerned with the Itô definition and try and give some of the main results.

First we define the space $H^2 := H^2[0,T]$ as the class of adapted random functions from $[0,T]$ to $\mathbb{R}$ such that

$$\mathbb{E}\left[\int_0^T f^2(\omega,t)dt\right] < \infty. \tag{2.1}$$

Now first of all this integral is defined for all functions in H^2, that can be written in the form

$$f(\omega,\cdot) = \sum_{i=0}^{n-1} a_i(\omega)1_{(t_i,t_{i+1}]} \tag{2.2}$$

For which, of course, the integral is defined as:

$$\int_0^T f(\omega,T)dW_t := \sum_{i=0}^{n-1} a_i(\omega)\left(W_{t_{i+1}} - W_{t_i}\right) \tag{2.3}$$

Now by limiting procedures this leads ultimately to a definition of the so-called Itô integral, satisfying (for f in H^2):

$$\mathbb{E}\left[\left(\int_0^T f(\omega,t)dW_t\right)^2\right] = \mathbb{E}\left[\int_0^T f^2(\omega,t)dt\right] \tag{2.4}$$

Equation (2.4) is called the *Itô isometry.*

Using a limiting procedure to calculate the Itô integral for each separate case is very tedious and complicated. Luckily, we have Itô's lemma. For this we first define a so-called Itô process to be a process of the form:

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dW_s \tag{2.5}$$

In most cases we will write Itô processes in differential form:

$$dX_t = udt + vdW_t \tag{2.6}$$

Now let g be a sufficiently smooth function from $[0,\infty)\times\mathbb{R}$ to $\mathbb{R}$, then Itô's lemma asserts that:

$$d\left(g(t,X_t)\right) = \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)(dX_t)^2, \tag{2.7}$$

where $(dt)^2 = dtdW_t = dB_t dt = 0$ and $(dW_t)^2 = dt$. Itô's lemma has proven to be a very powerful tool for solving stochastic differential equations or solving other types of problems concerning Brownian motions. For example, the value of many financial instruments is given by non-affine functions of Brownian motion, so integrals with respect to their values can be evaluated.

2.2 Cylindrical Brownian Motion

Even more than in the previous section, the technical details of the theory in this section go well beyond the scope of this thesis, yet, to understand the framework and to appreciate the simplicity of the analysis in the forthcoming sections, it will be useful to mention some of the results. All of the results were obtained quite recently mostly by Van Neerven and Weis and published in [7] and [10]. A full development of the theory from a relatively basic level can be found in [11].

First, we let $W_H := (W_H(t))_{t\in[0,T]}$ be a cylindrical Brownian motion and we regard a function $\Phi : (0,T) \to \mathscr{L}(H,E)$, where E is a separable Banach space.

Definition 2.1. The function Φ is called *stochastically integrable with respect to* W_H if there exists a sequence of finite rank step functions Φ_n such that

1. for all $h \in H$ the limit in measure $\lim_{n\to\infty}\Phi_n h = \Phi h$
2. there exists an E-valued random variable X such that $\lim_{n\to\infty}\int_0^T \Phi_n dW_H = X$ in probability.

Of course the stochastic integral of Φ is then defined to be the limit in probability

$$\int_0^T \Phi dW_H = \lim_{n\to\infty}\int_0^T \Phi_n dW_H,$$

where the integral of the step function $1_{(s,t)}\otimes(h\otimes x)$ is of course defined as

$$\int_0^T \left[1_{(s,t)}\otimes(h\otimes x)\right]dW_H = (W_H(t)h - W_H(s)h)\otimes x.$$

Several quite technical conditions for integrability exist that will not be mentioned here, yet these conditions do not require any specific properties of the Banach space other than perhaps separability. Of course one can look at the integral over a specific subinterval of $[0,T]$ by multipliying Φ with an indicator function.

Next we allow Φ to be a process, i.e. a function $\Phi : [0,T]\times\Omega \to \mathscr{L}(H,E)$, which we assume to be adapted. This strongly complicates the analysis of stochastic integrals. To this end we will first need a couple of really technical definitions. First of all, if one has a martingale M_n, we define $d_n := M_n - M_{n-1}$ to be its *difference sequence*. We assume for the rest of this section that $1 < p < \infty$.

Definition 2.2. Let M_n be an E-valued martingale with $M_n \in L^p(\Omega, E)$ and let d_n be its difference sequence. If there exists a constant β such that for any such d_n and any sequence of signs $\epsilon_n \in \{\pm 1\}$ one has:

$$\mathbb{E}\left\|\sum_{n=1}^{\infty} \epsilon_n d_n\right\|^p \leq \beta^p \mathbb{E}\left\|\sum_{n=1}^{\infty} d_n\right\|^p, \tag{2.8}$$

the space E is called a UMD_p-space.

The abbreviation UMD stands for *uniform martingale differences.* A typical example of a UMD_p-space is the L^p-space over some measure space. Next, a very nontrivial result on these spaces is that the UMD_p-property is independent of the parameter p, so we can just talk about *UMD-spaces.*

The most recent ground-breaking result in this context is the fact that if E is a UMD Banach space and $\Phi : (0, T) \times \Omega \to \mathscr{L}(H, E)$ is an adapted process satisfying certain measurability and integrability conditions, then the integral $\int_0^T \Phi dW_H$ exists in some way. This result was obtained in [10]

2.3 First Jump Processes

Let $(t_i)_i$ be a sequence of Poisson times with a parameter λ in $[0, T]$, i.e. the lengths $t_i - t_{i-1}$ are independent and exponentially distributed with the same parameter λ and let $(X_i)_i$ be a sequence of E-valued random variables. Now we define the *jump process* $\left(J_t^\lambda\right)_{t\in[0,T]}$ by

$$J_t^\lambda := \sum_{i\,:\,t_i \leq t} X_i \tag{2.9}$$

We can make the following observations

1. The process is by definition càdlàg
2. By the memoryless property of the Poisson times, we know that if the X_i are independent, J_t^λ is a Markov process
3. If the X_i are independent and have expectation zero, J_t^λ is a martingale.

