Contents

Chapter 1

Introduction

Stochastic differential equations (SDEs) are studied for various reasons. One of them is the fact that one can express the solution of certain parabolic second order partial differential equations (PDEs) by the solution of a corresponding SDE, more exactly by its expected value. This leads to new methods of solving a PDE numerically by using Monte Carlo simulation. However, this approach gets numerically problematic, especially if the dimension of the space is rather high.

In [9] Terry Lyons and Nicolas Victoir describe a new method of solving PDEs numerically by exploiting the equivalence of solving a PDE and a certain SDE. The method is motivated by the use of cubature formulae in numerical integration in finite dimensions. Iterated Stratonovich integrals of Brownian motion play the role of polynomials in classical cubature formulae, i. e. the cubature formula is exact for iterated Stratonovich integrals of a given order. Since the solution of the SDE can be approximated by finite sums of iterated Stratonovich integrals, one gets an approximation to the expected value of the solution of the SDE.

By doing some kind of Laplace transformation we get an extension of that method applicable for some higher order differential operators like the bi-Laplacian.

Special thanks go to Josef Teichmann, the supervisor of this work, who was exceptionally friendly, helpful and encouraging throughout the work. Of course, he also gave useful advice and explanations.

Chapter 2

Stochastic Analysis

2.1 Introduction

We repeat some of the basic definitions and concepts of stochastic analysis in this section, rather in order to fix notations than to give a proper introduction into that subject. The reader is referred to [12] for proofs of the concepts mentioned here.

With $\mathbb{N}$ we denote the set of the natural numbers $\{0, 1, 2, \ldots\}$. Fix some $d \in \mathbb{N} \setminus \{0\}$. Let $(\Omega, \mathcal{F}, P)$ be a probability space "rich enough". A stochastic process $W : \mathbb{R}_{\geq 0} \times \Omega \to \mathbb{R}$, where $\mathbb{R}_{\geq 0} = [0, \infty[$, is called *Wiener process* if W is *Gaussian*, i. e. any linear combination $\sum_{i=1}^{n} c_i W(t_i)$ with $c_i \in \mathbb{R}$, $t_i \in \mathbb{R}_{\geq 0}$, $i = 1, \ldots, n$, and $n \in \mathbb{N}$ is a Gaussian random variable, if $W(0) = 0$ almost surely and if $E(W(t)) = 0$, $E(W(t)W(s)) = \min(t, s)$ for all $t, s \in \mathbb{R}_{\geq 0}$, where $E(\cdot)$ denotes expectation with respect to P. By the Kolmogorov-Centsov Theorem, see [1], there is a version of the Wiener process having continuous paths, i. e. there is a stochastic process $\tilde{W}$ with $P\left(W(t) = \tilde{W}(t)\right) = 1$ for all $t \in \mathbb{R}_{\geq 0}$ such that $t \mapsto W(t)(\omega)$ is continuous for all $\omega \in \Omega$. $W(t) = (W_1(t), \ldots, W_d(t))$, $t \in \mathbb{R}_{\geq 0}$, is a d-dimensional Wiener process if the coordinates $W_1(t), \ldots, W_d(t)$, $t \in \mathbb{R}_{\geq 0}$, are independent Wiener processes.

A *filtration* is a family $(\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}$ of σ-algebras $\mathcal{F}_t \subseteq \mathcal{F}$ with $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$. $\left(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}\right)$ is called *filtered probability space.* $(\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}$ satisfies the *usual conditions* if $\mathcal{F}_0$ contains all sets with P-measure 0 of $\mathcal{F}$ and if $\mathcal{F}_t = \bigcap_{u > t} \mathcal{F}_u$ for all $t \in \mathbb{R}_{\geq 0}$. A stochastic process $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ is *adapted* to $(\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}$ if the random variables X_t are $\mathcal{F}_t$-measurable for all $t \in \mathbb{R}_{\geq 0}$.

Let $\left(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}\right)$ be a filtered probability space, such that $\mathcal{F}$ is the σ-algebra generated by $\bigcup_{t=0}^{\infty} \mathcal{F}_t$. A Wiener process $(B_t)_{t \in \mathbb{R}_{\geq 0}}$ with values in $\mathbb{R}^d$ is called *Brownian motion* if

1. $(B_t)_{t \in \mathbb{R}_{\geq 0}}$ is adapted to $(\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}$ and has continuous paths.

2. The filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}$ satisfies the usual conditions.

3. For $s \geq t$ the increments $B_s - B_t$ are independent of $\mathcal{F}_t$.

We may assume that $(\mathcal{F}_t)_{t \in \mathbb{R}_{\geq 0}}$ is the natural filtration of the Brownian motion, i. e. that for each $t \in \mathbb{R}_{\geq 0}$ $\mathcal{F}_t$ is the σ-algebra generated by $(B_s : 0 \leq s \leq t)$.

Therefore, for fixed $\omega \in \Omega$, the map $\mathbb{R}_{\geq 0} \to \mathbb{R}^d$, $t \mapsto B_t(\omega)$ is a continuous map with $B_0(\omega) = 0$. We denote the space of all such functions with $C_0^0\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right)$, i. e.

$$C_0^0\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right) = \left\{ f : \mathbb{R}_{\geq 0} \to \mathbb{R}^d \middle| f(0) = 0,\ f \text{ continuous} \right\}.$$

Brownian motion interpreted as a map from Ω to $C_0^0\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right)$ thus induces a probability measure on $C_0^0\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right)$ endowed with its Borel σ-algebra. Then the stochastic process $(ev_t)_{t \in \mathbb{R}_{\geq 0}}$, defined on the probability space $\left(C_0^0\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right), \mathcal{F}, P\right)$ by $ev_t(f) = f(t)$ for $f \in C_0^0\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right)$, where $\mathcal{F}$ is the Borel σ-algebra of $C_0^0\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right)$ and P is the probability measure induced by Brownian motion, satisfies all the requirements of Brownian motion.

$\left(C_0^0\left(\mathbb{R}_{\geq 0}, \mathbb{R}^d\right), \mathcal{F}, P\right)$ is called *Wiener space*. By the Brownian motion B_t we understand the evaluation functional ev_t, B_t^i denotes the i-th coordinate of B_t. Furthermore we define $B_t^0 = t$ for $t \in \mathbb{R}_{\geq 0}$. Analoguously we define the Wiener space $\left(C_0^0\left([0,T], \mathbb{R}^d\right), \mathcal{F}, P\right)$ and Brownian motion thereon, where $T > 0$.

For $N \in \mathbb{N} \setminus \{0\}$, $C_b^\infty\left(\mathbb{R}^N, \mathbb{R}^N\right)$ is the set of all functions $f : \mathbb{R}^N \to \mathbb{R}^N$, such that derivatives of any order exist and are bounded. By a *vector field* on $\mathbb{R}^N$ we understand such a function. More exactly, we identify $f \in C_b^\infty\left(\mathbb{R}^N, \mathbb{R}^N\right)$ with the vector field $\sum_{j=1}^N f^j(x) \frac{\partial}{\partial x^j}$, where $\frac{\partial}{\partial x^j}$ denotes the j-th partial derivative, i. e. $\frac{\partial}{\partial x^j} g = \frac{\partial g}{\partial x^j}$ for a smooth function $g : \mathbb{R}^N \to \mathbb{R}$.

The stochastic process $\xi_{t,x}$, $0 \leq t \leq T$, defined on the probability space $\left(C_0^0\left([0,T], \mathbb{R}^d\right), \mathcal{F}, P\right)$ with

$$\begin{cases} d\xi_{t,x} = \sum_{i=0}^d V_i(\xi_{t,x}) \circ dB_t^i \\ \xi_{0,x} = x \end{cases}, \tag{2.1}$$

is called *solution to the SDE corresponding to the vector fields* $V_0, V_1, \ldots, V_d$. Here $\circ$ indicates that the SDE is understood in the Stratonovich sense, see [12]. $x \in \mathbb{R}^N$ is the initial condition of the SDE. Under the conditions on vector fields stated above, such a solution indeed always exists and is uniquely determined (in L^2). It can be constructed using a Picard-Lindelöf kind of iteration that converges in L^2. Therefore a mapping from $\left(C_0^0\left([0,T], \mathbb{R}^d\right), \mathcal{F}, P\right)$ to $\mathbb{R}^N$ is defined by

$$\Phi_{T,x}(\omega) = \xi_{T,x}(\omega). \tag{2.2}$$

For $\omega \in C_{0,bv}^0\left([0,T], \mathbb{R}^d\right) = \left\{\omega \in C_0^0\left([0,T], \mathbb{R}^d\right) \mid \omega \text{ has bounded variation}\right\}$, (2.1) makes sense as a pathwise differential equation, too, i. e.

$$dy_{t,x} = \sum_{i=0}^d V_i(y_{t,x}) d\omega_t^i \tag{2.3}$$

has a uniquely determined solution $y_{t,x}$, $t \in [0,T]$, with $y_{0,x} = x$. For such a ω the Picard-Lindelöf iteration mentioned above converges pointwise. Since $P\left(C_{0,bv}^0\left([0,T], \mathbb{R}^d\right)\right) = 0$, we may choose a version of the stochastic process $\xi_{t,x}$ satisfying $\xi_{t,x}(\omega) = y_{t,x}$, where $y_{t,x}$ is the solution of the ODE (2.3) corresponding to $\omega \in C_{0,bv}^0\left([0,T], \mathbb{R}^d\right)$ for any such ω. Whenever we consider the solution to a SDE, we will assume that it indeed satisfies this property.

$\xi_{t,x}$ is strongly related to the solution of the parabolic PDE

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) = -Lu(t,x) \\ u(T,x) = f(x) \end{cases}, \tag{2.4}$$

where the differential operator L is defined as

$$L = V_0 + \frac{1}{2}\sum_{j=1}^{d} V_j^2 \tag{2.5}$$

and f is a smooth (i. e. infinitely often differentiable) or Lipschitz function. In (2.5) the square of a vector field V is the operator $V^2 f = V(Vf)$. Indeed, $(t,x) \mapsto E\left(f(\xi_{T-t,x})\right)$ defines the solution to the PDE (2.4).

2.2 Stochastic Taylor Approximation

For $f \in C^\infty(\mathbb{R}^N)$, we have

$$f(\xi_{t,x}) = f(x) + \sum_{i=0}^{d} \int_0^t V_i f(\xi_{s,x}) \circ dB_s^i, \tag{2.6}$$

where, of course, $\xi_{t,x}$ denotes the solution of the Stratonovich SDE corresponding to the vector fields $V_0, V_1, \ldots, V_d$ with initial value x.

Definition 2.2.1. For $m \in \mathbb{N}$ define

$$\mathcal{A}_m = \left\{ (i_1, \ldots, i_k) \in \{0, \ldots, d\}^k \middle| k \in \mathbb{N}, k + |\{j | 1 \le j \le k, i_j = 0\}| \le m \right\},$$

where $|\cdot|$ denotes the cardinality of a set.

The so defined sets $\mathcal{A}_m$ are needed in order to formulate the following stochastic Taylor approximation. First we introduce the following notation:

$$\iint\limits_{0<t_1<\cdots<t_k<t} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k} = \int_0^t \int_0^{t_k} \cdots \int_0^{t_2} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k}.$$

Proposition 2.2.2. *Let $f \in C^\infty(\mathbb{R}^N)$ and $m \in \mathbb{N}$. Let $\xi_{t,x}$ be the solution of the Stratonovich SDE corresponding to the vector fields $V_0, \ldots, V_d$ with initial value $x \in \mathbb{R}^N$. Then*

$$f(\xi_{t,x}) = \sum_{(i_1,\ldots,i_k) \in \mathcal{A}_m} V_{i_1} \cdots V_{i_k} f(x) \iint\limits_{0<t_1<\cdots<t_k<t} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k} + R_m(t,x,f) \tag{2.7}$$

with a remainder process R_m satisfying

$$\sup_{x \in \mathbb{R}^n} \sqrt{E(R_m(t,x,f)^2)} \le C t^{\frac{m+1+\mathbf{1}_{t>1}}{2}} \sup_{(i_1,\ldots,i_k) \in \mathcal{A}_{m+2} \setminus \mathcal{A}_m} \|V_{i_1} \cdots V_{i_k} f\|_\infty$$

with a constant C depending on d and m.

Proof. In a first step we show (by induction) that the remainder can be written as

$$R_m(t,x,f) = \sum_{\substack{(i_1,\dots,i_k)\in\mathcal{A}_m \\ (i_0,i_1,\dots,i_k)\notin\mathcal{A}_m}} \iint\limits_{0<t_0<\dots<t_k<t} V_{i_0}\cdots V_{i_k} f(\xi_{t_0,x}) \circ dB^{i_0}_{t_0} \cdots \circ dB^{i_k}_{t_k}. \tag{2.8}$$

For $m = 0$ we just have to look at (2.6). Now assume the formula is true for $m \in \mathbb{N}$, i. e. we have

$$f(\xi_{t,x}) = T_m(t,x,f) + R_m(t,x,f),$$

where

$$T_m(t,x,f) = \sum_{(i_1,\dots,i_k)\in\mathcal{A}_m} V_{i_1}\cdots V_{i_k} f(x) \iint\limits_{0<t_1<\dots<t_k<t} \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k}$$

and $R_m(t,x,f)$ as in (2.8). By replacing f with $V_{i_0}\cdots V_{i_k} f$ in (2.6) we get

$$V_{i_0}\cdots V_{i_k} f(\xi_{t_0,x}) = V_{i_0}\cdots V_{i_k} f(x) + \sum_{i=0}^{d} \int_0^{t_0} V_i V_{i_0}\cdots V_{i_k} f(\xi_{s,x}) \circ dB^i_s.$$

Applying this identity to the induction hypothesis (2.8) yields

$$\begin{aligned} f(\xi_{t,x}) = T_m(t,x,f) &+ \sum_{\substack{(i_1,\dots,i_k)\in\mathcal{A}_m \\ (i_0,i_1,\dots,i_k)\notin\mathcal{A}_m}} V_{i_0}\cdots V_{i_k} f(x) \iint\limits_{0<t_0<\dots<t_k<t} \circ dB^{i_0}_{t_0} \cdots \circ dB^{i_k}_{t_k} + \\ &+ \sum_{\substack{(i_1,\dots,i_k)\in\mathcal{A}_m \\ (i_0,i_1,\dots,i_k)\notin\mathcal{A}_m}} \sum_{i=0}^{d} \iint\limits_{0<s<t_0<\dots<t_k<t} V_i V_{i_0}\cdots V_{i_k} f(\xi_{s,x}) \circ dB^i_s \circ dB^{i_0}_{t_0} \cdots \circ dB^{i_k}_{t_k}. \end{aligned}$$

For $(i_1,\dots,i_k) \in \{0,\dots,d\}^k$ define $\mathrm{ord}(i_1,\dots,i_k) = k + |\{j | 1 \le j \le k, i_j = 0\}|$. Then $(i_1,\dots,i_k) \in \mathcal{A}_m$ while $(i_0,i_1,\dots,i_k) \notin \mathcal{A}_m$ if and only if $\mathrm{ord}(i_1,\dots,i_k) \le m$ and $\mathrm{ord}(i_0,i_1,\dots,i_k) = \mathrm{ord}(i_1,\dots,i_k) + 1 + \mathbf{1}_{i_0=0} > m$. There are three possibilities for that situation to occur:

1. $\mathrm{ord}(i_1,\dots,i_k) = m$ and $i_0 \in \{1,\dots,d\}$
2. $\mathrm{ord}(i_1,\dots,i_k) = m-1$ and $i_0 = 0$
3. $\mathrm{ord}(i_1,\dots,i_k) = m$ and $i_0 = 0$.

In the first and in the second case $(i_0,\dots,i_k) \in \mathcal{A}_{m+1}$, therefore

$$V_{i_0}\cdots V_{i_k} f(x) \iint\limits_{0<t_0<\dots<t_k<t} \circ dB^{i_0}_{t_0} \cdots \circ dB^{i_k}_{t_k}$$

occurs in $T_{m+1}(t,x,f)$ whereas

$$\iint\limits_{0<s<t_0<\dots<t_k<t} V_i V_{i_0}\cdots V_{i_k} f(\xi_{s,x}) \circ dB^i_s \circ dB^{i_0}_{t_0} \cdots \circ dB^{i_k}_{t_k}$$

occurs in $R_{m+1}(t,x,f)$ for any $i \in \{0,\dots,d\}$.

