Malliavin's calculus and applications in stochastic control and finance

Warsaw, March, April 2008

Peter Imkeller

version of 5. April 2008

Malliavin's calculus has been developed for the study of the smoothness of measures on infinite dimensional spaces. It provides a stochastic access to the analytic problem of smoothness of solutions of parabolic partial differential equations. The mathematical framework for this access is given by measures on spaces of trajectories.

In the one-dimensional framework it is clear what is meant by smoothness of measures. We look for a direct analogy to the smoothness problem in infinite-dimensional spaces. For this purpose we start interpreting the Wiener space as a sequence space, to which the theory of differentiation and integration in Euclidean spaces is generalized by extension to infinite families of real numbers instead of finite ones.

The calculus possesses applications to many areas of stochastics, in particular finance stochastics, as is underpinned by the recently published book by Malliavin and Thalmayer. In this course I will report on recent applications to the theory of backward stochastic differential equations (BSDE), and their application to problems of the fine structure of option pricing and hedging in incomplete finance or insurance markets.

At first we want to present an access to the Wiener space as sequence space.

1 The Wiener space as sequence space

Definition 1.1 *A probability space* $(\Omega, \mathcal{F}, P)$ *is called* Gaussian *if there is a family* $(X_k)_{1\le k\le n}$ *or a sequence* $(X_k)_{k\in\mathbf{N}}$ *of independent Gaussian unit random variables such that*

$$\mathcal{F} = \sigma(X_k : 1 \le k \le n) \quad \textit{resp.} \quad \sigma(X_k : k \in \mathbf{N})$$

(completed by sets of P*-measure 0).*

Example 1:
Let $\Omega = C(\mathbf{R}_+, \mathbf{R}^m)$, $\mathcal{F}$ the Borel sets on Ω generated by the topology of uniform convergence on compact sets of $\mathbf{R}_+$, P the m-dimensional canonical Wiener measure on $\mathcal{F}$. Let further $W = (W^1, \cdots, W^m)$ be the canonical m-dimensional Wiener process defined by the projections on the coordinates.

Claim: $(\Omega, \mathcal{F}, P)$ is Gaussian.

Proof:

Let $(g_i)_{i\in\mathbf{N}}$ be an orthonormal basis of $L^2(\mathbf{R}_+)$,

$$W^j(g_i) = \int g_i(s) dW^j_s, \quad i \in \mathbf{N}, 1 \leq j \leq m,$$

in the sense of L^2-limits of Itô integrals. Then (modulo completion) we have

$$\mathcal{F} = \sigma(W_t : t \geq 0).$$

Let $t \geq 0$, $(a_i)_{i\in\mathbf{N}}$ a sequence in l^2 such that

$$1_{[0,t]} = \sum_{i\in\mathbf{N}} a_i \, g_i.$$

Then we have for $1 \leq j \leq m$

$$W^j_t = \lim_{n\to\infty} \sum_{i=1}^{n} a_i \, W^j(g_i) = \sum_{i=1}^{\infty} a_i W^j(g_i),$$

hence W^j_t is (modulo completion) measurable with respect to $\sigma(W^j(g_i) : i \in \mathbf{N})$. Therefore (modulo completion)

$$\mathcal{F} = \sigma(W^j(g_i) : i \in \mathbf{N}, 1 \leq j \leq m).$$

Moreover, due to

$$E(W^j(g_i) W^k(g_l)) = \delta_{jk} \langle g_i, g_l \rangle = \delta_{jk} \, \delta_{il}, \quad i, l \in \mathbf{N}, 1 \leq j, k \leq m,$$

hence the $W^j(g_i)$ are independent Gaussian unit variables. •

In the following we shall construct an abstract isomorphism between the canonical Wiener space and a sequence space. Since we are finally interested in infinite dimensional spaces, we assume from now on

Assumption: the Gaussian space considered is generated by *infinitely many independent Gaussian unit variables.*

Let $\mathbf{R}^\mathbf{N} = \{(x_i)_{i\in\mathbf{N}} : x_i \in \mathbf{R}, i \in \mathbf{N}\}$ be the set of all real-valued sequences, and for $n \in \mathbf{N}$ denote by

$$\pi_n : \mathbf{R}^\mathbf{N} \to \mathbf{R}^n, \quad (x_i)_{i\in\mathbf{N}} \mapsto (x_i)_{1\le i\le n},$$

the projection on the first n coordinates. Let $\mathbf{B}^n$ be the σ-algebra of Borel sets in $\mathbf{R}^n$,

$$\mathbf{B}^\mathbf{N} = \sigma(\cup_{n\in\mathbf{N}} \overset{-1}{\pi_n}[\mathbf{B}^n]).$$