When the parameter λ doesn't play an essential role, we either assume it to be 1 and drop the superscript, or, even worse, drop the superscript all together. So J_t is the process that waits an exponentially distributed time before making a jump and then makes a jump according to a certain distribution. We define the so-called *Heaviside unit step function* by:

$$H(t) := \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geq 0 \end{cases} \tag{2.10}$$

Although this function isn't differentiable it can be regarded as an antiderivative to the so-called *Dirac δ-distribution*, the generalised function $\delta(\cdot)$ with $\delta(t) = 0$ for $t \neq 0$ and by definition $\int_{\mathbb{R}} f(t)\delta(t)dt = f(0)$ for all suficiently smooth functions $f(\cdot)$[1]. So one can write symbolically $dH(t) = \delta(t)dt$.

[1]for more information on this distribution one can see for example [19]

Using the function defined in (2.10), we can rewrite (2.9) as:

$$J_t^\lambda := \sum_i X_i H(t - t_i) \tag{2.11}$$

Since Banach space elements merely feature as (multiplicative or additive) constants, one can write:

$$dJ_t = \sum_i X_i \delta(t - t_i) dt$$

Now let $f(\cdot, \omega)$ be a function from $[0,T] \times \Omega$ to the space of linear operators from E to another Banach space E', $\mathscr{L}(E, E')$, then one can write down the integral

$$\int_0^T f(t,\omega) dJ_t = \int_0^T \sum_i \delta(t - t_i) f(t,\omega) X_i dt = \sum_i f(t_i, \omega) X_i \tag{2.12}$$

Using the fact that there are only finitely many t_i in $[0,T]$, almost surely.

If $f(t,\omega)$ only makes finitely many jumps and is independent of J_t, the probability of the jumps of f and J coinciding is zero and we won't have to worry about what value of $f(t_i)$ to take. Yet, one can easily imagine having to evaluate an integral like

$$\int_0^T J_t dJ_t.$$

Following the conventions for the Stieltjes integral in this cas we will take a "càglàd" version of the integrand for a càdlàg integrator. So we refine the expression in (2.12) as follows:

$$\int_0^T f(t,\omega) dJ_t = \sum_i f(t_i-, \omega) X_i, \tag{2.13}$$

with the notation $f(t_i-, \omega) := lim_{t \nearrow t_i} f(t, \omega)$.

Chapter 3

The Stochastic Abstract Cauchy Problem

3.1 The SACP with additive noise

First, we can present the simplest form of a *stochastic abstract Cauchy problem*. This is the stochastic analogon of the classical Cauchy problem. We define the SACP with additive noise to be an equation for the E-valued process $(X(t))_{t\in[0,T]}$

$$\begin{cases} dX(t) = AX(t)dt + BdL(t) & t \in [0,T] \\ X(0) = x & \text{for some } x \in E \end{cases} \tag{3.1}$$

In this equation, we assume that A is an operator from $D(A) \subset E$ to itself that generates a C_0-semigroup, $L(t)$ is an E'-valued Lévy process (we will be making several assumptions on the form $L(t)$ later on) and B is a bounded linear operator from E' to E. We will assume that the expression $dL(t)$ is in some way well-defined. Note that, a priori, the equation (3.1) doesn't necessarily make sense as $X(t)$ doesn't need to be in the domain of A.

We will define a *mild solution* of (3.1) to be an adapted E-valued process $X(t)$ satisfying

$$X(t) = S(t)x + \int_0^t S(t-s)BdL(s), \tag{3.2}$$

where $(S(t))_t$ is the semigroup generated by A. Note that this process is always well-defined (by the properties of A) and its definition is fully unambiguous, as long as we have the right definition for the integral. First of all we would like to see the relation between the mild solution and the solution of (3.1). To this end, we first establish a uniqueness result:

Lemma 3.1. *The problem* (3.1) *has at most one solution in the sense that any two solutions are (almost surely) equal for all t*

Proof. Let $X_1 := X_1(\cdot)$ and $X_2 := X_2(\cdot)$ be two solutions of (3.1). Then one has for $X = X_1 - X_2$ the (deterministic) equation

$$\begin{cases} dX(t) = AX_1(t)dt + BdL(t) - (AX_2(t)dt + BdL(t)) = AX(t)dt & t \in [0,T] \\ X(0) = 0 \end{cases} \tag{3.3}$$

This is a linear evolution equation with only the trivial solution that $X = X_1 - X_2 \equiv 0$, or that, for all t one has $X_1(t) = X_2(t)$. QED

Theorem 3.2. *Let $X(t)$ be such that $X(t) \in D(A)$ for all $t \in [0,T]$ and satisfy* (3.1) *where the process $L(t)$ satisfies a proper stochastic Fubini criterion, then $X(t)$ satisfies* (3.2)

Proof. We proof this theorem by first stating that if the process defined by (3.2) satisfies $X(t) \in D(A)$, that it is a solution to (3.1). Then invoking lemma 3.1 we see that this is the only solution to (3.1), thus proving our theorem.