In the third case, when $(i_1,\dots,i_k) \in \mathcal{A}_m \setminus \mathcal{A}_{m-1}$ and $i_0 = 0$ the corresponding term is

$$V_0 V_{i_1} \cdots V_{i_k} f(x) \iint\limits_{0<t_0<\dots<t_k<t} dt_0 \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k} +$$

$$+ \sum_{i=0}^{d} \iint\limits_{0<s<t_0<\dots<t_k<t} V_i V_0 V_{i_1} \cdots V_{i_k} f(\xi_{s,x}) \circ dB^i_s dt_0 \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k} =$$

$$= \iint\limits_{0<t_0<\dots<t_k<t} \left[V_0 V_{i_1} \cdots V_{i_k} f(x) + \sum_{i=0}^{d} \int_0^{t_0} V_i V_0 V_{i_1} \cdots V_{i_k} f(\xi_{s,x}) \circ dB^i_s \right] \circ dB^{i_0}_{t_0} \cdots \circ dB^{i_k}_{t_k} =$$

$$= \iint\limits_{0<t_0<\dots<t_k<t} V_0 V_{i_1} \cdots V_{i_k} f(\xi_{t_0,x}) dt_0 \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k},$$

where we once more used (2.6) in the last line. Thus the terms corresponding to the third case occur in $R_{m+1}(t,x,f)$, too. Obviously, all the terms in $T_{m+1}(t,x,f)$ and $R_{m+1}(t,x,f)$ either occur in $T_m(t,x,f)$ or correspond to a multi-index in one of the three cases above. Furthermore, no multi-index occurs more than once. Therefore (2.8) is proved.

We now express the Stratonovich integrals in (2.8) by Itô integrals using

$$\int_0^t V_{i_0} \cdots V_{i_k} f(\xi_{t_0,x}) \circ dB^{i_0}_{t_0} =$$

$$= \int_0^t V_{i_0} \cdots V_{i_k} f(\xi_{t_0,x}) dB^{i_0}_{t_0} + \frac{1 - \mathbf{1}_{i_0=0}}{2} \int_0^t V_{i_0}^2 V_{i_1} \cdots V_{i_k} f(\xi_{t_0,x}) dt_0.$$

No correction term occurs while converting a Stratonovich integral with absolutely continuous integrand, so after the conversion we get a sum of terms of the form

$$\iint\limits_{0<t_0<\dots<t_k<t} V_{i_0} \cdots V_{i_k} f(\xi_{t_0,x}) dB^{i_0}_{t_0} \cdots dB^{i_k}_{t_k} \tag{2.9}$$

and

$$\frac{1}{2} \iint\limits_{0<t_0<\dots<t_k<t} V_{i_j} V_{i_0} \cdots V_{i_k} f(\xi_{t_0,x}) dB^{i_0}_{t_0} \cdots dB^{i_{j-1}}_{t_{j-1}} dt_j dB^{i_{j+1}}_{t_{j+1}} \cdots dB^{i_k}_{t_k}. \tag{2.10}$$

The estimate in the Proposition follows from the following inequality

$$E\left(\left(\iint\limits_{0<t_0<\dots<t_k<t} V_{i_0} \cdots V_{i_k} f(\xi_{t_0,x}) dB^{i_0}_{t_0} \cdots dB^{i_k}_{t_k} \right)^2 \right) \le t^{\operatorname{ord}(i_0,\dots,i_k)} \| V_{i_0} \cdots V_{i_k} f \|_\infty^2, \tag{2.11}$$

which we prove by induction. Let h be a bounded continuous function. For $k = 0$ we have to distinguish between two cases. First, assume $i = i_0 = i_k = 0$. We have

$$E\left(\left(\int_0^t h(\xi_{s,x})ds\right)^2\right) \leq E\left(\left(\int_0^t |1| \cdot |h(\xi_{s,x})|ds\right)^2\right) \leq$$

$$\leq E\left(\left(\sqrt{\int_0^t 1^2 ds}\sqrt{\int_0^t h(\xi_{s,x})^2 ds}\right)^2\right) = t\int_0^t E\left(h(\xi_{s,x})^2\right) ds \leq t^2\|h\|_\infty^2,$$

where we used Hölder's inequality. For $i_0 \neq 0$ we simply use the Itô Lemma to get

$$E\left(\left(\int_0^t h(\xi_{s,x})dB(s)\right)^2\right) = \int_0^t E\left(h(\xi_{s,x})^2\right) ds \leq t\|h\|_\infty^2.$$

The induction step looks similar, therefore (2.11) is proved, if we choose $h(\xi_{s,x}) = V_{i_0} \cdots V_{i_k} f(\xi_{s,x})$. For terms of the form (2.10) the same inequality holds, if $\operatorname{ord}(i_0, \ldots, i_k)$ is replaced with $\operatorname{ord}(i_0, \ldots, i_k) + 1$. Proposition 2.2.2 is now proved.

The definition of $\mathcal{A}_m$ is motivated by

$$\iint\limits_{0<t_1<\cdots<t_k<t} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k} =^{\mathcal{L}} \sqrt{t}^{\operatorname{ord}(i_1,\ldots,i_k)} \iint\limits_{0<t_1<\cdots<t_k<1} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k}, \qquad (2.12)$$

where "$=^{\mathcal{L}}$" stands for equality in law. Equation (2.12) holds because the left hand side is the limit of a sum of products of the form $\pm B^{i_1}(ts_n^1) \cdots B^{i_k}(ts_n^k)$, $0 \leq s_n^j \leq 1$. In law, the sum remains unchanged if each factor is replaced by $\pm\sqrt{t}^{\operatorname{ord}(i_1,\ldots,i_k)} B^{i_1}(s_n^1) \cdots B^{i_k}(s_n^k)$. But those sums converge to the right hand side of (2.12).

Chapter 3

Cubature Formulae

As we have seen, approximating the solution of the PDE (2.4) means approximating

$$E(f(\xi_{t,x})) = \int_{\Omega} f(\xi_{t,x})(\omega)P(d\omega).$$

Therefore we want to do some kind of numerical integration with respect to the Wiener measure P on the infinite-dimensional space $\Omega = C_0^0(\mathbb{R}_{\geq 0}, \mathbb{R}^d)$. There are various methods of numerical integration available for integration on finite-dimensional spaces and it seems obvious to try to find an analogous adaption for the infinite-dimensional case. One class of those methods is numerical integration with cubature formulae. In general, the results of numerical integration by cubature formulae, which are also called quadrature formulae, especially in the one-dimensional case, tend to be fairly precise compared to other methods of comparable complexity. However, there are disadvantages, too: This method is not adaptive, i. e. if you want to reduce the error you will have to redo all your calculations in order to get a cubature formula of a higher degree, whereas in the case of many other methods you just have to do some additional calculations together with your previous results.

3.1 Tchakaloff's Theorem

Tchakaloff's Theorem is an existence theorem for cubature formulae for various measures on $\mathbb{R}^d$.

First we need to define the support of a measure. A *Borel measure* is a measure μ defined on the Borel σ-algebra of some Hausdorff space X, such that for all $x \in X$ there is an open neighborhood U of x with $\mu(U) < \infty$. A Borel measure μ is called *Radon measure* if for all Borel sets A

$$\mu(A) = \sup\{\mu(K) | K \subset A,\ K \text{ compact}\}.$$

Definition 3.1.1. Let X be a Hausdorff space and let μ be a Radon measure on the Borel σ-algebra $\mathcal{B}$. Then the *support* of μ is defined by

$$\operatorname{supp}(\mu) = X \setminus O,$$

where O is the biggest open set with μ-measure 0.

This definition actually makes sense, because given the assumptions of Definition 3.1.1 the union of any family of open sets with measure 0 is a set with measure 0, see [5]. Note that any Borel measure on $\mathbb{R}^d$ is a Radon measure, since $\mathbb{R}^d$ is a polish space.

We are now able to give a formal definition of the notion of a cubature formula.

Definition 3.1.2. Let $\mathbb{R}_m[X_1, \ldots, X_d]$ be the space of all real polynomials in d varibles of total degree up to m ($m \in \mathbb{N}$), furthermore let μ be a positive Borel measure on $\mathbb{R}^d$ with support K. The points $x_1, \ldots, x_n \in K$ and the positive scalars $\lambda_1, \ldots, \lambda_n$ define a *cubature formula of degree m with respect to μ* if and only if for all polynomials $P \in \mathbb{R}_m[X_1, \ldots, X_d]$:

$$\int_{\mathbb{R}^d} P(x)\mu(dx) = \sum_{i=1}^{n} \lambda_i P(x_i).$$

$x_1, \ldots, x_n$ are called *nodes* and $\lambda_1, \ldots, \lambda_n$ are called *weights* of the cubature formula. n is called *size* of the cubature formula.

Tchakaloff [15] proved the existence of a cubature formula with degree n, where n is the dimension of the space of polynomials $\mathbb{R}_m[X_1, \ldots, X_d]$, provided that the support of the measure is compact.

Theorem 3.1.3 (Tchakaloff's Theorem). *A) Let μ be a positive, finite Borel measure with compact support in $\mathbb{R}^d$ and let $m \in \mathbb{N}\backslash\{0\}$. Define $N_{m,d} = \dim \mathbb{R}_m[X_1, \ldots, X_d]$. Then there exists a cubature formula of degree m with respect to μ with size $N_{m,d}$.*
B) Let μ be a positive Borel measure on $\mathbb{R}^d$ having convergent moments up to at least degree $m+1$, $m \in \mathbb{N} \setminus \{0\}$. Then there exists a cubature formula for μ of degree m once again with size $N_{m,d}$.

Remark 3.1.4. The boundary in Theorem 3.1.3 can be refined by replacing $N_{m,d}$ with $N_{m,d;\mu} = \dim \left\{ P|_{\mathrm{supp}(\mu)} \middle| P \in \mathbb{R}_m[X_1, \ldots, X_d] \right\}$. The boundary is sharp in the case of a compactly supported measure, as shown in [15], but in special cases there are far "smaller" cubature formulae, e. g. the Gaussian quadrature for $d = 1$ with size $\leq \left[\frac{m}{2}\right] + 1$. The generalization B) is due to Curto and Fialkow [4]. They actually proved the existence of a cubature formula for μ of degree m with size $N_{m,d} + 1$ in case B), but the higher constant is not necessary: The "cubature measure" $\sum_{i=1}^{N_{m,d}+1} \lambda_i \delta_{x_i}$, where δ_x denotes the Dirac measure centered at $x \in \mathbb{R}^d$, $\lambda_i, x_i, i = 1, \ldots, N_{m,d} + 1$ as in Definition 3.1.2, is itself a compactly supported Borel measure. Therefore, we may apply part A) of the Theorem and get a cubature formula for μ of degree m with size smaller than or equal to $N_{m,d}$.

We will only prove Theorem 3.1.3 A) as it is done in [4]. The proof of the generalized version B) can be done using essentially the same method with some additional difficulties, see [4].

An element $\alpha \in \mathbb{N}^d$, where $\mathbb{N} = \{0, 1, \ldots\}$, is called *multi-index*. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$ and some $x = (x^1, \ldots, x^d) \in \mathbb{R}^d$ we define $x^\alpha = (x^1)^{\alpha_1} \cdots (x^d)^{\alpha_d}$ and $|\alpha| = \alpha_1 + \cdots + \alpha_d$, $|\alpha|$ is called *order* of α. A *multisequence* is a sequence indexed by multi-indices.

Note that Tchakaloff's Theorem is closely related to the so called *Truncated Multivariable Moment Problem* (TMMP). Given a multisequence $\beta = \beta^{(m)} = (\beta_\alpha : |\alpha| \leq m)$ a *representing measure* for β is a positive Borel measure μ on $\mathbb{R}^d$ such that

$$\forall |\alpha| \leq m : \beta_\alpha = \int_{\mathbb{R}^d} x^\alpha \mu(dx). \tag{3.1}$$

Now the TMMP asks whether there are representing measures for special multisequences. For a given closed subset $K \subset \mathbb{R}^d$, the K-moment problem furthermore asks, if there are representing measures supported in K. There are many open questions related to the TMMP, especially the following one: Given a multisequence $\beta^{(m)}$ having at least one representing measure supported in K, is there a finitely atomic representing measure supported in K, i. e. a representing measure $\mu = \sum_{i=1}^n a_i \delta_{x_i}$ where $x_i \in K, i = 1, \ldots, n$.

The connections between Theorem 3.1.3 and the TMMP are clear: Given a measure μ as in the theorem use (3.1) to define a multisequence β. Of course, μ is a representing measure supported in K. Tchakaloff's Theorem now guarantees the existence of a finitely atomic representing measure supported in K such that the cardinality of its support is less than or equal to $N_{m,d}$.

In the following, let $K = \operatorname{supp}(\mu)$, where μ is the measure in the theorem. Then K is compact. Define $\beta = \beta^{(m)}$ by (3.1). Consider the set

$$M_K(\beta) = \left\{ \nu \in M(\mathbb{R}^d) \middle| \nu \text{ is a representing measure for } \beta,\ \operatorname{supp}(\nu) \subset K \right\}, \tag{3.2}$$

where $M(\mathbb{R}^d)$ is the vector space of all finite signed Borel measures on $\mathbb{R}^d$.

In order to prove Theorem 3.1.3 we have to show that $M_K(\beta)$ contains at least one finitely atomic measure ν satisfying $|\operatorname{supp}(\nu)| \leq N_{m,d}$. In a first step we prove that any extreme point of the convex set $M_K(\beta)$ fullfills those requirements, in a second step we show that $M_K(\beta)$ indeed always has some extreme points.

If $\nu \in M_K(\beta)$ then, of course, any polynomial with total degree less than or equal to m is absolutely integrable with respect to ν. Let $\mathbb{R}_m[X_1, \ldots, X_d](\nu)$ denote $\mathbb{R}_m[X_1, \ldots, X_d]$ understood as a subset of $L^1(\nu)$ and define
$\mathcal{N}_{m,d;\nu} = \dim \mathbb{R}_m[X_1, \ldots, X_d](\nu)$.

Proposition 3.1.5. *If $M_K(\beta)$ has an extreme point ν, then ν is finitely atomic with*

$$|\operatorname{supp}(\nu)| \leq \mathcal{N}_{m,d;\nu} \leq N_{m,d;\mu} \leq N_{m,d}.$$

Proof. Denote $\mathbb{R}_m[X_1, \ldots, X_d](\nu)$ by $\mathcal{L}$. We claim that $\mathcal{L}$ is dense in $L^1(\nu)$ if ν is an extreme point of $M_K(\beta)$. Since ν is a finite measure, we have $L^1(\nu)^* = L^\infty(\nu)$, where X^* denotes the (algebraical and topological) dual space of a Banach space X. Thus we have to show that for all $f \in L^\infty(\nu)$, $f \neq 0$, there is a $P \in \mathcal{L}$ such that

$$\int_{\mathbb{R}^d} f(x)P(x)\nu(dx) \neq 0.$$

Assume the contrary and let $f \neq 0$ be an element of $L^\infty(\nu)$ that annihilates $\mathcal{L}$. By replacing f with $\frac{1}{2\|f\|_\infty} f$ if necessary, we may assume that $\|f\|_\infty \leq 1/2$. Therefore we can define new measures ν_1 and ν_2 by $\nu_1 = (1+f)\nu$ and $\nu_2 = (1-f)\nu$, i. e. ν_1 is the measure with density $1+f(x)$ with respect to ν. ν_i, $i=1,2$, belongs to $M_K(\beta)$ as for $P(x) = x^\alpha$ with $|\alpha| \leq m$ we have

$$\beta_\alpha = \int P(x)\nu(dx) = \int P(x)\nu(dx) \pm \int P(x)f(x)\nu(dx) = \int P(x)\nu_i(dx).$$

But now $\nu = \frac{1}{2}\nu_1 + \frac{1}{2}\nu_2$, which is a contradiction to the fact that ν is an extreme point. So $\mathcal{L}$ is a finite dimensional dense subset of $L^1(\nu)$, thus $\mathcal{L}$ is equal to $L^1(\nu)$. It follows that the dimensions of $\mathcal{L}$, $L^1(\nu)$ and $L^\infty(\nu)$ are all the same. Define $r = \mathcal{N}_{m,d;\nu}$. Suppose that $\text{supp}(\nu)$ contains at least $r+1$ distinct points $x_1, \dots, x_{r+1}$. For $i = 1, \dots, r+1$ let D_i denote mutually disjoint balls centered at x_i, respectively, with common positive radii. Then clearly the indicator functions $\mathbf{1}_{D_i}$ are linearly independent in $L^\infty(\nu)$, a contradiction to $r = \dim L^\infty(\nu)$. The other inequalities in the statement of the proposition are obvious.

Before being able to show that the set $M_K(\beta)$ always has an extreme point, we collect some facts from functional analysis that can be found in [16].