Let for $n \in \mathbf{N}$

$$\nu_1(dx) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})\, dx, \quad \nu = P_{(X_n)_{n\in\mathbf{N}}} = \otimes_{i\in\mathbf{N}} \nu_1, \quad \nu_n = \nu \circ \overset{-1}{\pi_n}.$$

This notation is consistent for $n = 1$.

We want to construct an isomorphism between the spaces of integrable functions on $(\Omega, \mathcal{F}, P)$ and $(\mathbf{R}^\mathbf{N}, \mathbf{B}^\mathbf{N}, \nu)$. For this purpose, it is necessary to know how functions on the two spaces are mapped to each other. It is clear that for $\mathbf{B}^\mathbf{N}$-measurable f on $\mathbf{R}^\mathbf{N}$ we have

$$F = f \circ ((X_n)_{n\in\mathbf{N}})$$

is $\mathcal{F}$-measurable on Ω.

Lemma 1.1 *Let F be $\mathcal{F}$-measurable on Ω. Then there exists a $\mathbf{B}^\mathbf{N}$-measurable function f on $\mathbf{R}^\mathbf{N}$ such that*

$$F = f \circ ((X_n)_{n\in\mathbf{N}}).$$

Proof

1. Let $F = 1_A$ with $A = ((X_i)_{1\le i\le n})^{-1}[B], B \in \mathbf{B}^n$. Then set $f = 1_{\pi_n^{-1}[B]}$. f is by definition $\mathbf{B}^\mathbf{N}$-measurable and we have

$$f((X_n)_{n\in\mathbf{N}}) = 1_B((X_i)_{1\le i\le n}) = 1_A = F.$$

Hence the asserted equation is verified by indicators of a generating set of $\mathcal{F}$ which is stable for intersections. Hence by Dynkin's theorem it is valid for all indicators of sets in $\mathcal{F}$.

2. By part 1. and by linearity the claim is then verified for linear combinations of indicator functions of $\mathcal{F}$-measurable sets. The assertion is stable for monotone limits in the set of functions for which it is verified. Hence it is valid for all $\mathcal{F}$-measurable functions by the monotone class theorem. •

Theorem 1.1 *Let $p \geq 1$. Then the mapping*

$$L^p(\mathbf{R}^\mathbf{N}, \mathbf{B}^\mathbf{N}, \nu)) \ni f \mapsto F = f \circ ((X_n)_{n \in \mathbf{N}}) \in L^p(\Omega, \mathcal{F}, P))$$

defines a linear isomorphism.

Proof

The mapping is well defined due to

$$\begin{aligned} ||F||_p^p &= E(|f((X_n)_{n\in\mathbf{N}})|^p) \\ &= \int |f(x)|^p \nu(dx) \qquad \text{(transformation theorem)} \\ &= ||f||_p^p, \end{aligned}$$

and bijective due to Lemma 1.1. Linearity is trivial. •

Theorem 1.1 allows us to develop a differential calculus on the sequence space $(\mathbf{R}^\mathbf{N}, \mathbf{B}^\mathbf{N}, \nu)$, and then to transfer it to the canonical space $(\Omega, \mathcal{F}, P)$. For this purpose we are stimulated by the treatment of the one-dimensional situation.

Questions of smoothness of probability measures are prevalent. We start considering them in the setting of $\mathbf{R}$.

2 Absolute continuity of measures on R

Our aim is to study laws of random variables defined on $(\Omega, \mathcal{F}, P)$, i.e. the probability measures P_X for random variables X. By means of Theorem 1.1 these measures correspond to the measures $\nu \circ f^{-1}$ for $\mathbf{B}^\mathbf{N}$-measurable functions f on $\mathbf{R}^\mathbf{N}$. The one-dimensional version of these measures is given by $\nu_1 \circ f^{-1}$ for $\mathbf{B}^1$-measurable functions f defined on $\mathbf{R}$.

We first discuss a simple analytic criterion for absolute continuity of measures of this type.