Let $X(t)$ be the process defined by (3.2) and such that $X(t) \in D(A)$ for all t. The initial condition holds trivially, for the differential equation we can write

$$\begin{aligned} AX(t) &= AS(t)x + A\int_0^t S(t-s)BdL(s) \\ &= AS(t)x + \int_0^t AS(t-s)BdL(s) \end{aligned}$$

Now we calculate $\int_0^\tau AX(t)dt$, for any given $\tau \in [0,T]$.

$$\begin{aligned} \int_0^\tau AX(t)dt &= \int_0^\tau AS(t)xdt + \int_0^\tau \int_0^t AS(t-s)BdL(s)dt \\ &= (S(\tau)x - S(0)x) + \int_0^\tau \int_0^t AS(t-s)BdL(s)dt \\ &= (S(\tau)x - x) + \int_0^\tau \int_s^\tau AS(t-s)dtBdL(s) \\ &= (S(\tau)x - x) + \int_0^\tau (S(\tau-s) - I)\, BdL(s) \\ &= (S(\tau)x - x) + \int_0^\tau (S(\tau-s))\, BdL(s) - \int_0^\tau BdL(s) \end{aligned}$$

On the other hand we can calculate $\int_0^\tau (dX(t) - BdL(t))$:

$$\begin{aligned} \int_0^\tau (dX(t) - BdL(t)) &= X(\tau) - X(0) - \int_0^\tau BdL(t) \\ &= X(\tau) - x - \int_0^\tau BdL(t) \\ &= S(\tau)x + \int_0^\tau (S(\tau-s))\, BdL(s) - x - \int_0^\tau BdL(s) \\ &= \int_0^\tau AX(t)dt \end{aligned}$$

So we have that $AX(t)dt = dX(t) - BdL(t)$, so $X(t)$ satisfies (3.1). QED

We will try and see how a solution of a specific SACP might work out in a simple example

Example 3.3 (A stochastic wave equation)*:* We take $L(t)$ to be a real valued form of the jump process defined in section 2.3, i.e. where all the X_i are just real valued variables. Furthermore, we start out with $u_0(\cdot), f(\cdot) \in L^p(\mathbb{R})$ for some p. Then we write down the equation:

$$\begin{cases} du(x,t) = \frac{\partial}{\partial x}u(x,t)dt + f(x)dL(t) & t \in [0,T] \\ u(x,0) = u_0(x) \end{cases} \tag{3.4}$$

We can now rewrite this equation into a more functional analytic form by writing for each t: $X(t) = u(\cdot, t)$, writing the operator A to be the differentiation operator on a suitable function space $E(\mathbb{R})$, in the case where we are working in $L^p(\mathbb{R})$, this will typically be the Sobolev space $W^{1,p}(\mathbb{R})$, defining the operator $B : \mathbb{R} \to E$ by $B\lambda = \lambda f(\cdot)$ and writing $X_0 = u_0(\cdot)$ as

$$\begin{cases} dX(t) = AX(t)dt + BdL(t) & t \in [0,T] \\ X(0) = X_0 \end{cases} \tag{3.5}$$

Looking, of course, awfully familiar. Now A generates the semigroup $(T(t))_{t\geq 0}$ given by $(T(t)f)(x) = f(x+t)$. So we can rewrite the mild solution given by (3.2) as:

$$\begin{aligned} u(x,t) &= u_0(x+t) + \int_0^t f(x+t-s)dL(s) \\ &= u_0(x+t) + \sum_{i:t_i\leq t} f(x+t-t_i)X_i \end{aligned}$$

So we obtain a superposition of the classical solution that merely propagates the initial condition, and a solution propagating more and more random shocks over time.

The next logical equation to consider is an analogon to the heat equation

Example 3.4 (A stochastic heat equation)*:* With everything as in the previous example, we can write

$$\begin{cases} du(x,t) = \frac{\partial^2}{\partial x^2}u(x,t)dt + f(x)dL(t) & t \in [0,T] \\ u(x,0) = u_0(x) \end{cases} \tag{3.6}$$

This can now be rewritten as

$$\begin{cases} dX(t) = \Delta X(t)dt + BdL(t) & t \in [0,T] \\ X(0) = X_0 \end{cases} \tag{3.7}$$

Now we know that the Laplace operator Δ generates the semigroup $(S(t))_t$ given by $(S(t)f)(x) = \int_{\mathbb{R}} h(x-y,t)f(y)dy$, where $h(\cdot,t)$ is the Gaussian density function with standard deviation $h(x,t)$.

$$\begin{aligned} u(x,t) &= \int_{\mathbb{R}} h(x-y,t)u_0(y)dy + \int_0^t \int_{\mathbb{R}} f(y)h(x-y,t-s)dydL(s) \\ &= \int_{\mathbb{R}} dy \left[h(x-y,t)u_0(y) + f(y)\int_0^t h(x-y,t-s)dL(s) \right] \\ &= \int_{\mathbb{R}} dy \left[h(x-y,t)u_0(y) + f(y)\sum_{i:t_i\leq t} h(x-y,t-t_i)X_i \right] \end{aligned}$$

As one can see, this looks like a superposition of solutions to heat equations starting at times $0, t_1, t_2, \ldots$. So we start with an initial "heat profile" which is spread out, and at every Poisson time t_i a profile $f(\cdot)X_i$ is superposed upon the existing profile and will start to spread out according to the heat equation from time t_i on.

Note that the way the process is chosen in the above two examples and the way B maps this into the function space is chosen in a very specific way. Of course a Banach space-valued process and a far more general way of mapping this into the proper function space wouldn't significantly influence the above analysis. As one might expect, both solutions look like superpositions of solutions of wave cq. heat equations starting at different times. This motivates us to solve these equations in an even more general setting. So, let $J(t)$ be the jump process from section 2.3 in its most general E-valued form and let B be a bounded operator in $\mathscr{L}(E, E')$. Now the mild solution to the equation for the E'-valued process $X(t)$:

$$\begin{cases} dX(t) = AX(t)dt + BdJ(t) & t \in [0,T] \\ X(0) = X_0 \end{cases} \tag{3.8}$$

can be given by[1]:

$$\begin{aligned} X(t) &= S(t)X_0 + \int_0^t S(t-s)BdJ(s) \\ &= S(t)X_0 + \sum_{i:t_i \le t} S(t-t_i)BX_i \end{aligned}$$

Which is completely in line with our expectations.