Let $(X, \|\cdot\|)$ be a normed vector space and let X^* be its dual space. For $F \subset X$ and $\varepsilon > 0$ define

$$U_{F,\varepsilon} = \{x^* \in X^* | \forall x \in F : |x^*(x)| \leq \varepsilon\}$$

and

$$\mathfrak{U} = \{U_{F,\varepsilon} | F \subset X, |F| < \infty, \varepsilon > 0\}.$$

By calling a subset $O \subset X^*$ open if and only if

$$\forall x^* \in O \exists U \in \mathfrak{U} : x^* + U \subset O,$$

a topology is defined on X^* that is called *weak*-topology*. X^* endowed with the weak*-topology is a locally convex topological vector space and a Hausdorff space.

Let Ω be a set with the σ-algebra $\mathcal{B}$. The vector space of all finite signed measures on $(\Omega, \mathcal{B})$ is denoted by $M_f(\Omega)$. For some $\nu \in M_f(\Omega)$, the non-negative real number given by

$$\|\nu\| = \sup_{\mathcal{Z}} \sum_{E \in \mathcal{Z}} |\nu(E)|,$$

where $\mathcal{Z}$ passes through all finite pairwise disjoint measurable partitions of Ω, is called the *variational norm* of ν. $(M_f(\Omega), \|\cdot\|)$ is a Banach space.

In the following, we need some of the most famous theorems of mathematics, which we cite without giving proofs.

Theorem 3.1.6 (Riesz Representation Theorem). *Let X be a compact topological space, let $C(X)$ be the Banach space of all continuous functions from X to $\mathbb{R}$ endowed with the supremum-norm. Furthermore, let $M_r(X)$ be the Banach space of finite signed regular Borel measures on X with the variational norm as defined above. Then $C(X)^*$ (with the norm topology) and $M_r(X)$ are isometrically isomorph via the isometry $T : M_r(X) \to C(X)^*$ given by*

$$(T\nu)(f) = \int_X f(x)\nu(dx).$$

Theorem 3.1.7 (Alaoglu's Theorem). *Let X be a normed vector space and let $B(X^*)$ be the unit ball of its dual space, i. e. $B(X^*) = \{x^* \in X^* | \|x^*\| \leq 1\}$. Then $B(X^*)$ is weak*-compact.*

Theorem 3.1.8 (Krein-Milman Theorem). *Let X be a locally convex Hausdorff space, $A \subset X$ compact, convex and nonempty. Then*

1. *the set* $\operatorname{ex} A$ *of extreme points of A is nonempty,*
2. *A is equal to the closed convex hull of* $\operatorname{ex} A$.

Proof of Theorem 3.1.3. The support K of the measure μ is a compact set in $\mathbb{R}^d$, therefore Theorem 3.1.6 implies that $C(K)^*$ is isometrically isomorph to $M_r(K) = M(K)$, where $M(K)$ once more denotes the space of all finite signed Borel measures. Thus we may regard $M_K(\beta) \subset M(K)$ as a subset of $C(K)^*$. Moreover, since μ is finite, we may assume that $\mu(K) = 1$. By definition of the variational norm, it follows that $\|\mu\| = 1$ and so $\mu \in B(C(K)^*)$. This is true for all other $\nu \in M_K(\beta)$, too (take $P(x) \equiv 1 \in \mathbb{R}_m[X_1, \ldots, X_d]$). So we have $M_K(\beta) \subset B(C(K)^*)$.

According to Theorem 3.1.7 $B(C(K)^*)$ is weak*-compact. Note that $B(C(K)^*)$ is weak*-metrizable, because $C(K)$ is separable, see [3], Theorem V.5.1.

Therefore, in order to prove weak*-closedness of $M_K(\beta)$, it suffices to show that the limit ν of any weak*-convergent series ν_n in $M_K(\beta)$ lies in $M_K(\beta)$. Of course, the restriction of any polynomial to K is continuous. $\nu_n \in M_K(\beta)$, i. e. $\forall |\alpha| \leq m$: $\int x^\alpha \nu_n(dx) = \beta_\alpha$. $\nu_n \to \nu$ means that for any continuous function f on K

$$\int f(x)\nu_n(dx) \longrightarrow \int f(x)\nu(dx)$$

as $n \to \infty$. For $f = x^\alpha$ follows

$$\beta_\alpha = \int_K x^\alpha \nu_n(dx) \longrightarrow \int_K x^\alpha \nu(dx).$$

Therefore $\nu \in M_K(\beta)$ and $M_K(\beta)$ indeed is weak*-closed. Since $C(K)^*$ with the weak*-topology is a Hausdorff space, this means that $M_K(\beta)$ is weak*-compact, too.

Because $\mu \in M_K(\beta)$, Theorem 3.1.8 guarantees the existence of an extreme point ν. According to Proposition 3.1.5 $\nu = \lambda_1 \delta_{x_1} + \cdots + \lambda_n \delta_{x_n}$ for some points $x_1, \ldots, x_n \in K$, some real numbers $\lambda_1, \ldots, \lambda_n$ and $n \leq N_{m;d}$. Since ν is a measure, the scalars are nonnegative and by reducing n if necessary, the λ_i even are positive. By linearity of the integral, a cubature formula is given by ν.

Remark 3.1.9. The proof of the generalized Tchakaloff theorem is more difficult. The problem is that one cannot use the Riesz Representation Theorem so easily. One has to work with continuous functions vanishing at infinity instead of merely continuous functions.

3.2 A Tchakaloff Theorem on Wiener Space

If we want to extend the notion of a cubature formula on stochastic integrals in the Stratonovich sense, we first have to identify special stochastic processes playing the role of polynomials. In the deterministic case cubature formulae can be used to approximate integrals of smooth function because smooth functions can be approximated by polynomials using the Taylor expansion. In Section 2.2, we saw that the solutions of SDE's can be approximated using Stratonovich iterated integrals. Therefore the definition below seems reasonable.

Definition 3.2.1. Let $m, n \in \mathbb{N} \setminus \{0\}$, $\omega_1, \dots, \omega_n \in C^0_{0,bv}([0,t], \mathbb{R}^d)$ and $\lambda_1, \dots, \lambda_n > 0$. $\omega_1, \dots, \omega_n$ and $\lambda_1, \dots, \lambda_n$ define a *cubature formula on Wiener space of degree m at time T* if and only if for all $(i_1, \dots, i_k) \in \mathcal{A}_m$

$$E\left(\iint\limits_{0<t_1<\cdots<t_k<T} \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k}\right) = \sum_{j=1}^{n} \lambda_j \iint\limits_{0<t_1<\cdots<t_k<T} d\omega_j^{i_1}(t_1) \cdots d\omega_j^{i_k}(t_k).$$

Remark 3.2.2. As mentioned before, when studiyng a solution of a SDE, we always choose a version that equals the solution of the corresponding deterministic differential equation when evaluated at a path with bounded variation. The iterated Stratonovich integrals can be interpreted as solutions of SDE's, therefore evaluating an iterated Stratonovich integral at a path of bounded variation ω actually makes sense and gives the respective iterated Riemann-Stieltjes integral of ω. Therefore the equation in Definition 3.2.1 can be rewritten as

$$E\left(\iint\limits_{0<t_1<\cdots<t_k<T} \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k}\right) = E_{\mathbb{Q}}\left(\iint\limits_{0<t_1<\cdots<t_k<T} \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k}\right),$$

where $E_{\mathbb{Q}}$ denotes expectation with respect to the probability measure $\mathbb{Q} = \sum_{j=1}^{n} \lambda_j \delta_{\omega_j}$.

We now state the analogon of Theorem 3.1.3 for cubature formulae on Wiener space.

Theorem 3.2.3. *For a natural number $m > 0$ there are n paths $\omega_1, \dots, \omega_n \in C^0_{0,bv}([0,T], \mathbb{R}^d)$ and n positive weights $\lambda_1, \dots, \lambda_n$ defining a cubature formula on Wiener space of degree m at time T for $n \leq |\mathcal{A}_m|$.*

We postpone the proof of Theorem 3.2.3 to the next chapter, see Lemma 4.7.1.

The following proposition ensures that it actually suffices to construct a cubature formula on Wiener space at time $T = 1$.

Proposition 3.2.4. *Assume $\omega_1, \dots, \omega_n \in C^0_{0,bv}([0,1], \mathbb{R}^d)$ and $\lambda_1, \dots \lambda_n > 0$ define a cubature formula on Wiener space of degree m at time 1. Then a cubature formula of degree m at time T is defined by the same weights $\lambda_1, \dots, \lambda_n$ and the paths $\omega_{T,1}, \dots, \omega_{T,n} \in C^0_{0,bv}([0,T], \mathbb{R}^d)$ with $\omega^j_{T,i}(t) = \sqrt{T}\omega^j_i(t/T)$ for $j = 1, \dots, d$, $i = 1, \dots, n$ and $0 \leq t \leq T$.*

Proof. By assumption we have

$$E\left(\iint\limits_{0<t_1<\cdots<t_k<1} \circ dB^{i_1}_{t_1} \cdots \circ dB^{i_k}_{t_k}\right) = \sum_{i=1}^{n} \lambda_i \iint\limits_{0<t_1<\cdots<t_k<1} d\omega_i^{i_1}(t_1) \cdots d\omega_i^{i_k}(t_k).$$

For $t \in [0,T]$, define $y = t/T \in [0,1]$ then we formally have $d\omega^j_{T,i}(t) = d\omega^j_{T,i}(yT) = \sqrt{T} d\omega^j_i(y)$ for $j = 1,\dots,d$ and $d\omega^0_{T,i}(t) = dt = d(yT) = T dy$. Using properties of Riemann-Stieltjes integrals we get

$$\sum_{i=1}^n \lambda_i \iint_{0<t_1<\cdots<t_k<T} d\omega^{i_1}_{T,i}(t_1)\cdots d\omega^{i_k}_{T,i}(t_k) = \sum_{i=1}^n \lambda_i \int_0^T \int_0^{t_k} \cdots \int_0^{t_2} d\omega^{i_1}_{T,i}(t_1)\cdots d\omega^{i_k}_{T,i}(t_k) =$$

$$= \sum_{i=1}^n \lambda_i \int_0^1 \int_0^{t_k/T} \cdots \int_0^{t_2/T} \sqrt{T}^{\operatorname{ord}(i_1,\dots,i_k)} d\omega^{i_1}_i(y_1)\cdots d\omega^{i_k}_i(y_k) =$$

$$= \sqrt{T}^{\operatorname{ord}(i_1,\dots,i_k)} \sum_{i=1}^n \lambda_i \iint_{0<t_1<\cdots<t_k<1} d\omega^{i_1}_i(t_1)\cdots d\omega^{i_k}_i(t_k),$$

where we substitute $y_l = t_l/T$ for $l = 1,\dots,k$. But by assumption, the last expression is equal to

$$\sqrt{T}^{\operatorname{ord}(i_1,\dots,i_k)} E\left(\iint_{0<t_1<\cdots<t_k<1} \circ dB^{i_1}_{t_1}\cdots\circ dB^{i_k}_{t_k}\right)$$

and by the scaling property (2.12), this is the same as

$$E\left(\iint_{0<t_1<\cdots<t_k<T} \circ dB^{i_1}_{t_1}\cdots\circ dB^{i_k}_{t_k}\right),$$

which concludes the proof.

3.3 Algorithmic Implementation

Assume that the paths $\omega_1,\dots,\omega_n$ and the weights $\lambda_1,\dots,\lambda_n$ define a cubature formula on Wiener space of degree m at time 1. Then by Proposition 3.2.4, using the notation therein, the paths $\omega_{T,1},\dots,\omega_{T,n}$ together with the weights $\lambda_1,\dots,\lambda_n$ define a cubature formula at time T. Following Remark 3.2.2, we can consider the probability measure

$$\mathbb{Q}_T = \sum_{i=1}^n \lambda_i \delta_{\omega_{T,i}}$$

on Wiener space and get for all multi-indices $(i_1,\dots,i_k) \in \mathcal{A}_m$:

$$E\left(\iint_{0<t_1<\cdots<t_k<T} \circ dB^{i_1}_{t_1}\cdots\circ dB^{i_k}_{t_k}\right) = E_{\mathbb{Q}_T}\left(\iint_{0<t_1<\cdots<t_k<T} \circ dB^{i_1}_{t_1}\cdots\circ dB^{i_k}_{t_k}\right),$$

where E, as usual, denotes expectation with respect to P and $E_{\mathbb{Q}_T}$ denotes expectation with respect to $\mathbb{Q}_T$.

As a first result, we derive an estimate for the remainder term in the stochastic Taylor expansion (2.7) in the new probability $\mathbb{Q}_T$.

Lemma 3.3.1. *Let $R_m(T,x,f)$ be the process defined in Proposition 2.2.2. Then we have*

$$\sup_{x\in\mathbb{R}^N} E_{\mathbb{Q}_T}\left(|R_m(t,x,f)|\right) \le C_{d,m,\mathbb{Q}_1} T^{\frac{m+1+\mathbf{1}_{T>1}}{2}} \sup_{(i_1,\dots,i_k)\in\mathcal{A}_{m+2}\setminus\mathcal{A}_m} \|V_{i_1}\cdots V_{i_k} f\|_\infty, \quad (3.3)$$

where $C_{m,d,\mathbb{Q}_1}$ is a constant depending on $\mathbb{Q}_1$, n and m.

Proof. In the proof of Proposition 2.2.2, we showed that

$$R_m(t,x,f) = \sum_{\substack{(i_1,\dots,i_k)\in\mathcal{A}_m \\ (i_0,\dots,i_k)\notin\mathcal{A}_m}} \iint\limits_{0<t_0<\cdots<t_k<T} V_{i_0}\cdots V_{i_k} f(\xi_{t_0,x}) \circ dB^{i_0}_{t_0}\cdots \circ dB^{i_k}_{t_k}.$$

Using the triangle inequality we see that $E_{\mathbb{Q}_T}(|R_m(t,x,f)|)$ is bounded by

$$\sum_{j=1}^n \lambda_j \sum_{\substack{(i_1,\dots,i_k)\in\mathcal{A}_m \\ (i_0,\dots,i_k)\notin\mathcal{A}_m}} \underbrace{\left| \iint\limits_{0<t_0<\cdots<t_k<T} V_{i_0}\cdots V_{i_k} f(\xi_{t_0,x}(\omega_{T,j}))d\omega^{i_0}_{T,j}(t_0)\cdots d\omega^{i_k}_{T,j}(t_k)\right|}_{=A_{j;i_0,\dots,i_k}(T,x,f)}.$$

Now we change variables just as in the proof of Proposition 3.2.4 and get

$$A_{j;i_0,\dots,i_k} = \sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \left| \int_0^1\int_0^{t_k}\cdots\int_0^{t_1} h(t_0)d\omega^{i_0}_j(t_0)\cdots d\omega^{i_{k-1}}_j(t_{k-1})d\omega^{i_k}_j(t_k)\right|,$$

where $h(t_0)$ denotes $V_{i_0}\cdots V_{i_k} f(\xi_{t_0,x}(\omega_{T,j}))$. In the next step we use the following inequality for Riemann-Stieltjes integrals:

$$\left|\int_0^t f(s)dg(s)\right| \le \|f\|_{\infty,[0,t]}\|g\|_{bv,[0,t]}, \quad (3.4)$$

where $\|f\|_{\infty,[0,t]}$ is the supremum-norm of the continuous function f on $[0,t]$ and $\|g\|_{bv,[0,t]}$ is the variation norm of the bounded variation function g on $[0,t]$. Applying (3.4) successively on the integrals we get

$$\begin{aligned}
A_{j;i_0,\dots,i_k} &\le \sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \|\omega^{i_k}_j\|_{bv,[0,1]}\cdot \\
&\cdot \left\| \int_0^{t_k}\cdots\int_0^{t_1} h(t_0)d\omega^{i_0}_j(t_0)\cdots d\omega^{i_{k-1}}_j(t_{k-1})\right\|_{\infty,[0,1]} = \\
&= \sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \|\omega^{i_k}_j\|_{bv,[0,1]} \sup_{0\le t_k\le 1} \left|\cdots\right| \le \sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \|\omega^{i_k}_j\|_{bv,[0,1]}\cdot \\
&\cdot \sup_{0\le t_k\le 1}\left(\|\omega^{i_{k-1}}_j\|_{bv,[0,t_k]} \left\| \int_0^{t_{k-1}}\cdots\int_0^{t_1} h(t_0)d\omega^{i_0}_j\cdots d\omega^{i_{k-2}}_j(t_{k-2})\right\|_{\infty,[0,t_k]}\right) = \\
&= \sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \|\omega^{i_k}_j\|_{bv,[0,1]} \|\omega^{i_{k-1}}_j\|_{bv,[0,1]} \left\| \int_0^{t_{k-1}} (\cdots)d\omega^{i_{k-2}}_j(t_{k-2})\right\|_{\infty,[0,1]}.
\end{aligned}$$

Finally by induction we get

$$\begin{aligned} A_{j;i_0,\dots,i_k} &\leq \sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \|\omega_j^{i_k}\|_{bv,[0,1]} \cdots \\ &\quad \cdots \|\omega_j^{i_0}\|_{bv,[0,1]} \|V_{i_0} \cdots V_{i_k} f(\xi_{t_0,x}(\omega_{T,j}))\|_{\infty,[0,1]} \\ &\leq \sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \|\omega_j^{i_k}\|_{bv,[0,1]} \cdots \|\omega_j^{i_0}\|_{bv,[0,1]} \|V_{i_0} \cdots V_{i_k} f\|_\infty. \end{aligned}$$

Here

$$\|V_{i_0} \cdots V_{i_k} f(\xi_{t_0,x}(\omega_{T,j}))\|_{\infty,[0,1]} = \sup_{t_0 \in [0,1]} |V_{i_0} \cdots V_{i_k} f(\xi_{t_0,x}(\omega_{T,j}))|,$$

whereas $\|V_{i_0} \cdots V_{i_k} f\|_\infty$ means $\sup_{x \in \mathbb{R}^N} |V_{i_0} \cdots V_{i_k} f(x)|$.