Lemma 2.1 *Let μ be a finite measure on $\mathbf{B}^1$. Suppose there exists $c \in \mathbf{R}$ such that for all $\phi \in C^1(\mathbf{R})$ we have*

$$|\int \phi'(x) \mu(dx)| \leq c ||\phi||_\infty.$$

Then $\mu << \lambda$, i.e. μ is absolutely continuous with respect to λ, the Lebesgue measure on $\mathbf{R}$.

Proof

Let $0 \leq f$ be continuous with compact support, and define

$$\phi(x) = \int_{-\infty}^{x} f(y)dy,$$

Then

$$\int f d\mu = \int \phi^{'}(x)\mu(dx) \leq c||\phi||_\infty = c\int f d\lambda.$$

By a measure theoretic standard argument this inequality follows for bounded measurable f. Therefore we conclude for $A \in \mathbf{B}^1$

$$\mu(A) \leq c\,\lambda(A),$$

which clearly implies $\mu << \lambda$. •

We aim at applying Lemma 2.1 to the probability measure $\nu_1 \circ f^{-1}$ with f $\mathbf{B}^1$-measurable. For this purpose we encounter for the first time the central technique of *integration by parts* on Gaussian spaces, which is at the heart of Malliavin's calculus.

For reasons of notational clarity we first recall the classical technique of integration by parts. Indeed, for $g, h \in C_0^\infty(\mathbf{R})$ (smooth functions with compact support) we have

$$(*)\langle g', h\rangle = \int g'(x)\,h(x)\,dx = -\int g(x)\,h'(x)\,dx = -\langle g, h'\rangle.$$

This relationship can be extended to functions $g, h \in L^2(\mathbf{R})$ which vanish at $\pm\infty$ and which possess derivatives in the distributional sense. Let us for a moment assume this setting and denote by dg the *derivative in distributional sense* of g, and in the sense of the duality (*) with δh its *adjoint operator*. Then for $h \in C_0^\infty(\mathbf{R})$ we have

$$\delta h = -h',$$

and we can interpret the duality relationship as

$$\langle dg, h\rangle = \langle g, \delta h\rangle.$$

Finally, the operator $\delta d = -\frac{d^2}{dx^2}$ plays an important role in the calculus. Here it is identical to the negative of the Laplace operator. In the just sketched classical calculus one does not have to distinguish between d and δ (modulo sign).

For the analysis on Gaussian spaces things are different. We sketch the analogue of a differential calculus with respect to duality on Gaussian spaces. For $g, h \in L^2(\mathbf{R}, \nu_1)$ denote

$$\langle g|h\rangle = \int g(x)h(x)\,\nu_1(dx).$$

To apply Lemma 1.1 formally to the measure $\mu = \nu_1 \circ f^{-1}$ for some $\mathbf{B}^1$-measurable f, we have to write, assuming all operations are justified,

$$\int \phi'(x)\nu_1 \circ f^{-1}(dx) = \int \phi' \circ f(x)\nu_1(dx) = \langle \phi' \circ f|1\rangle = \langle (\phi \circ f)'|\frac{1}{f'}\rangle.$$

Now, as in the classical setting, we want to transfer the derivation to the other argument. For this purpose we continue calculating for $g, h \in C_0^\infty(\mathbf{R})$

$$\begin{aligned}
\langle g'|h\rangle &= \frac{1}{\sqrt{2\pi}}\int g'(x)\,h(x)\,\exp(-\frac{x^2}{2})\,dx \\
&= -\frac{1}{\sqrt{2\pi}}\int g(x)\,\frac{d}{dx}[h(x)\,\exp(-\frac{x^2}{2})]\,dx \\
&= -\int g(x)\,\exp(\frac{x^2}{2})\,\frac{d}{dx}[h(x)\,\exp(-\frac{x^2}{2})]\,\nu_1(dx) \\
&= \langle g| - h' + xh\rangle.
\end{aligned}$$

So in the setting of Gaussian spaces, if we define as before dg as distributional derivative in the generalized sense, its dual operator on a suitable space of functions (to be described later) has to be defined by

$$\delta h = -h' + xh.$$

In this sense we have the following duality relationship, completely analogously to the classical formula

$$\langle dg|h\rangle = \langle g|\delta h\rangle.$$

For the combination of the derivative operator and its dual we obtain this time the following operator

$$L = \delta d = -\frac{d^2}{dx^2} + x\frac{d}{dx},$$

in the suitable distributional sense.

d will be called *Malliavin derivative*, δ *Skorokhod integral*, and L *Ornstein-Uhlenbeck operator.* The domains of these operators will be more precisely defined in the higher dimensional setting. The present exposition is given for motivating the notions to be studied.