3.2 More general SACP's

Equations like (3.1) are quite interesting to study in their own right, yet, in most cases, they are far too simplistic to be useful in practical situations. Therefore, we allow the equations to become more complicated and inhomogeneous. The most general of these is the inhomogeneous SACP with multiplicative noise for an E-valued process $U(t)$ on $[0,T]$:

$$\begin{cases} dU(t) = [AU(t) + F(t, U(t))]\, dt + B(t, U(t))dL(t) & \text{for } t \in [0,T] \\ U(0) = U_0 \end{cases} \tag{3.9}$$

Where A is a linear operator from E to itself, $F(\cdot,\cdot)$ a not necessarily linear function from $[0,T] \times E$ to E and $B(\cdot,\cdot)$ is a function from $[0,T] \times E$ to the space of bouded operators $\mathscr{L}(E', E)$ and $L(t)$ is an E'-valued (Lévy) process. Again our specific approach will be to take the jump process as our leading process and I will try to establish existence and/or uniqueness of mild solutions.

Definition 3.5. An E-valued process $U(t)$ is called a mild solution of (3.9) if and only if it satisfies the following integral equation:

$$U(t) = S(t)U_0 + \int_0^t S(t-s)F(s, U(s))ds + \int_0^t S(t-s)B(s, U(s))dL(s) \tag{3.10}$$

Where $S(t)$ is the semigroup generated by A

Note now that the mild solution isn't clearly and unambiguously defined as it was in the homogeneous, additive case. It's not even a priori certain if a process satisfying (3.10) exists. Also the connection between the mild solution and the SACP itsel is harder to make, because the nonlinearity of (3.9) doesn't allow a proof for uniqueness as we used to prove lemma 3.1.

We need to go about proving uniqueness in a very different way, first we look at a specific realisation of the process, i.e. we fix a certain point in our probability space and the corresponding path of our jump process (i.e., a specific sequence of t_i's and X_i's). This gives

[1] where for $A \in \mathscr{L}(E')$ and $B \in \mathscr{L}(E, E')$ we write $AB := (A \circ B) \in \mathscr{L}(E, E')$

(with probability one) a finite number of subintervals $[0, t_1], [t_1, t_2], \ldots, [t_{n-1}, t_n], [t_n, T]$. First we restrict equation (3.10) to $[0, t_1]$, giving the deterministic equation:

$$U(t) = S(t)U_0 + \int_0^t S(t-s)F(s, U(s))ds \qquad \text{for } t \in [0, t_1] \tag{3.11}$$

From the theory of evolution equations[2], we know by an application of the Banach contraction theorem that this equation has a unique solution $U : [0, T] \to E$,[3] as long as F is continuous with respect to the first variable and locally Lipschitz continuous with respect to the second variable with at most linear growth, i.e. there exists constants c_1 and c_2, such that for all $t$and x: $\|F(t, x)\| \leq c_1 + c_2|x|$. Now one can look at the equation in $[t_1, t_2]$, knowing that $U(t_1-)$ is uniquely defined. Now we get the equation

$$U(t) = S(t-t_1)U(t_1-) + B(t_1-, U(t_1-))X_1 + \int_{t_1}^t S(t-s)F(s, U(s))ds \qquad \text{for } t \in [t_1, t_2] \tag{3.12}$$

Again this equation has a unique solution from $[t_1, t_2]$ to E. Continuing inductively we see that on every interval $[t_i, t_{i+1}]$ the equation

$$U(t) = S(t-t_i)U(t_i-) + \sum_{k=1}^{i} B(t_k-, U(t_k-))X_k + \int_{t_i}^t S(t-s)F(s, U(s))ds \qquad \text{for } t \in [t_i, t_{i+1}] \tag{3.13}$$

has a unique solution, so we have the following result (probably the most important one in this thesis), stated here as a theorem:

Theorem 3.6. *Let $J(t)$ be a jump process as defined in section 2.3, then the stochastic abstract Cauchy problem*

$$\begin{cases} dU(t) = [AU(t) + F(t, U(t))]\,dt + B(t, U(t))dJ(t) & \textit{for } t \in [0, T] \\ U(0) = U_0 \end{cases} \tag{3.14}$$

in E admits a unique mild solution from $[0, T]$ to E.

Perhaps the most surprising part of theorem 3.6 is that there is no condition what so ever upon the form of B. Yet again, in retrospect the result itself could have been expected. Just like in the additive, homogeneous, linear case, the mild solution is just a superposition of mild solutions all starting at one of the times t_i, etcetera. The ease of the derivation, as well as the mildness of the relevant conditions is in shrill contrast with the theory where the underlying process is a Brownian motion, where very nontrivial results only yield solutions under very hard to verify conditions.

3.3 Musiela's SPDE as a SACP

One of the main applications of stochastic differential equations can be found in financial mathematics. Methods for contingent claim valuation rely heavily on stochastic calculus. The first contribution in this area was made by Black and Scholes in [4], who, by applying Itô's lemma twice, found a PDE that unambiguously fixed the arbitrage-free price for a

[2]see, for example sections 6.1 and 6.2 of [17]

[3]or, more specifically a unique solution in a certain function space from $[0, T]$ to E

european call option on a financial instrument whose value followed a geometric Brownian motion. Although their model used a number of nonrealistic assumptions, the obtained formula proved to be very applicable and powerful.