Now $\operatorname{ord}(i_0, \dots, i_k)$ is either $m+1$ or $m+2$, therefore for $T > 1$ we have $\sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \leq T^{\frac{m+2}{2}}$, whereas for $0 < T \leq 1$ $\sqrt{T}^{\operatorname{ord}(i_0,\dots,i_k)} \leq T^{\frac{m+1}{2}}$. So we get the desired result by defining

$$C_{d,m,\mathbb{Q}_1} = \sum_{j=1}^n \lambda_j \sum_{\substack{(i_1,\dots,i_k) \in \mathcal{A}_m \\ (i_0,\dots,i_k) \notin \mathcal{A}_m}} \|\omega_j^{i_0}\|_{bv,[0,1]} \cdots \|\omega_j^{i_k}\|_{bv,[0,1]}.$$

In order to get an approximation for the solution of a parabolic PDE, we have to approximate $E(f(\xi_{T,x}))$ for some smooth (i. e. C^∞) or Lipschitz function f. Assume f is smooth.

For some $\omega \in C^0_{0,bv}([0,T], \mathbb{R}^d)$ the solution $\Phi_{T,x}(\omega) = \xi_{T,x}(\omega)$ of the SDE (2.1) at ω is given by the solution of the ordinary differential equation

$$dy_{t,x} = \sum_{i=0}^d V_i(y_{t,x}) d\omega^i(t) \tag{3.5}$$

with initial condition $y_{0,x} = x$, $x \in \mathbb{R}^N$.

Proposition 3.3.2. *If the paths $\omega_1, \dots, \omega_n$ and the weights $\lambda_1, \dots, \lambda_n$ define a cubature formula on Wiener space of degree m at time 1, then*

$$\sup_{x \in \mathbb{R}^N} \left| E(f(\xi_{T,x})) - \sum_{i=1}^n \lambda_i f(\Phi_{T,x}(\omega_{T,i})) \right| \leq C \sqrt{T}^{m+1+\mathbf{1}_{T>1}} \sup_{(i_1,\dots,i_k) \in \mathcal{A}_{m+2} \setminus \mathcal{A}_m} \|V_{i_1} \cdots V_{i_k} f\|_\infty, \tag{3.6}$$

where C is a constant independent of T.

Proof. With the probability $\mathbb{Q}_T$ defined as above, we get $\sum_{i=1}^n \lambda_i f(\Phi_{T,x}(\omega_{T,i})) = E_{\mathbb{Q}_T}(f(\xi_{T,x}))$. Using the notation $f(\xi_{T,x}) = T_m(T,x,f) + R_m(T,x,f)$ as in the proof of Proposition 2.2.2, we have

$$(E - E_{\mathbb{Q}_T})(T_m(T,x,f)) = \\ = \sum_{(i_1,\dots,i_k) \in \mathcal{A}_m} V_{i_1} \cdots V_{i_k} f(x) (E - E_{\mathbb{Q}_T}) \left(\iint_{0<t_1<\dots<t_k<T} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k} \right) = 0.$$

Therefore

$$|(E - E_{\mathbb{Q}_T})(f(\xi_{T,x}))| \leq (E + E_{\mathbb{Q}_T})(|R_m(T,x,f)|) \leq \\ \leq \sqrt{E(R_m(T,x,f)^2)} + E_{\mathbb{Q}_T}(|R_m(T,x,f)|)$$

and the desired result follows by using the boundaries in Proposition 2.2.2 and in Lemma 3.3.1.

In order to approximate the expectation of $f(\xi_{T,x})$ with respect to P by the expectation with respect to $\mathbb{Q}_T$, the upper boundary of the difference given in Proposition 3.3.2 should be small, which is usually not the case unless T or $\sup_{(i_1,\dots,i_k)\in\mathcal{A}_{m+2}\setminus\mathcal{A}_m} \|V_{i_1}\cdots V_{i_k} f\|_\infty$ is small. Therefore we will divide the intervall $[0,T]$ into many small subintervals. Then we consecutively apply the approximation to these intervals in the following way: Let $0 = t_0 < t_1 < \cdots < t_k = T$ and define $s_l = t_l - t_{l-1}$ for $l = 1,\dots,k$. Now consider the random variable $(Y_l)_{0\leq l\leq k}$ given by

$$P(Y_{l+1} = \Phi_{s_{l+1},x}(\omega_{s_{l+1},i})|Y_l = x) = \lambda_i, \tag{3.7}$$

for $l = 0,\dots,k-1$ and $i = 1,\dots,n$. So we get Y_{l+1} by following (with probability λ_i) the solution of the ODE (3.5) driven by $\omega_{s_{l+1},i}$ starting at Y_l until time s_{l+1}. Of course, $(Y_l)_{0\leq l\leq k}$ satisfies the Markov property, i. e. the distribution of Y_{l+1} given Y_l is independent of $Y_0,\dots,Y_{l-1}$.

Theorem 3.3.3. *With the random variable $(Y_l)_{0\leq l\leq k}$ as defined above and with $P_t f(x) = E(f(\xi_{t,x}))$, we have*

$$\sup_{x\in\mathbb{R}^N} |E(f(Y_k)|Y_0 = x) - E(f(\xi_{T,x}))| \leq \\ \leq C\sum_{j=1}^{k} s_j^{\frac{m+1+\mathbf{1}_{s_j>1}}{2}} \sup_{(i_1,\dots,i_k)\in\mathcal{A}_{m+2}\setminus\mathcal{A}_m} \|V_{i_1}\cdots V_{i_k} P_{T-t_j} f\|_\infty \tag{3.8}$$

for some constant C.

Proof. Because of $P_0(Y_k) = f(Y_k)$ we get

$$P_T f(x) - E(f(Y_k)|Y_0 = x) = \sum_{j=0}^{k-1} E(P_{T-t_j} f(Y_j) - P_{T-t_{j+1}} f(Y_{j+1})|Y_0 = x).$$

Recall the *weak Markov property* of the solution of a SDE: For $t,h > 0$ we have

$$E(f(\xi_{t+h,x})|\mathcal{F}_t) = E(f(\xi_{h,y}))|_{y=\xi_{t,x}}.$$

Thus we get

$$P_{s_{j+1}} P_{T-t_{j+1}} f(x) = E\left(E(f(\xi_{T-t_{j+1},y}))\Big|_{y=\xi_{s_{j+1},x}} \right) = \\ = E\left(E(f(\xi_{T-t_j,x})|\mathcal{F}_{T-t_{j+1}})\right) = E\left(f(\xi_{T-t_j,x})\right) = P_{T-t_j} f(x).$$

Thus, for $0 \le j \le k-1$ we get

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^N} \left|E\left(P_{T-t_j} f(Y_j) P_{T-t_{j+1}} f(Y_{j+1}) | Y_0 = x\right)\right| = \\
&= \sup_{x\in\mathbb{R}^N} \left|E\left(P_{s_{j+1}} P_{T-t_{j+1}} f(Y_j) - P_{T-t_{j+1}} f(Y_{j+1}) | Y_0 = x\right)\right| \le \\
&\le \sup_{x\in\mathbb{R}^N} E\left(\left|E\left(P_{s_{j+1}} P_{T-t_{j+1}} f(Y_j) - P_{T-t_{j+1}} f(Y_{j+1}) | Y_j\right)\right| \big| Y_0 = x\right) \le \\
&\le \sup_{x\in\mathbb{R}^N} \left|E\left(P_{s_{j+1}} P_{T-t_{j+1}} f(Y_j) - P_{T-t_{j+1}} f(Y_{j+1}) | Y_j = x\right)\right| = \\
&= \sup_{x\in\mathbb{R}^N} \left|E\left(P_{s_{j+1}} (P_{T-t_{j+1}} f)(Y_j) | Y_j = x\right) - E\left(P_{T-t_{j+1}} f(Y_{j+1}) | Y_j = x\right)\right| = \\
&= \sup_{x\in\mathbb{R}^N} \left|E\left((P_{T-t_{j+1}} f)(\xi_{s_{j+1},x})\right) - \sum_{i=1}^n \lambda_i (P_{T-t_{j+1}} f)(\Phi_{s_{j+1}}(\omega_{s_{j+1},i}))\right| \le \\
&\le C s_{j+1}^{\frac{m+1+\mathbf{1}_{s_{j+1}>1}}{2}} \sup_{(i_1,\dots,i_k)\in\mathcal{A}_{m+2}\setminus\mathcal{A}_m} \|V_{i_1}\cdots V_{i_k} P_{T-t_{j+1}} f\|_\infty,
\end{aligned}$$

where we used Proposition 3.3.2 with $P_{T-t_{j+1}} f$ instead of f in the last line. The theorem follows by summing up for $j = 0, \dots, k-1$.

Theorem 3.3.3 enables us to approximate $E(f(\xi_{T,x}))$ by using the random variable Y_k. For $t_l \le t \le t_{l+1}, 0 \le l \le k-1$, let

$$\omega_{s_1,i_1} \otimes \cdots \otimes \omega_{s_k,i_k}(t) = \omega_{s_1,i_1}(s_1) + \cdots + \omega_{s_l,i_l}(s_l) + \omega_{s_{l+1},i_{l+1}}(t - t_l),$$

then $\omega_{s_1,i_1} \otimes \cdots \otimes \omega_{s_k,i_k} \in C^0_{0,bv}([0,T],\mathbb{R}^d)$ is the path one gets by consecutively traversing the paths $\omega_{s_1,i_1}, \dots, \omega_{s_k,i_k}$, where ω_{s_2,i_2} is understood to be "attached" to the endpoint of ω_{s_1,i_1} and so on. Then, given $Y_0 = x$, the random variable Y_k takes the value $\Phi_{T,x}(\omega_{s_1,i_1} \otimes \cdots \otimes \omega_{s_k,i_k})$ with probability $\lambda_{i_1} \cdots \lambda_{i_k}$ by the semi-group property of solutions of ODEs.

Therefore, $E(f(\xi_{T,x}))$ can be approximated by

$$E(f(Y_k)|Y_0 = x) = E_{\mathbb{Q}_T^k}(f(\xi_{T,x})),$$

where

$$\mathbb{Q}_T^k = \sum_{i_1=1}^n \cdots \sum_{i_k=1}^n \lambda_{i_1} \cdots \lambda_{i_k} \delta_{\omega_{s_1,i_1} \otimes \cdots \otimes \omega_{s_k,i_k}}.$$

An implementation of this method would look like this: We start at time 0 at the starting value x with probability 1. In the first step compute the n solutions at time s_1 of the ODEs (3.5) corresponding to the paths $\omega_{s_1,1}, \dots, \omega_{s_1,n}$ with initial value x, e. g. by some numerical algorithm. Denote these solutions by $\Phi_{s_1,x}(\omega_{s_1,1}), \dots, \Phi_{s_n,x}(\omega_{s_1,n})$. Then for each $1 \le j \le n$ compute the solutions of the ODEs corresponding to $\omega_{s_2,1}, \dots, \omega_{s_2,n}$ at time s_2 with initial value $\Phi_{s_1,x}(\omega_{s_1,j})$. Iterate n times, then compute the weighted sum as indicated above.

If we we can compute the exact solutions of all these ODEs, Theorem 3.3.3 guarantees that the total error is less than $D \sum_{j=1}^k s_j^{(m+1)/2}$ for some constant D and for $s_j < 1$. For $m > 1$ we have

$$\sum_{j=1}^k s_j^{(m+1)/2} \le T \max_{j=1,\dots,k} s_j^{(m-1)/2},$$

so by increasing the number of steps we can make the error as small as desired. Of course, if we numerically solve the ODE's, the error made by the ODE solver might get bigger if the number of steps is increased.

Chapter 4

Tensor Algebra

In the following, we adopt the Lie algebraic interpretation of iterated integrals, as e. g. laid out in [10]. First we will need the notion of a tensor algebra.

4.1 Tensor Products

Before actually being able to define the tensor algebra that we are going to use, we will repeat some basic facts about the tensor product of two finite-dimensional (real) vector spaces V and W.

First, let S be a set. Consider the set of all maps $F : S \to \mathbb{R}$ such that $F(s) = 0$ for all but finitely many $s \in S$. Under pointwise addition and scalar multiplication, this set is a real vector space, called the *free vector space on* S and denoted by $\mathbb{R}\langle S\rangle$. Obviously, S can be embedded into $\mathbb{R}\langle S\rangle$ by identifying $s \in S$ with the indicator function $\mathbf{1}_{\{s\}} : S \to \{0, 1\}$. Then S is a basis of its free vector space.

Now, let $\mathcal{R}$ be the linear hull of all the elements of $\mathbb{R}\langle V \times W\rangle$ which have the form

$$\begin{aligned} a(v, w) &- (av, w),\\ a(v, w) &- (v, aw),\\ (v, w) + (v', w) &- (v + v', w),\\ (v, w) + (v, w') &- (v, w + w'), \end{aligned}$$

for $a \in \mathbb{R}$, $v, v' \in V$ and $w, w' \in W$. We now define the *tensor product* of V and W to be the quotient space $\mathbb{R}\langle V \times W\rangle/\mathcal{R}$ and denote it by $V \otimes W$. $v \otimes w$ with $v \in V$ and $w \in W$ denotes the equivalence class of the element $(v, w) \in V \times W$ and is called the tensor product of v and w. From the definition of $\mathcal{R}$ follows bilinearity of the tensor product $\otimes$. However, not all elements of $V \otimes W$ are of the form $v \otimes w$, but all of them can be written as a linear combination of elements of the form $v \otimes w$.

One of the most important properties of the tensor product is the so called characteristic property.

Proposition 4.1.1 (Characteristic Property of Tensor Products). *If $A : V \times W \to X$ is a bilinear map into any vector space X, there is a unique linear map*

$\tilde{A} : V \otimes W \to X$, such that the following diagram commutes:

$$\begin{array}{ccc} V \times W & \xrightarrow{A} & X \\ \pi \downarrow & \nearrow_{\tilde{A}} & \\ V \otimes W & & \end{array}$$

where $\pi(v, w) := v \otimes w$.

Proof. First of all, A extends to a uniquely determined linear map A' from $\mathbb{R}\langle V \times W \rangle$ to X, like any other map from $V \times W$ to X, since $V \times W$ is a basis of the free vector space. The bilinearity of A implies that $\mathcal{R}$ lies in the kernel of the extended map. Therefore, the extension of A to the free vector space descends to a linear map $\tilde{A} : V \otimes W \to X$. For $v \in V$, $w \in W$:

$$\tilde{A}(v \otimes w) = \tilde{A}((v, w) + \mathcal{R}) = A'(v, w) = A(v, w).$$

So the diagram is commutative and $\tilde{A}$ is uniquely determined for all $v \otimes w$. Since $V \otimes W$ is spanned by these elements, the proposition holds.

Note that for any three finite-dimensional real vector spaces V, W and X, there is a uniquely determined isomorphism $(V \otimes W) \otimes X \to V \otimes (W \otimes X)$, mapping $(v \otimes w) \otimes x$ to $v \otimes (w \otimes x)$, see e. g. [8] for details. Therefore, we may define the tensor product of more than two vector spaces in the obvious way without bothering about brackets.