Let us return to the problem of smoothness of the measure $\nu_1 \circ f^{-1}$.

Lemma 2.2 *Let $g, h \in L^2(\mathbf{R}, \nu_1)$ be such that $dg, \delta h \in L^2(\mathbf{R}, \nu_1)$. Then we have*

$$\langle dg|h\rangle = \langle g|\delta h\rangle.$$

Moreover for $f \in L^2(\mathbf{R}, \nu_1)$ such that $\delta(\frac{1}{df}) \in L^2(\mathbf{R}, \nu_1)$ we have

$$\nu_1 \circ f^{-1} << \lambda.$$

Proof

We continue the above calculation in the notation chosen. We have by duality and the inequality of Cauchy-Schwarz

$$\begin{aligned} |\langle d(\phi \circ f)|\frac{1}{df}\rangle| &= |\langle \phi \circ f|\delta(\frac{1}{df})\rangle \\ &= |\int \phi \circ f(x)\, \delta(\frac{1}{df})(x)\, \nu_1(dx) \end{aligned}$$

$$\leq ||\phi||_\infty ||\delta(\frac{1}{df})||_2.$$

Hence Lemma 1.1 can be applied with $c = ||\delta(\frac{1}{df})||_2$ which yields the desired absolute continuity. •

With this lemma the program for the development of Gaussian differential calculus in finite and infinite dimensional spaces is sketched. We have to develop rigorously in this framework the calculus of the three operators. We shall hereby, for brevity, mostly concentrate on the operators d and δ. One natural orthonormal basis of $L^2(\mathbf{R}, \nu_1)$ proves to be very useful hereby.

3 Hermite polynomials; orthogonal developments

We continue denoting d, δ and L the operators studied above. They are at least (and this is the sense in which we use them) well defined on

$C_0^\infty(\mathbf{R})$, even, due to the integrability properties of the Gaussian density, on the space of polynomials in one real variable.

Definition 3.1 *For $n \geq 0$ let*

$$H_n = \delta^n 1.$$

H_n is called Hermite polynomial *of degree n.*

By definition we have for $x \in \mathbf{R}$

$$\begin{aligned} H_0(x) &= 1, \\ H_1(x) &= \delta 1 = x, \\ H_2(x) &= \delta x = -1 + x^2, \\ H_3(x) &= \delta(-1 + x^2) = -x - 2x + x^3 = x^3 - 3x. \end{aligned}$$

Theorem 3.1 *H_n is a polynomial of degree n, with leading coefficient 1. Moreover for $n \in \mathbf{N}$*

$$\begin{aligned} (i) &\quad \delta H_n = H_{n+1}, \\ (ii) &\quad dH_n = nH_{n-1}, \\ (iii) &\quad LH_n = n\, H_n. \end{aligned}$$

In particular, H_n is an eigenvector of L with eigenvalue n.

Proof
(i): follows by definition.

(ii): We first briefly investigate the commutator of d and δ. In fact, we have for $h \in C_0^\infty(\mathbf{R})$

$$(d\delta - \delta d)\, h = d(-h' + xh) - (-h'' + xh') = -h'' + xh' + h - (-h'' + xh') = h.$$

This means that

$$(d\delta - \delta d) = id.$$

With this in mind we proceed by induction on the degree n. The claim is clear for $n = 1$. Assume it holds for $n - 1$. Then

$$\begin{aligned} dH_n &= d\delta H_{n-1} = \delta d H_{n-1} + H_{n-1} \\ &= (n-1)\delta H_{n-2} + H_{n-1} = nH_{n-1}. \end{aligned}$$

(iii): $LH_n = \delta d H_n = n\delta H_{n-1} = nH_n$. •

Corollary 3.1 *For* $g \in L^2(\mathbf{R})$ *define the* Fourier transform *by*

$$\hat{g}(u) = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{iux} g(x)\, dx, \quad u \in \mathbf{R}.$$