The Black-Scholes model is very useful to evaluate contingent claims on stocks and stock portfolio's, because the "state" of these instruments are determined solely by their value. Studying bonds requires a far more complicated analysis, because bonds not only have a value at each time, but also have their *maturity*, a time at which the price of the bond is assumed to be fixed. One of the first rigourous mathematical studies of the behaviour of bonds was done by Heath, Jarrow and Morton in [6]. They assumed the bond to be interest-free (zero coupon) and built a model for the so-called *forward rate process*, the logarithm of the derivative of the bond price. Their quite complicated analysis was simplified later by a notational trick. In later literature the forward rate of the bond was modeled as a random function on $\mathbb{R}^2$: $r(t,x)$ was defined to be the forward rate of a bond at time t, with maturity at time $t+x$, instead of working with fixed maturities as the original article did.

By notational preference I will be using $u(t,x)$ instead of $r(t,x)$. The forward rate, according to the Musiela model, now satisfies *Musiela's SPDE:*

$$du(t,x) = \left[\frac{\partial}{\partial x}u(t,x) + f(t,x,u|_{[0,x]}(t,x))\right]dt + \langle \sigma(t,x,u(t,x)), dM(t)\rangle \tag{3.15}$$

In this model σ is the volatitility coefficient, taking values in a Hilbert space H and $M(t)$ is mostly taken to be a locally square integrable semimartingale. A priori, there is no obvious reason not to take $M(t) \in E$ and σ taking values in E' for some Banach space E, but as the most considered process for $M(t)$ is (cylindrical) Brownian motion, the Hilbert space case is the most general one considered. In case this is a Brownian motion and f is independent of t and u, the Musiela SPDE can be written as the type of linear SACP considered in chapter 14 of [11]. An analysis on this equation for a more general form of $M(t)$ can be found in [8].

The case where $M(t)$ is a Brownian motion was, as already mentioned, first considered by Heath, Jarrow and Mortion in [6], who proved that for a no-arbitrage condition to hold, there had to be a very specific relation between the *drift* coefficient f and the *volatility* σ. The astonishingly easy relation they found was that $f = \frac{1}{2}|\sigma|^2$. For other forms for $M(t)$ such a relation can not in general be established.

First of all, we take the bond to be issued at time 0 and have maturity T. In this case, we know that we can restrict our analysis to $x \in [0,T]$ and model $U(t) := u(t,x) \in E$, where E is some function space on $[0,T]$ (e.g an L^p-space). Now we can rewrite (3.15) as an inhomogeneous SACP:

$$dU(t) = [AU(t) + F(t,U(t))]\,dt + B(t,U(t))dM(t) \qquad \text{for } t \in [0,T] \tag{3.16}$$

Where A denotes the differentiation operator, and so on. As the differentiation operator generates the strongly continuous translation semigroup on almost every relevant function space, we know that under a Lipschitz condition for F (and whence for f), equation (3.16) admits a mild solution for any initial forward rate curve.

Chapter 4

Stochastic Integration (II)

The results from previous chapters on the jump process defined might look astonishingly easy for a process that is not even continuous and that might also be quite realistic. From both theoretical and practical considerations it might be useful to extend the process defined in section 2.3 and the corresponding integral to more general and/or realistic processes.

4.1 Geometrical and multiplicative noise

When using a process like the one defined in section 2.3, one has the problem that the changes in the process are purely additive. When $J(t)$ models something that influences the price of a financial instrument, for example, one has to impose some rather strict conditions on the boundedness of the X_i and the way they are injected into the state space to assure positivity of solutions (note that, in the Musiela model, the forward rate is allowed to become negative). In a way, taking a Banach space as the space for $J(t)$ limits the generality of our analyis, because multiplication is generally not allowed for our process. So, so far, a model for e.g. a stock price as a jump process is not possible. Since we do not really have an Itô's lemma for $J(t)$, it is not possible to define $df(J(t))$ if f is a non-affine function, like the exponential. We can conclude that the way Brownian motion is transformed into geometrical Brownian motion doesn't work as well for jump processes.

One might have a "geometrical" real-valued jump process, for $(Y_i)_i$ a sequence of real-valued random variables:

$$I(t) = \prod_{i:t_i \le t} (1 + Y_i),$$

for which the differential will be

$$dI(t) = \left[\sum_{i:t_i \le t} Y_i \delta(t - t_i) \prod_{k=1}^{i-1} (1 + Y_k) \right] dt$$

As can be shown by complete induction and application of the product rule. This expression shows the problem one might have with jump processes: a slight variation in the form of the process leads to a very complicated analysis that has to be done for each variation and yields a very complicated and hard to use result.

4.2 Brownian motion

We consider the real-valued jump process $(J^\epsilon(t))_{t\geq 0}$ that jumps at fixed deterministic times $n\epsilon$, $n = 1, 2, \dots$ and makes a normally distributed jump with fixed variance ϵ that is independent of all previous jumps[1]. This process is both a martingale and a Markov process. The idea is of course to let the ϵ tend to zero in some way and then apply the central limit theorem to look at the limit in distribution of the random variables corresponding to the resulting process at some time. This limiting procedure gives a very intuitive way to think about Brownian motion as being the limit of a random walk making infinitely small steps at infinitely small time intervals.

The theoretic backup for this limiting procedure is given by *Donsker's Theorem*, that can be found in among others [12]. First we consider the vector spaces $C := C[0, 1]$ and $D := D[0, 1]$ of all continuous and all càdlàg functions on $[0, 1]$, respectively. Of course $C \subset D$ and C is a separable Banach space with the supremumnorm (giving rise to the uniform topology). The space D is given the so-called *relative Skorohod topology*, which is defined as follows: first we define the class Λ of strictly increasing continuous onto functions from $[0, 1]$ to itself. Then we define convergence:

Definition 4.1. A sequence of functions $(x_n)_n$ in D is said to converge to $x \in D$ with respect to the relative Skorohod topology if there exists a sequence $(\lambda_n)_n$ in Λ such that $\lambda_n(\cdot)$ tends to the identity and $x_n(\lambda_n(\cdot)) \to x(\cdot)$, both uniformly, as n tends to infinity.