4.2 Iterated Integrals

Why might the notion of tensor products be useful in the context of iterated integrals? Suppose $x_t = \sum_{i=1}^{d} x_t^i e_i$ is a path in a real vector space V with basis $(e_1, \ldots, e_d)$. Then the second iterated integral is a matrix $X^2(0, t)$ defined in the following way (with respect to the basis induced by $(e_1, \ldots, e_d)$):

$$(X^2(0, t))^{ij} := \iint\limits_{0<u_1<u_2<t} dx_{u_1}^i dx_{u_2}^j. \tag{4.1}$$

Of course, this matrix could also be interpreted as the coordinate representation of a tensor with respect to the basis $(e_i \otimes e_j, i, j = 1, \ldots, d)$ of $V \otimes V$, i. e. $X^2(0, t) = \sum_{i,j=1}^{d} (X^2(0, t))^{ij} e_i \otimes e_j$. We will write

$$X^2(0, t) = \iint\limits_{0<u_1<u_2<t} dx_{u_1} \otimes dx_{u_2}, \tag{4.2}$$

if we want to emphasize the second interpretaion of the iterated integral. Analoguously, we can think of the k-th iterated integral as an element of $V^{\otimes k}$, where $V^{\otimes k}$ is inductively defined by

$$V^{\otimes 0} = \mathbb{R},$$

$$V^{\otimes 1} = V, \tag{4.3}$$

$$V^{\otimes j+1} = V^{\otimes j} \otimes V, \quad j \geq 1. \tag{4.4}$$

In this case we will write

$$X^k(0,t) = \iint\limits_{0<u_1<\cdots<u_k<t} dx_{u_1} \otimes \cdots \otimes dx_{u_k}. \tag{4.5}$$

4.3 Tensor Algebra

As indicated in Section 4.2 above, one might use the *tensor algebra* over V defined as

$$T = \bigoplus_{k=0}^{\infty} V^{\otimes k} \tag{4.6}$$

in order to study the iterated integrals of paths in V. However, this approach would not be appropriate in our situation. We want to study paths $\omega \in C_0^0([0,T],\mathbb{R}^d)$ with the additional component $\omega^0(t) = t$. So our paths are in fact $\mathbb{R} \oplus \mathbb{R}^d$-valued. But the 0-th projection does play a different role than the other ones. If we chose $V = \mathbb{R}^{d+1} \simeq \mathbb{R} \oplus \mathbb{R}^d$ and studied T as defined in (4.6), that difference would be neglected.

Moreover, as we have seen in Section 2 before, ω^0 somehow has another order of magnitude than the other projections: The index 0 counts twice as much as the other indices in the definition of $\mathcal{A}_m$. Ideally, this property should be reflected by the definition of the appropriate tensor algebra. As a consequence, now the order of the indices matters: A sequence of indices containing 0 must be treated differently than a sequence without 0.

Let S_k denote the group of permutations of $\{1,\dots,k\}$. Then for vector spaces $W_1,\dots,W_k$ define the *symmetrized product*

$$(W_1,\dots,W_k) = \bigoplus_{V \in M} V, \tag{4.7}$$

where $M = \left\{V \middle| \exists \sigma \in S_k : V = W_{\sigma(1)} \otimes \cdots \otimes W_{\sigma(k)}\right\}$. The symmetrized product is the direct sum of all tensor products of $W_1,\dots,W_k$ in any order, but such that the same term does not appear twice in the direct sum. As a shorthand notation we will use

$$(W_1,W_2)^{p,q} = (\underbrace{W_1,\dots,W_1}_{p \text{ times}},\underbrace{W_2,\dots,W_2}_{q \text{ times}}), \tag{4.8}$$

for natural numbers p and q.

Define

$$U_k(\mathbb{R},\mathbb{R}^d) = \bigoplus_{\substack{(i,j)\in\mathbb{N} \\ 2i+j=k}} (\mathbb{R},\mathbb{R}^d)^{i,j}. \tag{4.9}$$

U_k now plays the role of $V^{\otimes k}$, but respects the special role of ω^0 as mentioned above. We seperately define $U_0(\mathbb{R},\mathbb{R}^d) = \mathbb{R}$.

Definition 4.3.1. The *tensor algebra* is defined as

$$T(\mathbb{R},\mathbb{R}^d) = \bigoplus_{k=0}^{\infty} U_k(\mathbb{R},\mathbb{R}^d).$$

For $a \in T(\mathbb{R}, \mathbb{R}^d)$ we will write $a = (a_0, a_1, a_2, \ldots)$, with $a_i \in U_i(\mathbb{R}, \mathbb{R}^d)$ for $i \in \mathbb{N}$. Alternatively we will also write $a = a_0 + a_1 + a_2 + \cdots$, where the projection of a_i to $U_j(\mathbb{R}, \mathbb{R}^d)$ equals 0 for all j but i.

Obviously, $T(\mathbb{R}, \mathbb{R}^d)$ is a vector space with the operations

$$\begin{aligned} a + b &= (a_0 + b_0, a_1 + b_1, \ldots) \\ \lambda a &= (\lambda a_0, \lambda a_1, \ldots), \end{aligned}$$

when $a,\ b \in T(\mathbb{R}, \mathbb{R}^d)$ and $\lambda \in \mathbb{R}$. Moreover, $T(\mathbb{R}, \mathbb{R}^d)$ is an associative algebra with

$$(a \otimes b)_i = \sum_{j=0}^{i} a_j \otimes b_{i-j}.$$

The 1-element of that algebra is the element $(1, 0, 0, \ldots)$. If $a \in U_k(\mathbb{R}, \mathbb{R}^d)$, i. e. $a = a_k$ in the notation previously defined, and $b \in U_l(\mathbb{R}, \mathbb{R}^d)$, then $a \otimes b \in U_{k+l}(\mathbb{R}, \mathbb{R}^d)$, because $(a \otimes b)_i = \sum_{j=0}^{i} a_j \otimes b_{i-j} \neq 0$ if and only if $i = k + l$ and $j = k$. An algebra $\mathcal{A}$ is called *graded*, if it allows a direct sum decomposition $\mathcal{A} = \bigoplus_k A_k$, such that $A_k A_l \subset A_{k+l}$. Thus $T(\mathbb{R}, \mathbb{R}^d)$ is a graded algebra.

For $i \in \{0, \ldots, d\}$ let ε_i be the vector $(\delta_i^0, \ldots, \delta_i^d) \in \mathbb{R} \oplus \mathbb{R}^d$, where δ_i^j denotes the Kronecker-Delta. Then $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_d$ form a basis of $\mathbb{R} \oplus \mathbb{R}^d$ and the following lemma holds.

Lemma 4.3.2. *The set* $\{\varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_k} \mid (i_1, \ldots, i_k) \in \mathcal{A}_n \setminus \mathcal{A}_{n-1}\}$ *is a basis of* $U_n(\mathbb{R}, \mathbb{R}^d)$.

Proof. Denote the set defined in the lemma with $\mathcal{E}$. Because of (4.9), we have to show that any space $(\mathbb{R}, \mathbb{R}^d)^{i,j}$ with $2i + j = n$ is spanned by some elements of $\mathcal{E}$. Since $\mathcal{E}$ is invariant under permutation of the indices $(i_1, \ldots, i_k)$, it suffices to show that $\mathbb{R} \otimes \cdots \otimes \mathbb{R} \otimes \mathbb{R}^d \otimes \cdots \otimes \mathbb{R}^d$, where one has i times the term $\mathbb{R}$ and j times the term $\mathbb{R}^d$, is spanned by some elements of $\mathcal{E}$. Obviously, if V and W are vector spaces with basis $(e_1, \ldots, e_k)$ and $(f_1, \ldots, f_m)$ respectively, then $(e_i \otimes f_j : 1 \leq i \leq k, 1 \leq j \leq m)$ is a basis of $V \otimes W$. Therefore $\mathbb{R} \otimes \cdots \otimes \mathbb{R} \otimes \mathbb{R}^d \otimes \cdots \otimes \mathbb{R}^d$ is spanned by the tensors

$$\underbrace{\varepsilon_0 \otimes \cdots \otimes \varepsilon_0}_{i \text{ times}} \otimes \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_j}, \quad \{i_1, \ldots, i_j\} \in \{1, \ldots, d\}^d. \tag{4.10}$$

But the multi-index $(0, \ldots, 0, i_1, \ldots, i_j)$ is in $\mathcal{A}_n \setminus \mathcal{A}_{n-1}$, because $2i + j = n$. Since all the elements of $\mathcal{E}$ are linearly independent, that concludes the proof.

Many tensors even are invertible. Let $a \in T(\mathbb{R}, \mathbb{R}^d)$ be a tensor with $a_0 \neq 0$. Define a tensor c implicitly by $a = a_0(1 + c)$. Then c equals $(0, c_1, c_2, \cdots)$. Now let $c_{(i)} = (0, c_i, (c^2)_i, (c^3)_i, \ldots)$. Since $c_0 = 0$, for all i $c_{(i)}$ has only finitely many components different from 0. Thus we may define (without bothering about topology)

$$a^{-1} = \frac{1}{a_0} \sum_{k \geq 0} (-1)^k c^k, \tag{4.11}$$

since the sum in (4.11) really is a finite one in each component. Then $a \otimes a^{-1} = (1 + c) \otimes \sum_{k \geq 0} (-1)^k c^k = \sum_{k \geq 0} (-1)^k (1 + c) \otimes c^k$. Since this sum actually is a finite sum in each degree, we may conclude $a \otimes a^{-1} = (1+c) + (-c-c^2) + (c^2+c^3) + \cdots = 1 + (c-c) + (-c^2+c^2) + \cdots = 1$. For $a_0 > 0$ we define the logarithm by

$$\log(a) = \log(a_0) + \sum_{k\geq 1} (-1)^{k-1} \frac{c^k}{k}. \tag{4.12}$$

Once more, there are only finite sums in each component in (4.12).

The exponential can be defined for all tensors a, but it involves an actually infinite sum. Therefore, a topology is required, but that problem is not quite critical. We will discuss that point later. The definition is

$$\exp(a) = \sum_{k\geq 0} \frac{a^k}{k!}. \tag{4.13}$$

The space $\bigoplus_{k=n+1}^{\infty} U_k(\mathbb{R}, \mathbb{R}^d)$ is a subspace of $T(\mathbb{R}, \mathbb{R}^d)$ that is closed under multiplication. Moreover, if $a \in T(\mathbb{R}, \mathbb{R}^d)$ and $b \in \bigoplus_{k=n+1}^{\infty} U_k(\mathbb{R}, \mathbb{R}^d)$, then $a \otimes b$ and $b \otimes a$ are in the subspace as well. So it is an ideal and we can define the truncated tensor algebra of degree n by

$$T^{(n)}(\mathbb{R}, \mathbb{R}^d) = T(\mathbb{R}, \mathbb{R}^d) / \bigoplus_{k=n+1}^{\infty} U_k(\mathbb{R}, \mathbb{R}^d). \tag{4.14}$$

$T^{(n)}(\mathbb{R}, \mathbb{R}^d)$ can be identified with the space $\bigoplus_{k=0}^{n} U_k(\mathbb{R}, \mathbb{R}^d)$, and so we define projections $\pi_n : T(\mathbb{R}, \mathbb{R}^d) \to T^{(n)}(\mathbb{R}, \mathbb{R}^d)$ by setting $\pi_n((a_0, a_1, \ldots)) = (a_0, \ldots, a_n)$. From the definition of the operations follows that π_n is an algebra homomorphism. Now the space $T^{(n)}(\mathbb{R}, \mathbb{R}^d)$ is a finite-dimensional vector space. So we could choose any norm topology on it. Then, if we chose any topology on the tensor algebra, such that the projection π_n is continuous (and such that exp is defined), π_n would commute with the inverse, the logarithm and also with the exponential. So e. g. choose the initial topology of the projections $(\pi_n : n \in \mathbb{N})$.

The product

$$[a, b] = a \otimes b - b \otimes a \tag{4.15}$$

defines a Lie bracket both on $T(\mathbb{R}, \mathbb{R}^d)$ and on $T^{(n)}(\mathbb{R}, \mathbb{R}^d)$, and now π_n also is a Lie algebra homomorphism. The Lie algebra generated by $W = \mathbb{R} \oplus \mathbb{R}^d$ is denoted by $\mathcal{U}$ and looks like

$$\mathcal{U} = W \oplus [W, W] \oplus [W, [W, W]] \oplus \cdots,$$

where e. g. $[W, [W, W]]$ is the subspace of $W^{\otimes 3}$ spanned by all the elements $[w_1, [w_2, w_3]]$, with $w_1, w_2, w_3 \in W$. Note that $\mathcal{U} \subset T(\mathbb{R}, \mathbb{R}^d)$ in the following way: $\pi_0(W) = \{0\}$, $\pi_1(W) = \mathbb{R}^d$ and so on.

If A is a linear map from W into some Lie algebra $\mathcal{B}$, then there is a unique extension of A to a Lie algebra homomorphism $\tilde{A} : \mathcal{U} \to \mathcal{B}$, namely by defining $\tilde{A}([w_1, [w_2, w_3]]) = [Aw_1, [Aw_2, Aw_3]]$, for example. A Lie algebra with that property (like $\mathcal{U}$) is called *free* over $W = \mathbb{R} \oplus \mathbb{R}^d$.

An element of $\mathcal{U}^{(n)} = \pi_n(\mathcal{U})$ is called a Lie polynomial of degree n, whereas an infinite sequence of Lie brackets will be called Lie series. The Lie algebra $\mathcal{U}^{(n)}$ is nilpotent of degree n, i. e. $[u_1, [u_2, \cdots, [u_n, u_{n+1}] \cdots] = 0$ for any $u_1, \ldots, u_{n+1} \in \mathcal{U}^{(n)}$, because any application of n Lie brackets results in a sum of terms involving $n+1$ tensor products each. And since $\pi_0(u_i) = 0$, the tensor product of $n+1$ elements of $\mathcal{U}^{(n)}$ is 0 as an element of $\mathcal{U}^{(n)}$.

Theorem 4.3.3. *Let* $G^{(n)} = \exp\left(\mathcal{U}^{(n)}\right) \subset T^{(n)}(\mathbb{R}, \mathbb{R}^d)$*, then* $G^{(n)}$ *is a Lie group in* $T^{(n)}(\mathbb{R}, \mathbb{R}^d)$ *and* $\mathcal{U}^{(n)}$ *is its Lie algebra.*

For details see [10].

4.4 Lie Basis

In this subsection we will learn more about the Lie algebras $\mathcal{U}$ and $\mathcal{U}^{(n)}$ defined above. First we will define the notion of a universal enveloping algebra. Note that any associative algebra also has the structure of a Lie algebra, if one takes the commutator as Lie bracket, i. e. $[a, b] = ab - ba$.

Definition 4.4.1. Let $\mathcal{G}$ be a Lie algebra. A *universal enveloping algebra* of $\mathcal{G}$ is an associative algebra U (with unit) over the same field (e. g. $\mathbb{R}$) together with a Lie algebra homomorphism $\iota : \mathcal{G} \to U$, such that if A is another associative algebra (over the same field) and $\phi : \mathcal{G} \to A$ is another Lie algebra homomorphism, then there exists a unique homomorphism of associative algebras $\psi : U \to A$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathcal{G} & \xrightarrow{\iota} & U \\ & \searrow^{\phi} & \downarrow \psi \\ & & A \end{array}$$

Now observe that $T(\mathbb{R}, \mathbb{R}^d)$ is a universal enveloping algebra of $\mathcal{U}$. To see this, let A be some associative algebra (just like in Definition 4.4.1) and let ι be the inclusion of $\mathcal{U}$ into $T(\mathbb{R}, \mathbb{R}^d)$. Now for any $a \in \mathcal{U}$ define $\psi(a) = \phi(a)$. Since $\mathbb{R} \oplus \mathbb{R}^d \subset \mathcal{U}$, ψ is now defined for $\varepsilon_i,\ 0 \le i \le d$. Let

$$\psi(\varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_k}) = \phi(\varepsilon_{i_1}) \cdots \phi(\varepsilon_{i_k}).$$

Because of Lemma 4.3.2 and of Definition 4.3.1, this extends to a uniquely defined algebra homomorphism on $T(\mathbb{R}, \mathbb{R}^d)$.

Theorem 4.4.2 (Poincaré-Birkhoff-Witt). *Let* $\mathcal{G}$ *be a Lie algebra with a basis* $\mathcal{B}$ *equipped with a total order* $\preccurlyeq$*. Then the set*

$$\bigcup_{n \ge 0} \{l_1 \cdots l_n | l_1, \dots, l_n \in \mathcal{B}, l_1 \preccurlyeq \cdots \preccurlyeq l_n\}$$

forms a basis of the universal enveloping algebra $U(\mathcal{G})$*.*

See [14] for a proof.