Then

$$(\widehat{H_n\, e^{-\frac{x^2}{2}}})(u) = (iu)^n\, e^{-\frac{u^2}{2}}.$$

Proof

Choose $u \in \mathbf{R}$. Then

$$\begin{aligned}
(\widehat{H_n\, e^{-\frac{x^2}{2}}})(u) &= (\widehat{\delta^n 1\, e^{-\frac{x^2}{2}}})(u) \\
&= \langle e^{iu\cdot} | \delta^n 1 \rangle \\
&= \langle d^n e^{iu\cdot} | 1 \rangle \\
&= (iu)^n \langle e^{iu\cdot} | 1 \rangle \\
&= (iu)^n e^{-\frac{u^2}{2}}.
\end{aligned}$$

•

With these preliminaries, we can show that the Hermite polynomials constitute an orthonormal basis of our Gaussian space in one dimension.

Theorem 3.2 $(\frac{1}{\sqrt{n!}} H_n)_{n\geq 0}$ *is an orthonormal basis of* $L^2(\mathbf{R}, \nu_1)$.

Proof

1. Let $n, k \in \mathbf{N}$, and suppose that $n < k$. Then by Theorem 3.1

$$\langle H_n | H_k \rangle = \langle \delta^n 1 | \delta^k 1 \rangle = \langle d^k \delta^n 1 | 1 \rangle = 0,$$

while

$$\langle H_n | H_n \rangle = \langle d^n \delta^n 1 | 1 \rangle = n! \langle 1 | 1 \rangle = n!.$$

2. It remains to show that $(H_n)_{n\geq 0}$ is complete in $L^2(\mathbf{R}, \nu_1)$, i.e. the set of linear combinations of Hermite polynomials is dense in $L^2(\mathbf{R}, \nu_1)$. For this purpose, it suffices

to show: if $\phi \in L^2(\mathbf{R}, \nu_1)$ satisfies for all $n \geq 0$ we have $\langle H_n | \phi \rangle = 0$, then $\phi = 0$.

For $z \in \mathbf{C}$ let

$$F(z) = \int_{\mathbf{R}} \phi(v) e^{ivz - \frac{1}{2}v^2} dv.$$

Then we have for $k \in \mathbf{N}, t \in \mathbf{R}$ by Cauchy-Schwarz

$$| \int_{\mathbf{R}} |v^k \phi(v)| e^{-vt - \frac{1}{2}v^2} dv \leq [\int_{\mathbf{R}} \phi^2(v) e^{-\frac{1}{2}v^2} dv \int_{\mathbf{R}} v^{2k} e^{-2vt - \frac{1}{2}v^2} dv]^{\frac{1}{2}} < \infty.$$

Hence F may be differentiated arbitrarily often under the integral sign, which implies that F is an entire function. Moreover, we have for $k \geq 0$ with $x^k = \sum_{l=0}^k a_l H_l(x)$

$$\begin{aligned} F^{(k)}(0) &= i^k \int_{\mathbf{R}} v^k \phi(v) e^{-\frac{v^2}{2}} dv = i^k \langle x^k | \phi \rangle \\ &= i^k \sum_{l=0}^k a_l \langle H_l | \phi \rangle = 0. \end{aligned}$$

This, however, implies that $F = 0$, and so by the uniqueness of Fourier transforms also $\phi = 0$. •

We now return to our target space, namely $\mathbf{R}^{\mathbf{N}}$, the sequence space version of our infinite dimensional Gaussian space. Our task will be to establish in this space suitable notions of the operators d and δ. For this purpose it will be convenient to have again an orthonormal basis of this Gaussian space. We have to define an infinite dimensional extension of Hermite polynomials.

Definition 3.2 *For $n \in \mathbf{N}$ let $E_n = \mathbf{Z}_+^n$, let E be the set of sequences in $\mathbf{Z}_+$ that vanish except for finitely many components. For $p = (p_1, \cdots, p_k, 0, \cdots) \in E$ let $|p| = \sum_{i=1}^k p_i, p! = \prod_{i=1}^k p_i!$. For $x \in \mathbf{R}^k$ resp. $x \in \mathbf{R}^{\mathbf{N}}$, and $p \in E_k$ resp. $p \in E$ let*

$$\begin{aligned} H_p(x) &= \prod_{i=1}^k H_{p_i}(x_i) \quad \textit{resp.} \\ & \quad \prod_{i \in \mathbf{N}} H_{p_i}(x_i). \end{aligned}$$

*H_p is called k-*dimensional *resp.* generalized Hermite polynomial.