By taking $\lambda_n(t) := t$, we see that by definition the relative Skorohod topology coincides with the uniform topology when restricted to C. The relative Skorohod topology is induced by the metric

$$d(x, y) := \inf \left\{ \epsilon > 0 \;\middle|\; \begin{array}{c} \text{there exists } \lambda \in \Lambda \text{ such that} \\ \sup_{0\leq t\leq 1} |\lambda(t) - t| \leq \epsilon \text{ and} \\ \sup_{0\leq t\leq 1} |y(\lambda(t)) - x(t)| \leq \epsilon \end{array} \right\} \qquad \text{for } x, y \in D \tag{4.1}$$

If we replace the first condition on ϵ and λ in the above expression by

$$\sup_{s\neq t} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right| \leq \epsilon,$$

one obtains a metric d_0 that is equivalent to d, but that makes D into a complete metric space. The problem with D is that it is not a separable Banach space with respect to any norm or metric, so, although we can define random elements in D as $\mathscr{F}$-measurable functions from a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ to D, we can not define their expectation.

Brownian motion can be viewed as a random element of the vector spaces C *and* D. On $(C, \mathscr{B}(C))$ one can define its distribution as the so-called *Wiener measure* W. Now having a topology we can define the relevant process and its convergence to "Brownian Motion". Now for every element of our probability space we can define the jump process $X_n(t, \omega)$ as follows: we define a sequence of independent identically distributed variables $(\xi_n(\omega))_n$ with expectation zero and fixed and existent variance σ, then we define the $X_n(\cdot, \omega)$ in D as follows:

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} \xi_i(\omega)$$

[1]several extensions can be made that will probably still work. For example, making Gaussian jumps in a Hilbert space with a fixed covariance operator or making independent identically distributed real-valued jumps with fixed (and existent) variance

Now we are ready to state Donsker's theorem without proof[2].

Theorem 4.2. *Let the sequence X_n of random variables in D be defined as above, then the random variables X_n in D converge in distribution to the Wiener process.*

Of course, we are able to define for any random function $f(\cdot, \omega)$ in D, and for every n the integral (leaving aside the notational dependence on ω:

$$\int_0^1 f(t)dX_n(t) := \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} f(\frac{i}{n}-)\xi_i$$

Now the question is if, in which way and towards which stochastic variable this integral converges.

A first attempt may be to consider the ξ_i to be standard Gaussian random variables, and the function f to be deterministic and continuous on $[0,1]$

$$\int_0^1 f(t)dX_n(t) := \sum_{i=1}^{n} f(\frac{i}{n}-)\frac{\xi_i}{\sqrt{n}}$$

is identically distributed as

$$\sum_{i=1}^{n} f(\frac{i}{n}-)\left(B_{\frac{i+1}{n}} - B_{\frac{i}{n}}\right)$$

Which is a step function approximation to the Itô integral (cf. for example [16]). So in this case there will be convergence in distrbution towards the Itô integral. This is quite a simple argument, motivating extension of this analysis towards more general integrands, jump processes or trying to assert stronger modes of convergence.

The first extension to this analysis might be to consider, instead of deterministic jumps after a time ϵ, jumps that occur according to a Poisson process with intensity ϵ, which will hopefully yield a same result.

4.3 Processes with countably many jumps

A more general question when it comes to stochastic integration is how to define a suitable stochastic integral for more general càdlàg Lévy processes, that are allowed to make infinitely many jumps. This question is addressed among others in [9]. I will try and address this question from a new angle. I will consider a Lévy process $L(t)$ on a Banach space E' making countably many jumps X_i on $[0,T]$ such that (almost) surely the sum of all discontinuities is finite:

$$\sum_i \|X_i\| < \infty. \tag{4.2}$$

Now I define for $\epsilon > 0$ the process $L^\epsilon(t)$ to be the sum of all discontinuities up to time t for which the difference between left and right limits is more than ϵ. Now for ϵ we define X^ϵ to be the mild solution of the following SACP:

$$\begin{cases} dX(t) = AX(t)dt + B(t)dL^\epsilon(t) & t \in [0,T] \\ X(0) = x & \text{for some } x \in E \end{cases} \tag{4.3}$$

[2] also, many subtleties are involved when dealing with the modes of convergence of random variable in abstract spaces.

Where we assume $B(\cdot)$ to be a continuous function from $[0, T]$ to $\mathscr{L}(E', E)$. We know that a solution X^ϵ exists for each $\epsilon > 0$.

Lemma 4.3. *The mild solutions $(X^\epsilon)_{\epsilon>0}$ form a Cauchy family in the space of càdlàg functions from $[0, T] \to E$ endowed with the topology of uniform convergence.*

Proof. We start by working with $0 < \delta < \epsilon$, such that δ tends to zero as ϵ does. Our lemma is proven if we can show that

$$\sup_{0\leq t\leq T} \|U^\epsilon(t) - U^\delta(t)\|$$

tends to zero as ϵ tends to zero.

Because the first propagation term in the expression for the mild solution cancels out, we can write down:

$$\sup_{0\leq t\leq T} \|U^\epsilon(t) - U^\delta(t)\| \quad = \quad \sup_{0\leq t\leq T} \left\| \int_0^t S(t-s)B(s) \left[dL^\epsilon(s) - dL^\delta(s)\right] \right\|$$

Now all jumps smaller than δ do not feature in this integral and all jumps larger than ϵ cancel out, so what remains is.

$$\sup_{0\leq t\leq T} \|U^\epsilon(t) - U^\delta(t)\| \quad = \quad \sup_{0\leq t\leq T} \left\| \sum_{i:t_i\leq t \text{ and } \delta\leq\|X_i\|<\epsilon} S(t-t_i)B(t_i)X_i \right\|$$