Let $\mathcal{B}_{\mathcal{U}}$ be a basis of the Lie algebra $\mathcal{U}$. Assume that there is a total order $\preccurlyeq$ defined on $\mathcal{B}_{\mathcal{U}}$ and that any element of the basis is contained in some $U_k(\mathbb{R}, \mathbb{R}^d)$. Such a basis exists, e. g. the Lyndon words basis has these properties, see [14]. From Theorem 4.4.2 follows that

$$\bigcup_{n\in\mathbb{N}} \{l_1 \otimes \cdots \otimes l_n | l_1, \ldots, l_n \in \mathcal{B}_{\mathcal{U}}, l_1 \preccurlyeq \cdots \preccurlyeq l_n\}$$

forms a basis of the tensor algebra $T(\mathbb{R}, \mathbb{R}^d)$.

We will use another basis of the tensor algebra.

Corollary 4.4.3. *The set*

$$\bigcup_{n\in\mathbb{N}} \{(l_1, \ldots, l_n) | l_1, \ldots, l_n \in \mathcal{B}_{\mathcal{U}}, l_1 \preccurlyeq \cdots \preccurlyeq l_n\},$$

where the symmetrized product of Lie polynomials is defined by $(l_1, \ldots, l_n) = \frac{1}{n!} \sum_{\sigma \in S_n} l_{\sigma(1)} \otimes \cdots \otimes l_{\sigma(n)}$*, is a basis of* $T(\mathbb{R}, \mathbb{R}^d)$*.*

The basis of Corollary 4.4.3 yields a good formula for the exponential of Lie polynomials.

Proposition 4.4.4. *For Lie polynomials* $l_1, \ldots, l_n$ *and scalars* $\beta_1, \ldots, \beta_n$*, the exponential is given by*

$$\exp\left(\sum_{i=1}^n \beta_i l_i\right) = \sum_{k=0}^{\infty} \sum_{i_1+\cdots+i_n=k} \frac{\beta_1^{i_1} \cdots \beta_n^{i_n}}{i_1! \cdots i_n!} (\underbrace{l_1, \ldots, l_1}_{i_1 \textit{ times}}, \ldots, \underbrace{l_n, \ldots, l_n}_{i_n \textit{ times}}).$$

Proof. By (4.13) we have to show that

$$\frac{1}{k!}\left(\sum_{i=1}^n \beta_i l_i\right)^k = \sum_{i_1+\cdots+i_n=k} \frac{\beta_1^{i_1} \cdots \beta_n^{i_n}}{i_1! \cdots i_n!} (l_1, \ldots, l_1, \ldots, l_n, \ldots, l_n),$$

for all $k \in \mathbb{N}$. We will do so by induction with respect to k. Using an appropriate interpretation of the terms, the formula holds true for $k = 0$ and for $k = 1$. Now suppose the formula is correct for $k \geq 1$. Then

$$\begin{aligned} \frac{1}{(k+1)!}\left(\sum_{i=1}^n \beta_i l_i\right)^{(k+1)} &= \\ &= \frac{1}{k+1}\left(\sum_{i=1}^n \beta_i l_i\right)\left(\sum_{i_1+\cdots+i_n=k} \frac{\beta_1^{i_1} \cdots \beta_n^{i_n}}{i_1! \cdots i_n!} (l_1, \ldots, l_1, \ldots, l_n, \ldots, l_n)\right) = \\ = \sum_{i_1+\cdots+i_n=k} \sum_{j=1}^n \frac{\beta_1^{i_1} \cdots \beta_j^{i_j+1} \cdots \beta_n^{i_n}}{i_1! \cdots (i_j+1)! \cdots i_n!} &\left(\frac{i_j+1}{(k+1)!} \sum_{\sigma \in S_n} l_{m(\sigma(1))} \otimes \cdots \otimes l_{m(\sigma(n))} \otimes l_j\right), \end{aligned}$$

where $m(1) = 1, \ldots, m(i_1) = 1, m(i_1+1) = 2, \ldots, m(i_1+i_2) = 2, \ldots$ (provided that $i_1, i_2 > 0$ and appropriately changed otherwise). Now suppose for a fixed j we want to reorder the sequence $(l_1, \ldots, l_n)$, where l_1 appears i_1 times, l_2 appears i_2 times and so on, but l_j appears $i_j + 1$ times, such that the last element of the reordered sequence is l_j. For any permutation of the remaining k elements there are $i_j + 1$ possibilities to do so, because we may put any of the l_j's in the last position. Therefore, if we sum up the terms for all the j's, we get the equation and so the proposition holds.

4.5 Chen Theorem

As we have already seen, the iterated integrals of paths of bounded variation can be interpreted as elements of tensor products of the respective spaces. Remember that in our situation this space is $W = \mathbb{R} \oplus \mathbb{R}^d$.

Definition 4.5.1. Let $\omega \in C^0_{0,bv}([0,T], \mathbb{R}^d)$ be a path of bounded variation. The *Chen series* or series of iterated integrals of ω is defined by

$$X_{s,t}(\omega) = \sum_{k=0}^{\infty} \iint\limits_{s<t_1<\cdots<t_k<t} d\omega(t_1) \otimes \cdots \otimes d\omega(t_k),$$

for $s < t \in [0,T]$. $X_{s,t}(\omega)$ is an element of $T(\mathbb{R}, \mathbb{R}^d)$.

As indicated above, since the collection of all the $\varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_k}$ makes up a basis of the tensor algebra, the Chen series can also be written as

$$X_{s,t}(\omega) = \sum_{k=0}^{\infty} \sum_{(i_1,\ldots,i_n) \in \mathcal{A}_k \setminus \mathcal{A}_{k-1}} \iint\limits_{s<t_1<\cdots<t_n<t} d\omega^{i_1}(t_1) \cdots d\omega^{i_n}(t_n) \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_n}. \tag{4.16}$$

Here, for each k, the respective term is an element of $U_k(\mathbb{R}, \mathbb{R}^d)$. The truncated series $\pi_n(X_{s,t}(\omega))$ is denoted by $X^{(n)}_{s,t}(\omega)$. We see that $X^{(n)}_{s,t}(\omega)$ is given by (4.16), if we sum up to n only.

Theorem 4.5.2 (Chen's Theorem). *For a process ω of bounded variation, the series of iterated integrals is* multiplicative, *i. e.*

$$X_{r,s}(\omega) \otimes X_{s,t}(\omega) = X_{r,t}(\omega).$$

$\log(X_{s,t}(\omega))$ *is a Lie series. On the other hand, if $\mathcal{L} \in \mathcal{U}^{(n)}$ is a Lie polynomial, then there exists a path ω of bounded variation, such that*

$$\pi_n(\log(X_{s,t}(\omega))) = \mathcal{L}.$$

Proof. First let ω be a path of bounded variation. As usual, we denote by $X^i_{s,t}$ the component of $X_{s,t} = X_{s,t}(\omega)$ in $U_i(\mathbb{R}, \mathbb{R}^d)$. As a preliminary result we show the following equation:

$$\iint\limits_{r<t_1<\cdots<t_i<t} d\omega^{l_1}(t_1) \cdots d\omega^{l_i}(t_i) = \\ = \sum_{j=0}^{i} \iint\limits_{s<t_{j+1}<\cdots<t_i<t} d\omega^{l_{j+1}}(t_{j+1}) \cdots d\omega^{l_i}(t_i) \iint\limits_{r<t_1<\cdots<t_j<s} d\omega^{l_1}(t_1) \cdots d\omega^{l_j}(t_j). \tag{4.17}$$

This is done by working with indicator functions. Per definition the left hand side of (4.17) is

$$\int_r^t \int_r^{t_i} \cdots \int_r^{t_2} d\omega^{l_1}(t_1) \cdots d\omega^{l_i}(t_i) =$$

$$= \int_r^t \cdots \int_r^{t_3} \left[\left(\int_r^s d\omega^{l_1}(t_1) + \int_s^{t_2} d\omega^{l_1}(t_1) \right) \mathbf{1}_{[t_2>s]} + \int_r^{t_2} d\omega^{l_1}(t_1) \mathbf{1}_{[t_2<s]} \right] d\omega^{l_2}(t_2) \cdots d\omega^{l_i}(t_i) =$$

$$= \int_s^t \cdots \int_s^{t_2} d\omega^{l_1}(t_1) \cdots d\omega^{l_i}(t_i) + \left(\int_s^t \cdots \int_s^{t_3} d\omega^{l_2}(t_2) \cdots d\omega^{l_i}(t_i) \right) \int_r^s d\omega^{l_1}(t_1) + \cdots .$$

By iterating this manipulation, one gets the right hand side of (4.17). We have to show that

$$X_{r,t}^k = \sum_{(i_1,\ldots,i_n)\in\mathcal{A}_k\setminus\mathcal{A}_{k-1}} \int_r^t \cdots \int_r^{t_2} d\omega^{i_1}(t_1) \cdots d\omega^{i_n}(t_n) \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_n}$$

equals

$$\sum_{l=0}^k X_{r,s}^l \otimes X_{s,t}^{k-l} = \sum_{l=0}^k \sum_{\substack{(i_1,\ldots,i_j)\in\mathcal{A}_l\setminus\mathcal{A}_{l-1} \\ (i_{j+1},\ldots,i_n)\in\mathcal{A}_{k-l}\setminus\mathcal{A}_{k-l-1}}} \left(\int_r^s \cdots \int_r^{t_2} d\omega^{i_1}(t_1) \cdots d\omega^{i_j}(t_j) \right)$$
$$\left(\int_s^t \cdots \int_s^{t_{j+2}} d\omega^{i_{j+1}}(t_{j+1}) \cdots d\omega^{i_n}(t_n) \right) \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_n}.$$

Now, by (4.17), we get

$$X_{r,t}^k = \sum_{(i_1,\ldots,i_n)\in\mathcal{A}_k\setminus\mathcal{A}_{k-1}} \sum_{j=0}^n \left(\int_r^s \cdots \int_r^{t_2} d\omega^{i_1}(t_1) \cdots d\omega^{i_j}(t_j) \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_j} \right) \otimes$$
$$\otimes \left(\int_s^t \cdots \int_s^{t_{j+2}} d\omega^{i_{j+1}}(t_{j+1}) \cdots d\omega^{i_n}(t_n) \varepsilon_{i_{j+1}} \otimes \cdots \otimes \varepsilon_{i_n} \right).$$

If we define

$$X_{r,s}^{i_1,\ldots,i_j} = \iint_{r<t_1<\cdots<t_j<s} d\omega^{i_1}(t_1) \cdots d\omega^{i_j}(t_j) \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_j},$$

it remains to show that

$$\sum_{(i_1,\ldots,i_n)\in\mathcal{A}_k\setminus\mathcal{A}_{k-1}} \sum_{j=0}^n X_{r,s}^{i_1,\ldots,i_j} \otimes X_{s,t}^{i_{j+1},\ldots,i_n} =$$
$$= \sum_{l=0}^k \sum_{\substack{(i_1,\ldots,i_j)\in\mathcal{A}_l\setminus\mathcal{A}_{l-1} \\ (i_{j+1},\ldots,i_n)\in\mathcal{A}_{k-l}\setminus\mathcal{A}_{k-l-1}}} X_{r,s}^{i_1,\ldots,i_j} \otimes X_{s,t}^{i_{j+1},\ldots,i_n}. \quad (4.18)$$

But if $(i_1, \dots, i_n) \in \mathcal{A}_k \setminus \mathcal{A}_{k-1}$ (i. e. $n + |\{m : i_m = 0\}| = k$), then for fixed j there is exactly one l, namely $l = j + |\{m : i_m = 0, 1 \le m \le j\}|$, such that $(i_1, \dots, i_j) \in \mathcal{A}_l \setminus \mathcal{A}_{l-1}$ and $(i_{j+1}, \dots, i_n) \in \mathcal{A}_{k-l} \setminus \mathcal{A}_{k-l-1}$. Since the converse is true, too, exactly the same summands occur in both sides of (4.18). So the first part of the theorem is proved. Since the projections π_n are algebra-homomorphisms, the truncated series $X_{s,t}^{(n)}(\omega)$ are multiplicative, too. They satisfy the differential equation

$$dX_{s,t}^{(n)}(\omega) = \sum_{i=0}^{d} X_{s,t}^{(n)}(\omega) \otimes \varepsilon^i d\omega^i(t) \tag{4.19}$$

with $X_{s,s}^{(n)} = 1 = (1, 0, \dots, 0)$. Note that (4.19) actually makes sense, because we are now working within a finite dimensional real vector space. Since the starting point 1 is an element of the Lie group $G^{(n)}$, the solution does not leave the group, see Remark 4.5.3. Therefore $\log\left(X_{s,t}^{(n)}\right) \in \mathcal{U}^{(n)}$ for all n, and so $X_{s,t}(\omega)$ is a Lie series.

Let $\mathcal{L} \in \mathcal{U}^{(n)}$ be a Lie polynomial and let $g = \exp(\mathcal{L}) \in G^{(n)}$. The group $G^{(n)}$ is connected and thus path connected, i. e. for $g \in G^{(n)}$ there is a path $h \in C^\infty(\mathbb{R}_{\ge 0}, G^{(n)}$ with $h(0) = 1$ and $h(1) = g$. Let $d^{(r)}$ denote the right logarithmic derivative, see [7], and define $\omega(t) = d^{(r)}h(t)$. Then $X_{0,1}^{(n)}(\omega) = Evol^{(r)}(\omega)(1) = \exp(\mathcal{L})$, which finishes the proof.

Remark 4.5.3. If ω is a smooth, i. e. infinitely often differentiable, path, then the differential equation (4.19) with initial condition $X_{s,s}^{(n)}(\omega) = 1$ has a solution $X_{s,t}^{(n)}(\omega) \in G^{(n)}$, $\forall t \ge s$, see [7]. Now any bounded variation path ω can be approximated by a sequence of smooth paths ω_k. For ω the differential equation (4.19) has a solution $X_{s,t}^{(n)}(\omega)$ in $T^{(n)}(\mathbb{R}, \mathbb{R}^d)$, which can be shown by a Picard-Lindelöf argument, whereas for the smooth paths ω_k the solution of the corresponding differential equation $X_{s,t}^{(n)}(\omega_k)$ is in the Lie group. $X_{s,t}^{(n)}(\omega_k)$ converges to $X_{s,t}^{(n)}(\omega)$ and so $X_{s,t}^{(n)}(\omega) \in G^{(n)}$, because $G^{(n)}$ is locally closed.

Example 4.5.4. As an example we compute the projection on $T^{(2)}$ of the logarithm of the series of iterated integrals of ω. By (4.16) we have

$$a = X_{0,1}(\omega) = \left(1\ , \sum_{i=1}^{d} \omega^i(1)\varepsilon_i\ ,\ \varepsilon_0 + \sum_{i,j=1}^{d} \int_0^1 \omega^i(t) d\omega^j(t) \varepsilon_i \otimes \varepsilon_j\ ,\ \dots\right).$$

Therefore, using the same notation as in (4.12), $c_i = a_i$ for $i > 0$, $c_0 = 0$ and

$$c^2 = \left(0, 0, \sum_{i,j=1}^{d} \omega^i(1)\omega^j(1)\varepsilon_i \otimes \varepsilon_j, \dots\right).$$

By partial integration we get

$$\omega^i(1)\omega^j(1) = \int_0^1 \omega^i(t) d\omega^j(t) + \int_0^1 \omega^j(t) d\omega^i(t). \tag{4.20}$$

Now by (4.20)

$$\log\left(X_{0,1}(\omega)\right) =$$
$$= \left(0\ , \sum_{i=1}^{d} \omega^i(1)\varepsilon_i\ ,\ \varepsilon_0 + \sum_{i,j=1}^{d} \left[\int_0^1 \omega^i(t)d\omega^j(t) - \frac{1}{2}\omega^i(1)\omega^j(1)\right] \varepsilon_i \otimes \varepsilon_j\ ,\ \ldots\right)$$
$$= \left(0\ , \sum_{i=1}^{d} \omega^i(1)\varepsilon_i\ ,\ \sum_{1\le i<j\le d} \left[\int_0^1 \omega^i(t)d\omega^j(t) - \frac{1}{2}\omega^i(1)\omega^j(1)\right] [\varepsilon_i, \varepsilon_j]\ ,\ \ldots\right).$$

Once more by (4.20) follows that

$$\pi_2\left(\log\left(X_{0,1}\right)\right) = \varepsilon_0 + \sum_{i=1}^{d} \omega^i(1)\varepsilon_i + \frac{1}{2} \sum_{1\le i<j\le d} \int_0^1 (\omega^i(t)d\omega^j(t) - \omega^j(t)d\omega^i(t))[\varepsilon_i, \varepsilon_j].$$

4.6 Stratonovich Integrals

In fact, Chen's Theorem does not only apply to functions of bounded variation, but also to the Brownian motion, if the integral is understood to be of Stratonovich type, see [10], of course resulting in random Lie series. The (random) series of iterated Stratonovich integrals will be denoted by $X_{s,t}(oB)$.