We can extend Theorem 3.2 to the multidimensional setting.

Theorem 3.3 *$(\frac{1}{\sqrt{p!}} H_p)_{p \in E_k}$ is an orthonormal basis of $L^2(\mathbf{R}^k, \mathbf{B}^k, \nu_k)$, $(\frac{1}{\sqrt{p!}} H_p)_{p \in E}$ is an orthonormal basis of $L^2(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu)$.*

Proof

1. For $k \in \mathbf{N}$ and $g, h \in L^2(\mathbf{R}^k, \nu_k)$ denote

$$\langle g | h \rangle = \int_{\mathbf{R}^k} g(x) h(x) \, \nu_k(dx).$$

Then for $p, q \in E_k$ we have, due to Fubini's theorem

$$\langle H_p | H_q \rangle = \prod_{i=1}^{k} \langle H_{p_i} | H_{q_i} \rangle.$$

Hence $(\frac{1}{\sqrt{p!}} H_p)_{p \in E_k}$ is an orthonormal system. Moreover, linear combinations of tensor products of functions of one of k variables are dense in $L^2(\mathbf{R}^k, \mathbf{B}^k, \nu_k)$. Hence the first claim follows from Theorem 3.2.

2. The set $\cup_{n \in \mathbf{N}} \pi_n^{-1}[L^2(\mathbf{R}^n, \mathbf{B}^n, \nu_n)]$ is dense in $L^2(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu)$. Hence, the second assertion follows from the first. •

We next define and study Sobolev spaces in finite and infinite dimension, for which we use our knowledge of the orthonormal bases just acquired.

4 Finite dimensional Gaussian Sobolev spaces

Let $k \in \mathbf{N}$. We consider k-dimensional spaces first. Before treating the Gaussian spaces, let us recall the most important facts about classical Sobolev spaces, i.e. Sobolev spaces with respect to Lebesgue measure on $\mathbf{R}^k$.

Definition 4.1 *Let $p \geq 1$. For $f \in L^p(\mathbf{R}^k), a \in \mathbf{R}^k$ we say that f* possesses a directional (generalized) derivative in direction a, *if there is a function $d_a f \in L^p(\mathbf{R}^k)$ such that*

$$||\frac{1}{\varepsilon}[f(\cdot + \varepsilon a) - f] - d_a f||_p \to 0$$

as $\varepsilon \to 0$. Let

$$W_1^p = \{f \in L^p(\mathbf{R}^k) : f \text{ possesses a directional derivative in direction } a \text{ for any } a \in \mathbf{R}^k\}$$

*(*Sobolev space of order $(1, p)$*).*

By linearity, it is clear that if for $1 \leq i \leq k$ we denote by e_i the ith canonical basis vector, then for $f \in W_1^p(\mathbf{R}^k)$ we have $d_a f = \sum_{i=1}^k a_i d_{e_i} f$, if $a = (a_1, \cdots, a_k)$. Let $d_i = d_{e_i}$ for a canonical basis $e_1, \cdots, e_k$ of $\mathbf{R}^k$.

Definition 4.2 *Let $p \geq 1, s \in \mathbf{N}$. We define recursively*

$$W_s^p = \{f \in W_1^p : d_a f \in W_{s-1}^p \ \ \textit{for any } a \in \mathbf{R}^k\}$$

*(*Sobolev space of order (s,p)*). For $f \in W_s^p$, $a_1, \cdots, a_s \in \mathbf{R}^k$ we define recursively*

$$d_{a_1} d_{a_2} \cdots d_{a_s} f.$$

We define the $(1,p)$-Sobolev norm *by*

$$||f||_{1,p} = ||f||_p + \sum_{i=1}^k ||d_i f||_p, \quad f \in W_1^p,$$

and analogous norms for higher derivative orders.

For any $p \geq 1, s \in \mathbf{N}$ we have

$$C_0^\infty(\mathbf{R}^k) \subset W_s^p$$

and for $g \in C_0^\infty(\mathbf{R}^k), a = (a_1, \cdots, a_k) \in \mathbf{R}^k$ we have

$$d_a g = \sum_{i=1}^k a_i \frac{\partial g}{\partial x_i}.$$

What is the relationship of our Sobolev spaces and the "weak derivatives" or "derivatives in distributional sense" encountered above?