First we take the norm within the sum and then see that the supremum of resulting expression is attained when all of the jumps are taken into account:

$$\begin{aligned}\sup_{0\leq t\leq T} \|U^\epsilon(t) - U^\delta(t)\| \quad &\leq \quad \sum_{\delta\leq\|X_i\|<\epsilon} \|S(t-t_i)\|\|B(t_i)\|\|X_i\| \\ &\leq \quad \sum_{\delta\leq\|X_i\|<\epsilon} K\|X_i\|\end{aligned}$$

for some constant K, now we can drop the $X_i \geq \delta$-condition in the sum at the cost of an inequality.

$$\sup_{0\leq t\leq T} \|U^\epsilon(t) - U^\delta(t)\| \quad \leq \quad \sum_{\|X_i\|<\epsilon} K\|X_i\|$$

Because all the elements of this series are positive, we can renumber them in decreasing order. Then letting ϵ tend to zero is essentially taking the tail of the sequence for increasingly higher indices, so this expression tends to zero as ϵ tends to zero. QED

Next we want to consider a nonadditive SACP and try and do the same analysis for this version of the problem, so we define U^ϵ as the unique mild solution to the equation.

$$\begin{cases} dU(t) = AU(t)dt + B(t, U(t))dL^\epsilon(t) & t \in [0, T] \\ U(0) = u & \text{for some } u \in E \end{cases} \tag{4.4}$$

Again we want to prove the same result as we did for additive noise, which I will formulate here as a proposition:

Proposition 4.4. *There exist conditions on A and B, and the space E, such that for almost surely any path of the process $L(t)$ the family of mild solutions $(U^\epsilon)_{\epsilon>0}$ to equation* (4.4) *forms a Cauchy family in the space of càdlàg functions from $[0,T]$ to E, with respect to some topology.*

To give an idea of the bounds that might play a role we assume that B is Lipschitz in the second variable with a Lipschitz constant that is bounded for $t \in [0,T]$ by some constant K. The problem with proving this proposition is that differences between the two solutions not only come from the differences in the jumps contributing to the respective integrals, but also from jumps that feature in both integrals that are mapped into the state space in a different way. We call $t_1, t_2, \dots$ with $t_i < t_{i+1}$ the jump times of the process $L(t)$, with the respective jumps $Y_{t_1}, Y_{t_2}, \dots \in E'$.

Again we have $0 < \delta < \epsilon$ and the difference between the two corresponding processes will now be

$$U^\epsilon(t) - U^\delta(t) = \int_0^t S(t-s)\left[B(s,U^\epsilon(s))dL^\epsilon(s) - B(s,U^\delta(s))dL^\delta(s)\right].$$

We want to mark the jumps that contribute to the difference between U^ϵ and U^δ. The times of these jumps will be denoted $\tilde{t}_1, \tilde{t}_2, \dots$. Up to the first jump that features in L^δ but not in L^ϵ the processes will be equal, so $\tilde{t}_1 = \min\{t_i : \delta \le Y_{t_i} < \epsilon\}$, from then on, every jump will matter, so for $i \ge 2$ we have

$$\tilde{t}_i = \min\left\{t_j, j = 1,2,\dots \mid t_j > \tilde{t}_{i-1} \text{ and } Y_{t_j} \ge \delta\right\}$$

Now we define V_i to be the size of the difference between the two processes at (or right after) $\tilde{t}_i$:

$$V_i = \left\| U^\epsilon\left(\tilde{t}_i\right) - U^\delta\left(\tilde{t}_i\right)\right\| \tag{4.5}$$

Now we see that V_1 is smaller than some constant (dependent on B) times ϵ, and for each $i \ge 2$ the jump can be either caused by a jump bigger than epsilon, in which case the difference between the two solutions just before $\tilde{t}_i$ is just a propagation of the previous jump, in which case:

$$V_i \le (1 + K\|X_i\|)\exp\left[\mu\left(\tilde{t}_i - \tilde{t}_{i-1}\right)\right]V_{i-1} \tag{4.6}$$

On the other hand, it could be a jump smaller than ϵ, in which case we only know that

$$V_i \le \left\|B\left(\tilde{t}_i, U^\delta(\tilde{t}_i)\right)\right\| \|X_i\| + \exp\left[\mu\left(\tilde{t}_i - \tilde{t}_{i-1}\right)\right]V_{i-1} \tag{4.7}$$

This leads to a recursive inequality that motivates us to reformulate proposition 4.4 as follows:

Proposition 4.5. *Consider equation* (4.4) *and assume the following conditions are met:*

1. *The operator A generates a strongly continuous contractive semigroup*
2. *$B(t,\cdot)$ is Lipschitz continuous with constant $C(t)$ and $K := \sup_{0\le t\le T} C(t) < \infty$*
3. *There exists a constant b such that for all t in $[0,T]$ and x in E: $\|B(t,x)\| < b$*

Then for almost each path of the process $L(t)$ the family of mild solutions $(U^\epsilon)_{\epsilon>0}$ to equation (4.4) *forms a Cauchy family in the space of càdlàg functions from $[0,T]$ to E, with respect to the topology of uniform convergence.*

Proof. Let U^ϵ, U^δ and V_i be as before. We begin by combining the inequalities (4.6) and (4.7) into

$$V_i \leq (1+\alpha_i)\, V_{i-1} + \beta_i \tag{4.8}$$

With $V_0 = 0$, and α_i and β_i defined by:

$$\alpha_i := \begin{cases} K\|X_i\| & \text{for } \|X_i\| \geq \epsilon \\ 0 & \text{for } \|X_i\| < \epsilon \end{cases}$$

and

$$\beta_i := \begin{cases} 0 & \text{for } \|X_i\| \geq \epsilon \\ b\|X_i\| & \text{for } \|X_i\| < \epsilon \end{cases}$$

So $\sum_i \alpha_i \leq \sum_i K\|X_i\| < \infty$ and for $\sum_i \beta_i$ an even better bound exists:

$$\sum_i \beta_i \leq b \sum_{\|X_i\|<\epsilon} \|X_i\|.$$

Furthermore $\alpha_1 = 0$.