Remark 4.6.1. Let $\lambda \in \mathbb{R}$ be a scalar and define a map $d_\lambda : T(\mathbb{R}, \mathbb{R}^d) \to T(\mathbb{R}, \mathbb{R}^d)$ by $d_\lambda(a) = (a_0, \lambda a_1, \lambda^2 a_2, \ldots)$. Then we can rewrite (2.12) in the following way

$$d_{\sqrt{t-s}}\left(X_{0,1}(oB)\right) =^{\mathcal{L}} X_{s,t}(oB).$$

Remark 4.6.2. Let Γ be the algebra homomorphism from the tensor algebra to the algebra of vector fields on $\mathbb{R}^d$ (with composition as multiplication) that maps ε_i to V_i. Then we have

$$\Gamma\left(X_{0,t}^{(m)}(oB)\right) f(x) =$$
$$= \Gamma\left(\sum_{(i_1,\ldots,i_k)\in\mathcal{A}_m} \iint_{0<t_1<\cdots<t_k<t} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k} \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_k}\right) f(x) =$$
$$= \sum_{(i_1,\ldots,i_k)\in\mathcal{A}_m} V_{i_1} \cdots V_{i_k} f(x) \iint_{0<t_1<\cdots<t_k<t} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k},$$

and by the stochastic Taylor expansion we get

$$f(\xi_{t,x}) = \Gamma\left(X_{0,t}^{(m)}(oB)\right) f(x) + R_m(t, x, f). \tag{4.21}$$

Lemma 4.6.3. *The paths $\omega_1, \ldots, \omega_n$ and the weights $\lambda_1, \ldots, \lambda_n$ define a cubature formula on Wiener space of degree m if and only if*

$$E\left(X_{0,1}^{(m)}(oB)\right) = \sum_{j=1}^{n} \lambda_j X_{0,1}^{(m)}(\omega_j).$$

Proof. Assume that $\omega_1, \dots, \omega_n$ and $\lambda_1, \dots, \lambda_n$ indeed define a cubature formula. According to Definition 3.2.1 this means that

$$E\left(\iint\limits_{0<t_1<\cdots<t_k<1} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k}\right) = \sum_{j=1}^{n} \lambda_j \iint\limits_{0<t_1<\cdots<t_k<1} d\omega_j^{i_1}(t_1) \cdots d\omega_j^{i_k}(t_k),$$

for all $(i_1, \dots, i_k) \in \mathcal{A}_m$. Therefore

$$E\left(X_{0,1}^{(m)}(oB)\right) = \sum_{(i_1,\dots,i_k)\in\mathcal{A}_m} E\left(\iint\limits_{0<t_1<\cdots<t_k} \circ dB_{t_1}^{i_1} \cdots \circ dB_{t_k}^{i_k}\right) \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_k} =$$
$$= \sum_{j=1}^{n} \lambda_j \sum_{(i_1,\dots,i_k)\in\mathcal{A}_m} \iint\limits_{0<t_1<\cdots<t_k<1} d\omega_j^{i_1}(t_1) \cdots d\omega_j^{i_k}(t_k) \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_k} =$$
$$= \sum_{j=1}^{n} \lambda_j X_{0,1}(\omega_j).$$

This formula proves the converse, too, since the coefficients of a vector with respect to a given basis are uniquely determined.

According to Chen's Theorem, $\log(X_{0,1}^{(m)}(\omega_j))$ is a Lie polynomial for ω_j with bounded variation. So the right hand side in Lemma 4.6.3 can be written as $\sum_{j=1}^{n} \lambda_j \pi_m(\exp(\mathcal{L}_j))$, for some Lie polynomials $\mathcal{L}_1, \dots, \mathcal{L}_n$. On the other hand, for any Lie polynomial $\mathcal{L} \in \mathcal{U}^{(m)}$ there is a path of bounded variation ω such that $\pi_m(\log(X_{0,1}(\omega))) = \mathcal{L}$. Therefore we can equivalently define the notion of a cubature formula in terms of Lie polynomials.

Definition 4.6.4. The Lie polynomials $\mathcal{L}_1, \dots, \mathcal{L}_n \in \mathcal{U}^{(m)}$ and the positive weights $\lambda_1, \dots, \lambda_n$ $(n, m \in \mathbb{N})$ define a *cubature formula on Wiener space of degree m* if and only if

$$E\left(X_{0,1}^{(m)}(oB)\right) = \sum_{j=1}^{n} \lambda_j \exp(\mathcal{L}_j),$$

where exp denotes the exponential function in $T^{(m)}(\mathbb{R}, \mathbb{R}^d)$, of course.

In order to find a cubature formula as in Definition 4.6.4, it would be nice if we knew the expectation of the series of iterated Stratonivich integrals in terms of elements of the tensor algebra.

Proposition 4.6.5.

$$E\left(X_{0,1}(oB)\right) = \exp\left(\varepsilon_0 + \frac{1}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right).$$

Proof. We proof the equation by using the connection between solutions of certain differential equations and expectations of solutions of stochastic differential equations, see Section 2.1. So $u(t,x) = E(f(\xi_{T-t,x}))$ is a solution of the PDE $\frac{\partial u}{\partial t}(t,x) = -Lu(t,x)$ with initial condition $u(T,x) = f(x)$, where

$$L = V_0 + \frac{1}{2}\left(V_1^2 + \cdots + V_d^2\right) = \Gamma\left(\varepsilon_0 + \frac{1}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right).$$

Therefore using Taylor expansion of $u(T-t,x) = E(f(\xi_{t,x}))$ we get

$$E(f(\xi_{t,x})) = \sum_{k=0}^{m} \frac{t^k L^k}{k!} f(x) + o(t^m), \tag{4.22}$$

for any m.

On the other hand, by Remark 4.6.2 and by the stochastic Taylor expansion, we have

$$E(f(\xi_{t,x})) = \Gamma\left(E\left(X_{0,t}^{(2m)}(oB)\right)\right) f(x) + o(t^m)$$

and with Remark 4.6.1 we get

$$E(f(\xi_{t,x})) = \Gamma\left(d_{\sqrt{t}} E\left(X_{0,1}^{(2m)}(oB)\right)\right) f(x) + o(t^m). \tag{4.23}$$

Since

$$\exp\left(t\varepsilon_0 + \frac{t}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right) = \sum_{k\geq 0} \frac{t^k}{k!}\left(\varepsilon_0 + \sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right)^k$$

and since $\left(\varepsilon_0 + \sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right)^k \in T^{(2m)}(\mathbb{R}, \mathbb{R}^d)$ for all $0 \leq k \leq m$, we additionally have

$$\sum_{k=0}^{m} \frac{t^k \Gamma\left(\varepsilon_0 + \frac{1}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right)^k}{k!} = \Gamma\left(\pi_{2m}\left(\exp\left(t\varepsilon_0 + \frac{t}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right)\right)\right). \tag{4.24}$$

Now (4.22) can be rewritten using (4.24)

$$\begin{aligned} E(f(\xi_{t,x})) &= \sum_{k=0}^{m} \frac{t^k \Gamma\left(\varepsilon_0 + \frac{1}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right)^k}{k!} f(x) + o(t^m) \\ &= \Gamma\left(\pi_{2m}\left(\exp\left(t\varepsilon_0 + \frac{t}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right)\right)\right) f(x) + o(t^m) \\ &= \Gamma\left(\pi_{2m}\left(d_{\sqrt{t}} \exp\left(\varepsilon_0 + \frac{1}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right)\right)\right) f(x) + o(t^m), \end{aligned}$$

where the last line is true because $t\varepsilon_0 + \frac{t}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i \in U_2(\mathbb{R}, \mathbb{R}^d)$ and thus its exponential looks like $(1, 0, t(\varepsilon_0 + 1/2\sum \varepsilon_i \otimes \varepsilon_i), 0, t^2(\varepsilon_0 + 1/2\sum \varepsilon_i \otimes \varepsilon_i)^2, \ldots)$. Comparing with (4.23), we get

$$d_{\sqrt{t}}\left(E\left(X_{0,1}^{(2m)}(oB)\right)\right) = \pi_{2m} d_{\sqrt{t}}\left(\exp\left(\varepsilon_0 + \frac{1}{2}\sum_{i=1}^{d} \varepsilon_i \otimes \varepsilon_i\right)\right).$$

Now take $t = 1$ and let $m \to \infty$ to get the desired result.

4.7 A Proof of the Tchakaloff Theorem on Wiener Space

As promised in Chapter 3, we now give a proof of Theorem 3.2.3. By Definition 4.6.4 and the comments above it, we have to show the following lemma.

Lemma 4.7.1. *Let m be a natural number. For some $n \leq |\mathcal{A}_m|$ there are n Lie polynomials $\mathcal{L}_1, \ldots, \mathcal{L}_n$ and positive weights $\lambda_1, \ldots, \lambda_n$ defining a cubature formula on Wiener space.*

Proof. As we have already observed in Section 4.6, the series of iterated Stratonovich integrals is the exponential of a random Lie series $\mathcal{L}$ and so $X_{0,1}^{(m)} = \pi_m(\exp(\mathcal{L})) = \exp(\pi_m(\mathcal{L}))$. Let $\mathcal{B}_{\mathcal{U}}$ be a basis of the Lie algebra $\mathcal{U}$. For $l \in \mathcal{B}_{\mathcal{U}}^{(m)} = \mathcal{B}_{\mathcal{U}} \cap \mathcal{U}^{(m)}$, the coefficient of $\pi_m(\mathcal{L})$ corresponding to l is a real valued random variable denoted by X_l. Note that all the multivariable moments of

$$X = (X_l)_{l \in \mathcal{B}_{\mathcal{U}}^{(m)}}$$

exist, because $E(X_{0,1}(oB)) = E(\exp(\mathcal{L}))$ exists and the coefficients of $\pi_k(E(\exp(\mathcal{L})))$ just are multivariable moments of X. X is understood to be a random variable with values in $\mathbb{R}^{|\mathcal{B}_{\mathcal{U}}^{(m)}|}$. By Theorem 3.1.3, part B), there is a natural number n and there are positive numbers $\lambda_1, \ldots, \lambda_n$ and vectors $\tilde{X}_1 = (\tilde{X}_{1,l})_{l \in \mathcal{B}_{\mathcal{U}}^{(m)}}, \ldots, \tilde{X}_n = (\tilde{X}_{n,l})_{l \in \mathcal{B}_{\mathcal{U}}^{(m)}}$, such that for any multi-index $p = (p_l)_{l \in \mathcal{B}_{\mathcal{U}}^{(m)}}$ with $|p| \leq m$

$$E\left(X^p\right) = \sum_{j=1}^{n} \lambda_j \tilde{X}_j^p.$$

With $\mathcal{L}_j = \sum_{l \in \mathcal{B}_{\mathcal{U}}^{(m)}} \tilde{X}_{j,l} l$ for $j = 1, \ldots, n$, we get

$$E(\pi_m(\exp(\mathcal{L}))) = \sum_{j=1}^{n} \lambda_j \pi_m(\exp(\mathcal{L}_j)).$$

So the $\mathcal{L}_1, \ldots, \mathcal{L}_n$ and $\lambda_1, \ldots, \lambda_n$ define a cubature formula on Wiener space with size n of degree m. According to Theorem 3.1.3 B), n can be chosen to be smaller than or equal to $|\mathcal{A}_m| = \dim T^{(m)}(\mathbb{R}, \mathbb{R}^d)$, which concludes the proof.

Chapter 5

Formulae for Some Other PDE's

Now we obtained an approximation method for the solution of the PDE (2.4) having a stochastic representation (2.1). But, of course, there are differential operators lacking such a representation, e. g. the bi-Laplacian operator

$$\Delta^2 = \left(\sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} \right)^2 .$$

In order to approximate the solution of the PDE (2.4) with $L = \Delta^2$ we have to find Lie polynomials $\mathcal{L}_1, \ldots, \mathcal{L}_n$ and weights $\lambda_1, \ldots, \lambda_n$ such that

$$\pi_m \exp\left(\left(\varepsilon_0 + \frac{1}{2} \sum_{j=1}^{d} \varepsilon_j \otimes \varepsilon_j \right)^2 \right) = \sum_{j=1}^{n} \lambda_j \exp(\mathcal{L}_j).$$

The calculations in this section are motivated by [6], especially by the remarks in Paragraph 4 therein.

5.1 Bochner's Subordination

Before being able to present Bochner's subordination, we have to give some definitions. Let X be a Banach space.

Definition 5.1.1. A family $(T_t)_{t \geq 0}$ of continuous (i. e. bounded) linear operators $X \to X$ is called a *strongly continuous semi-group* of operators, if

$$T_0 = Id, \text{ where } Id \text{ is the identiy on } X, \tag{5.1}$$

$$T_{s+t} = T_s T_t \text{ for all } s, t \geq 0, \tag{5.2}$$

$$\lim_{t \to 0} T_t x = x \text{ for all } x \in X. \tag{5.3}$$

A strongly continuous semi-group $(T_t)_{t \geq 0}$ is called *continuous in norm*, if additionally or instead of (5.3)

$$\lim_{t \to 0} \|T_t - Id\| = 0, \tag{5.4}$$

where $\| \cdot \|$ is the operator norm on the space $L(X)$ of continuous linear operators on X.

For details see [16].

Definition 5.1.2. Let $(T_t)_{t\geq 0}$ be a strongly continuous semi-group on X. The operator A defined on

$$\operatorname{dom}(A) = \left\{ x \in X \,\middle|\, \lim_{h\to 0} \frac{T_h x - x}{h} \text{ exists} \right\}$$

by

$$Ax = \lim_{h\to 0} \frac{T_h x - x}{h}$$

is called *infinitesimal generator* of the semi-group $(T_t)_{t\geq 0}$.

Definition 5.1.3. A measure μ on $\mathbb{R}$ is called *infinitely divisible*, if for all $n \in \mathbb{N} \setminus \{0\}$ there is a measure μ_n on $\mathbb{R}$, such that

$$\mu = \mu_n * \cdots * \mu_n \qquad (n \text{ factors}),$$

where $*$ denotes convolution.

Of course, the condition in Definition 5.1.3 means that $\hat{\mu} = (\hat{\mu}_n)^n$, where $\hat{\mu}$ is the characteristic function of μ.

Definition 5.1.4. A family $(\mu_t)_{t\geq 0}$ of probability measures on $\mathbb{R}$ is called *convolution semi-group*, if

$$\forall s, t \geq 0 : \mu_{s+t} = \mu_s * \mu_t. \tag{5.5}$$

There is a strong relation between infinitely divisible probability measures and convolution semi-groups, that is formulated in the following Proposition 5.1.5.

Proposition 5.1.5. *Let $(\mu_t)_{t\geq 0}$ be a convolution semi-group. Then for all $t \geq 0$ the corresponding measure μ_t is infinitely divisible. Conversely, let μ be an infinitely divisible probability measure and $t_0 > 0$. Then there is one and only one weakly continuous convolution semi-group $(\mu_t)_{t\geq 0}$ such that $\mu_{t_0} = \mu$.*

A proof of Proposition 5.1.5 and other information on infinitely divisible measures and convolution semi-groups can be found in [1].

Theorem 5.1.6. *Let $A \in T^{(m)}(\mathbb{R}, \mathbb{R}^d)$ with $\pi_0(A) = 0$ and let $(\mu_t)_{t\geq 0}$ be a convolution semi-group defined on $\mathbb{R}$ such that*

$$\int_{\mathbb{R}} e^{\lambda x} \mu_t(dx) = e^{t\psi(\lambda)}$$

for some analytic function ψ defined on a neighborhood of 0. *Define*

$$S_t = \int_{\mathbb{R}} \exp(xA)\, \mu_t(dx),$$

where exp *denotes the exponential function in the truncated tensor algebra $T^{(m)}(\mathbb{R}, \mathbb{R}^d)$. Then*

$$S_t = \exp(t\psi(A)).$$

Proof. ψ is analytic, i. e. we have $\psi(\lambda) = \sum_{k=0}^{\infty} \frac{\psi^{(k)}(0)}{k!}\lambda^k$, and so for an element A of the tensor algebra we define

$$\psi(A) = \sum_{k=0}^{\infty} \frac{\psi^{(k)}(0)}{k!} A^k,$$

provided the sum exists. If $A \in T^{(m)}(\mathbb{R}, \mathbb{R}^d)$ with $\pi_0(A) = 0$, as in the Theorem, we do not have to worry about convergence because of the nilpotency of those elements of the truncated algebra.