Now we define $(z_i)_i$ to be the unique solution of the equation

$$\begin{cases} z_i = (1+\alpha_i)\, z_{i-1} + \beta_i \\ z_0 = 0 \end{cases} \tag{4.9}$$

So $V_i \leq z_i$. Now one can prove by induction that

$$z_i = \sum_{j=1}^{i} \beta_j \prod_{l=j}^{i} (1+\alpha_l)$$

Introducing the variable $\gamma_l := \ln(1+\alpha_l) \leq \alpha_l$, we can write:

$$\prod_{l=j}^{i}(1+\alpha_l) = \exp\left(\sum_{l=j}^{i} \gamma_l\right)$$

which can be estimated by

$$\begin{aligned} \exp\left(\sum_{l=j}^{i} \gamma_l\right) &\leq \exp\left(\sum_{l=j}^{i} \alpha_l\right) \\ &\leq \exp\left(K \sum_{l=j}^{i} \|X_l\|\right) \\ &=: R \end{aligned}$$

So we have established that

$$V_i \leq \sum_{j=1}^{i} R\beta_j$$

So the supremum of the difference between U^ϵ and U^δ is bounded as follows:

$$\begin{aligned}\sup_{0\leq t\leq T}\|U^\epsilon(t)-U^\delta(t)\| &\leq \sup_i V_i \\ &\leq R\sum_{j=1}^{i}\beta_j \\ &\leq R\sum_{j=1}^{\infty}\beta_j \\ &\leq bR\sum_{j:\|X_j\|<\epsilon}\|X_j\|\end{aligned}$$

As we saw before, this last expression tends to zero, so $(U^\epsilon)_{\epsilon>0}$ forms a Cauchy family in $D([0,T],E)$ with the topology of uniform convergence. QED

Chapter 5

Conclusive remarks

As I went about the analysis from a very new angle, leaving Brownian motion aside and purely considering jump processes, I obtained some results that might not be that astounding or shocking, yet some of the obtained expressions might turn out to be useful in future analysis. The last result, which I obtained in co-operation with my thesis advisor, Onno van Gaans, gives something of a new step in the theory of stochastic equations with Lévy noise. By using more sophisticated methods, maybe even stronger results can be obtained. As the nature of this thesis was very analytic, using more of the stochastic properties might yield some very interesting results.

Other option for future research are being held possible by the highly nonconstructive nature of our analysis. Although one can conclude existence results for solution to certain quite general classes of processes from the results in this thesis, uniqueness does not follow and very few properties of the solutions are known. Deeper analysis, use of the stochastic properties of the processes and good numerical simulations might yield a better insight towards problems concerning Lévy noise. Also the Cauchy property derived in section 4.3 gives us a motivation that (finite) numerical simulations can be used for modeling the behaviour of solutions.

As to the jump process itself, specific forms of the X_i might yield some interesting results. Also, statistic analysis can be used to model real-life processes (e.g. stock prices, day-to-day interest rates, etc.) as jump processes and determine realistic distributions for the X_i in order to make the results more applicable. An interesting analysis might be to consider the Musiela PDE with jump process noise and trying to determine a relation between drift and volatility so as to make the forward rate process satisfy a no arbitrage condition.

Bibliography

[1] Louis Bachelier. Theorie de la spéculation. *Annales Scientifiques de l'ENS*, 17:21–68, 1900.

[2] Albert Einstein. Zur theorie der brownschen bewegung. *Ann. Phys. (Leipzig)*, 19:371–381, 1906.

[3] Kiyosi Itô. Stochastic integral. *Proceedings of the Imperial Academy*, 20(8), 1944.

[4] Fisher Black and Myron Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3), 1973.

[5] J. Michael Harrison and Stanley R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3):215–260, Aug 1981.

[6] David Heath, Robert Jarrow, and Andrew Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1), January 1992.

[7] J.M.A.M. van Neerven and L. Weis. Stochastic integration of functions with values in a banach space. *Studia Math*, 166, 2005.

[8] Carlo Marinelli. Local well-posedness of musiela's spde with lévy noise. April 2007.

[9] Markus Riedle and Onno van Gaans. Stochastic integration for lévy processes with values in banach spaces. July 2007.

[10] J.M.A.M. van Neerven, M.C. Veraar, and L. Weis. Stochastic integration in umd banach spaces. *Ann Probab.*, 35, 2007.

[11] J.M.A.M. van Neerven. Internet seminar on stochastic evolution equations, 2008.

[12] Patrick Billingsley. *Convergence of Probability Measures*. Wiley series in probability and mathematical statistics. John Wiley & Sons, Inc., 1968.

[13] Giuseppe da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 1992.

[14] R.M. Dudley. *Real analysis and probability*. Cambridge studies in advanced Mathematics. Cambridge University Press, 2002.

[15] Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics. Springer Verlag, 2000.

[16] B. Øksendal. *Stochastic Differential Equations.* Springer Verlag, 1985.

[17] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Applied Mathematical Sciences. Springer Verlag, 1983.

[18] Martin Schechter. *Principles of Functional Analysis.* Graduate Studies in Mathematics. American Mathematical Society, 2000.

[19] Laurent Schwartz. *Théorie des Distributions.* Hermann, 1978.

[20] Albert Nikolaevich Shiryaev. *Essentials of Stochastic Finance.* Advanced Series on statistical science & applied probability. World Scientific, 1999.

[21] J. Michael Steele. *Stochastic Calculus and Financial Applications.* Applications of Mathematics. Springer Verlag, 2001.

[22] David Williams. *Probability with Martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, 1991.