By plugging the series representation of ψ into the series representation of the exponential function, we get

$$e^{t\psi(\lambda)} = 1 + p_1(t)\lambda + \cdots + p_m(t)\lambda^m + g(t,\lambda)\lambda^{m+1},$$

where

$$p_j(t)j! = \frac{\partial^j}{\partial \lambda^j}\bigg|_{\lambda=0} e^{t\psi(\lambda)} = \int_{\mathbb{R}} x^j \mu_t(dx)$$

is the j-th moment of μ_t for $j = 1, \ldots, m$. Therefore we get

$$\exp\left(t\psi(A)\right) = 1 + p_1(t)A + \cdots + p_m(t)A^m. \tag{5.6}$$

On the other hand,

$$\begin{aligned} S_t = \int_{\mathbb{R}} \exp\left(xA\right)\mu_t(dx) &= \int_{\mathbb{R}} \sum_{k=0}^{m} \frac{x^k}{k!} A^k \mu_t(dx) = \\ &= 1 + \sum_{k=1}^{m} \frac{1}{k!} \int_{\mathbb{R}} x^k \mu_t(dx) A^k = 1 + p_1(t)A + \cdots + p_m(t)A^m. \end{aligned}$$

Comparing with (5.6) we get $S_t = \exp\left(t\psi(A)\right)$.

Remark 5.1.7. Theorem 5.1.6 is closely related to Bochner's Subordination, see [2], [11] and [13]. In [13], R. S. Phillips proves a much more general version of the theorem: Let T_t be a strongly continuous semi-group of linear operators on a Banach space X with infinitesimal generator A, denote $\omega(t) = \log\|T_t\|$ and define $\omega_0 = \inf_{t>0} \frac{\omega(t)}{t}$. Suppose μ_t is a semi-group of complex Lebesgue-Stieltjes measures defined on the Borel σ-algebra of $[0,\infty)$, such that

$$\int_0^{\infty} e^{\omega(x)} |\mu_t|(dx) < \infty,$$

for all $t \geq 0$ and such that

$$t^{-1} \log \int_0^{\infty} e^{\lambda x} \mu_t(dx) = m\lambda + \int_0^{\infty} \left(e^{(\lambda-\omega_0)x} - 1\right) \nu(dx) + a,$$

for some positive measure ν and some constants $m \geq 0$ and $a \in \mathbb{R}$, defined for all complex λ with real part smaller than or equal to ω_0. Define

$$S_t x = \int_0^\infty T_s x \ \mu_t(ds),$$

where $x \in X$. Then $(S_t)_{t \geq 0}$ defines a strongly continuous semi-group of linear operators on X with infinitesimal generator B, whose domain contains the domain of A. For $x \in \text{dom}(A)$ we have

$$Bx = mAx + \int_0^\infty \left(e^{-\omega_0 s} T_s x - x\right) \nu(ds) + ax.$$

5.2 A First Example

We will use Theorem 5.1.6 to find a cubature formula for the bi-Laplacian operator (or rather for $\frac{1}{8}\Delta^2$). For $s \in \mathbb{R}_{\geq 0}$, let the Lie polynomials $\mathcal{L}_1(s), \ldots, \mathcal{L}_n(s)$ together with the positive weights $\lambda_1, \ldots, \lambda_n$ define a cubature formula of degree m in the sense of Definition 4.6.4, i. e.

$$\pi_m \left(E(X_{0,s}(oB))\right) = \pi_m \left(\exp\left(s(\varepsilon_0 + \frac{1}{2}\sum_{i=1}^d \varepsilon_i \otimes \varepsilon_i)\right)\right) = \sum_{j=1}^n \lambda_j \pi_m \left(X_{0,s}(\omega_{s,j})\right), \tag{5.7}$$

where $\omega_{s,1}, \ldots, \omega_{s,n} \in C^0_{0,bv}([0,s], \mathbb{R}^d)$ such that $\exp(\mathcal{L}_j(s)) = \pi_m \left(X_{0,s}(\omega_{s,j})\right)$, $j = 1, \ldots, n$.

For $t > 0$, let F_t be the distribution of a normal random variable X with mean 0 and variance t; we write $X \sim \mathcal{N}(0,t)$. Furthermore, let $F_0 = \delta_0$. Then $(F_t)_{t \geq 0}$ is a convolution semi-group, see [1], and $\psi(\lambda) = \lambda^2/2$, where ψ is defined as in Theorem 5.1.6. Therefore, with $A = \varepsilon_0 + \frac{1}{2}\sum_{j=1}^d \varepsilon_j \otimes \varepsilon_j$, we get

$$\int_{\mathbb{R}} \pi_m \left(\exp(sA)\right) F_t(dx) = \pi_m \left(\exp\left(\frac{t}{2}\left(\varepsilon_0 + \frac{1}{2}\sum_{j=1}^d \varepsilon_j \otimes \varepsilon_j\right)^2\right)\right). \tag{5.8}$$

Of course, F_t is supported on the entire real line, not just on $\mathbb{R}_{\geq 0}$, and (5.7) is only valid for $s \geq 0$. However, the left hand side in (5.7) makes sense for negative s, too.

With $V_0 = 0$ and $V_i = \frac{\partial}{\partial x_i}$, the last expression in (5.8) gives a cubature formula for approximating solutions of the PDE for $L = \frac{1}{8}\Delta^2$, provided we can extend (5.7) to the entire real line.

Following Funaki [6] we define

$$\bar{B}^j_s = \begin{cases} B^j_s, & s \geq 0 \\ iB^j_{-s}, & s \leq 0 \end{cases}, j = 1, \ldots, d, \tag{5.9}$$

where i is the imaginary unit $\sqrt{-1}$. Once more, $\bar{B}^0_s = s$ for $s \in \mathbb{R}$. Fix $N = d$ and let $V_0 = 0$, $V_j = \frac{\partial}{\partial x_j}$, $j = 1, \ldots, d$. The solution of the corresponding SDE for $s \in \mathbb{R}$ is given by

$$\xi_{s,x} = \bar{B}_s + x.$$

Definition 5.2.1. $\mathcal{D}_1$ is defined as the set of all functions $f : \mathbb{R}^d \to \mathbb{R}$ that can be extended to holomorphic functions defined on $\mathbb{C}^d$ mapping to $\mathbb{C}$ denoted by $\bar{f}$, such that

$$\forall x \in \mathbb{R}^d : \bar{f}(x) = f(x)$$
$$|f(z)| \exp(-h|z|^2), \ \left|\frac{\partial f}{\partial x_j}(z)\right| \exp(-h|z|^2), \text{ and } \left|\frac{\partial f}{y_j}(z)\right| \exp(-h|z|^2)$$

are bounded on $\mathbb{C}^d$ for all $h > 0$ and $j = 1, \ldots, d$, where $z = (z_1, \ldots, z_d)$ and $z_j = x_j + iy_j$, $j = 1, \ldots, d$.

For $f \in \mathcal{D}_1$, $u(s, x) = E(f(\xi_{s,x}))$ is a solution to

$$\begin{cases} \frac{\partial u}{\partial s} = \frac{1}{2}\Delta u(s, x), & s \in \mathbb{R} \setminus \{0\}, \ x \in \mathbb{R}^d \\ u(0, x) = f(x), & x \in \mathbb{R}^d \end{cases} . \tag{5.10}$$

Remark 5.2.2. In general, the heat equation (5.10) does not allow an extension to the past, i. e. to the negative half-line. Therefore it is necessary to choose the "initial functions" f from a class like $\mathcal{D}_1$ that guarantees the existence of a solution in both time directions.

As a consequence, we may proceed just as in the proof of Proposition 4.6.5 and get for $s \in \mathbb{R}$

$$E\left(X_{0,s}^{(m)}(o\bar{B})\right) = \pi_m \exp\left(s\left(\varepsilon_0 + \frac{1}{2}\sum_{j=1}^{d} \varepsilon_j^2\right)\right),$$

where $X_{0,s}(o\bar{B})$ is the series of iterated Stratonovich integrals of $\bar{B}$ with

$$\iint\limits_{0>t_1>\cdots>t_k>s} \circ d\bar{B}_{t_1}^{j_1} \cdots \circ d\bar{B}_{t_k}^{j_k} = i^{\operatorname{ord}(j_1,\ldots,j_k)} \iint\limits_{0<t_1<\cdots<t_k<-s} \circ dB_{t_1}^{j_1} \cdots \circ dB_{t_k}^{j_k}$$

for $s < 0$. Note that $X_{0,s}(o\bar{B})$ is a random Lie series in $T(W \oplus iW)$, not in $T(W)$ as before!

If $\omega_1, \ldots, \omega_n \in C_{0,bv}^0([0,1], \mathbb{R}^d)$ and $\lambda_1, \ldots, \lambda_n > 0$ define a cubature formula on Wiener space at time 1, we have constructed paths $\omega_{1,s}, \ldots, \omega_{n,s}$ defining a cubature formula at time $s > 0$. We extend this construction to $\mathbb{R}$ by defining

$$\bar{\omega}_{l,s}^j(x) = \begin{cases} \sqrt{|s|}\omega_l^j(x/|s|), & x \in [0, |s|] \\ i\sqrt{|s|}\omega_l^j(-x/|s|), & x \in [-|s|, 0] \end{cases} , \tag{5.11}$$

for $s \in \mathbb{R}$, $j = 1, \ldots, d$ and $l = 1, \ldots, n$. Obviously, $\bar{\omega}_{1,s}, \ldots, \bar{\omega}_{n,s}$ now define a cubature formula on Wiener space for $s \in \mathbb{R}$, i. e. for $(j_1, \ldots, j_k) \in \mathcal{A}_m$

$$E\left(\iint\limits_{0>t_1>\cdots>t_k>s} \circ d\bar{B}_{t_1}^{j_1} \cdots \circ d\bar{B}_{t_k}^{j_k}\right) = \sum_{l=1}^{n} \lambda_l \iint\limits_{0>t_1>\cdots>t_k>s} d\bar{\omega}_{l,s}^{j_1}(t_1) \cdots d\bar{\omega}_{l,s}^{j_k}(t_k),$$

for $s < 0$. Define Lie polynomials $\mathcal{L}_l(s) \in U^{(m)}(W \oplus iW)$ for $l = 1, \ldots, n$ and $s < 0$ by

$$\exp\left(\mathcal{L}_l(s)\right) = X_{0,s}^{(m)}(\bar{\omega}_{l,s}),$$

then (5.7) holds for $s \in \mathbb{R}$ and we get

$$\pi_m \exp\left(\frac{t}{2}\left(\varepsilon_0 + \frac{1}{2}\sum_{j=1}^{d}\varepsilon_j^2\right)^2\right) = \sum_{l=1}^{n}\lambda_l \int_{\mathbb{R}} \exp(\mathcal{L}_l(s))F_t(ds), \tag{5.12}$$

for $t > 0$. From (5.11) we see that $\exp(\mathcal{L}_l(s)) = X_{0,s}^{(m)}(\bar{\omega}_{l,s}) = d_{\sqrt{s}}\left(X_{0,1}^{(m)}(\omega_l)\right)$ is a polynomial in $\sqrt{s}$ of degree $|\mathcal{A}_m|$, where $\sqrt{-x} = i\sqrt{x}$ for $x > 0$, compare with Remark 4.6.1 and the calculations in the proof of Proposition 3.2.4. Therefore, we may apply a complex version of Tchakaloff's Theorem and get

$$\int_{\mathbb{R}} \exp(\mathcal{L}_l(s))F_t(ds) = \sum_{j=1}^{k}\tilde{\lambda}_j \exp(\mathcal{L}_l(t_j)), \tag{5.13}$$

where $\tilde{\lambda}_1, \dots, \tilde{\lambda}_k > 0$ and $t_1, \dots, t_k \in \mathbb{R}$ define a (complex) cubature formula of degree $|\mathcal{A}_m|$ for the distribution of $\sqrt{X}$ for $X \sim \mathcal{N}(0, t)$.

Remark 5.2.3. More exactly, we do apply the real version of Tchakaloff's Theorem twice: First we use Tchakaloff's Theorem for

$$\int_{\mathbb{R}_{\geq 0}} \exp(\mathcal{L}_l(s))F_t(ds),$$

then for

$$\int_{]-\infty,0[} \exp(\mathcal{L}_l(s))F_t(ds),$$

where the integrand is purely imaginary.

Combining (5.13) and (5.12) we get

$$\pi_m \exp\left(\frac{t}{2}\left(\varepsilon_0 + \frac{1}{2}\sum_{j=1}^{d}\varepsilon_j \otimes \varepsilon_j\right)^2\right) = \sum_{l=1}^{n}\sum_{j=1}^{k}\lambda_l\tilde{\lambda}_j \exp(\mathcal{L}_l(t_j)). \tag{5.14}$$

Clearly, (5.14) defines a (complex) cubature formula on Wiener space for some fourth order PDE.

Let $z_{j,l}$ denote the solution of the ODE

$$dy_{t,x} = \sum_{i=0}^{d} V_i(y_{t,x})d\bar{\omega}_l^i(t)$$

with initial value $y_{0,x} = x$ at time t_j, where $V_i(u) = u^i$, i. e. the i-th component of the vector u, and $V_0 = 0$. Note that $z_{j,l}$ can be a complex number now. Then

$$\Gamma\left(\exp\left(\frac{1}{2}\left(\varepsilon_0 + \frac{1}{2}\sum_{i=1}^{d}\varepsilon_i \otimes \varepsilon_i\right)^2\right)\right) f(x) = \sum_{l=1}^{n}\sum_{j=1}^{k}\lambda_l\tilde{\lambda}_j f(z_{j,l})$$

is an approximation to the solution of the PDE

$$\frac{\partial u}{\partial t} = \frac{1}{8}\Delta^2 u(t, x),$$

with $u(0, x) = f(x)$ for $f \in \mathcal{D}_1$ at time $t = 1$.

Bibliography

[1] Heinz Bauer: *Wahrscheinlichkeitstheorie*, de Gruyter Lehrbuch, Walter de Gruyter, Berlin, New York 2002.

[2] Salomon Bochner: *Harmonic Analysis and the Theory of Probability*, Univ. of California Press, Berkeley and Los Angeles, 1955.

[3] John B. Conway: *A Course in Functional Analysis*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1985.

[4] Raúl E. Curto, Lawrence A. Fialkow: *A Duality Proof of Tchakaloff's Theorem*, J. Math. Anal. Appl. 269, No.2, 519-532 (2002).

[5] Jürgen Elstrodt: *Mass- und Integrationstheorie*, Springer-Verlag, Berlin, 1999.

[6] Tadahisa Funaki: *Probabilistic Construction of the Solution of Some Higher Order Parabolic Differential Equation*, Proc. Japan Academy Ser. A, 55, 176–179.

[7] Andreas Kriegl, Peter Michor: *The Convenient Setting of Global Analysis*, Mathematical Surveys and Monographs 53, AMS, 119.

[8] John M. Lee: *Introduction to Smooth Manifolds*, Graduate Texts in Mathematics, Springer-Verlag, New York, 2003.

[9] Terry Lyons, Nicolas Victoir: *Cubature on Wiener Space*, Proc. Royal Soc. A, to appear

[10] Terry Lyons: *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana 14 (1998), No. 2, 215–310.

[11] Edward Nelson: *A Functional Calculus Using Singular Laplace Integrals*, Trans. Amer. Math. Soc. 88 (1958), 400–413.

[12] Bernt Øksendal: *Stochastic Differential Equations. An Introduction with Applications*, Universitext, Springer-Verlag, Berlin Heidelberg New York, 2000.

[13] Ralph S. Phillips: *On the Generation of Semigroups of Linear Operators*, Pacific J. Math. 2 (1952), 343–369.

[14] Christophe Reutenauer: *Free Lie Algebras*, London Mathematical Society Monographs, New Series 7. Oxford Science Publications, 1993.

[15] Vladimir Tchakaloff: *Formules de cubatures mécaniques à coefficients non négatifs*, Bull. Sci. Math. (2) 81(1957), 123–134.

[16] Dirk Werner: *Funktionalanalysis*, Springer-Verlag, Berlin Heidelberg New York, 2000.