Definition 4.3 *Let $f \in L^1_{loc}(\mathbf{R}^k), a \in \mathbf{R}^k$. Then $u_a \in L^1_{loc}(\mathbf{R}^k)$ is called* weak derivative *of f* in direction *a, if for any $\phi \in C_0^\infty(\mathbf{R}^k)$ the equation*

$$\langle f, d_a \phi \rangle = -\langle u_a, \phi \rangle$$

is satisfied.

Theorem 4.1 *Let $p \geq 1, f \in L^p(\mathbf{R}^k$. Then the following are equivalent:*
(i) $f \in W_1^p$,
(ii) for $a \in \mathbf{R}^k$ f possesses a weak derivative u_a, and we have $u_a \in L^p(\mathbf{R}^k)$.
In this case, moreover, $d_a f = u_a$.

Proof

1. Let us show that (i) implies (ii). For this purpose, assume $a \in \mathbf{R}^k, \phi \in C_0^\infty(\mathbf{R}^k)$. Then for $\varepsilon > 0$ by translational invariance of Lebesgue measure

$$\int \frac{1}{\varepsilon}[f(x+\varepsilon a) - f(x)]\phi(x)dx = \int \frac{1}{\varepsilon} f(x)[\phi(x-\varepsilon a) - \phi(x)]dx.$$

Now use (i), let $\varepsilon \to 0$, to identify the limit as $\int d_a f(x)\phi(x)dx$ on the left hand side, and as $-\int f(x) d_a\phi(x)dx$ on the right hand side. Hence for any $\phi \in C_0^\infty(\mathbf{R}^k)$

$$\langle d_a f, \phi\rangle = -\langle f, d_a\phi\rangle.$$

This means by definition that f possesses the weak derivative $d_a f$ which belongs to $L^p(\mathbf{R}^k)$.

2. Let us now prove that (ii) implies (i). Fix $\phi \in C_0^\infty(\mathbf{R}^k), a = (a_1, \cdots, a_k) \in \mathbf{R}^k$. Then by Taylor's formula with integral remainder term and Fubini's theorem we have for any $\varepsilon > 0$

$$\begin{aligned} & \int_{\mathbf{R}^k} \frac{1}{\varepsilon}[f(x+\varepsilon a) - f(x)]\phi(x)dx \\ = & \int_{\mathbf{R}^k} \frac{1}{\varepsilon} f(x)[\phi(x-\varepsilon a) - \phi(x)]dx \\ = & -\int_{\mathbf{R}^k} f(x)[\frac{1}{\varepsilon}\int_0^\varepsilon \sum_{i=1}^k a_i \frac{\partial\phi}{\partial x_i}(x-\xi a)d\xi]dx \\ = & -\frac{1}{\varepsilon}\int_0^\varepsilon [\int_{\mathbf{R}^k} \sum_{i=1}^k a_i \frac{\partial\phi}{\partial x_i}(x-\xi a) f(x)dx]d\xi \\ = & \frac{1}{\varepsilon}\int_0^\varepsilon [\int_{\mathbf{R}^k} \phi(x-\xi a)u_a(x)dx]d\xi \\ = & \frac{1}{\varepsilon}\int_0^\varepsilon [\int_{\mathbf{R}^k} \phi(x)u_a(x+\xi a)dx]d\xi \\ = & \int_{\mathbf{R}^k} [\frac{1}{\varepsilon}\int_0^\varepsilon u_a(x+\xi a)d\xi]\phi(x)dx. \end{aligned}$$

It remains to prove that $\frac{1}{\varepsilon}\int_0^\varepsilon u_a(\cdot + \xi a)d\xi$ converges to u_a in $L^p(\mathbf{R}^k)$. This is certainly true provided $u_a \in C_0^\infty(\mathbf{R}^k)$. But for any $f, g \in L^p(\mathbf{R}^k)$ we have uniformly in $\varepsilon > 0$

$$\begin{aligned} ||\frac{1}{\varepsilon}\int_0^\varepsilon f(\cdot+\xi a)d\xi \ - & \ \frac{1}{\varepsilon}\int_0^\varepsilon g(\cdot+\xi a)d\xi||_p \\ & \le \frac{1}{\varepsilon}\int_0^\varepsilon ||f(\cdot+\xi a) - g(\cdot+\xi a)||_p d\xi \\ = & \ ||g-f||_p. \end{aligned}$$

By means of this observation we can transfer the desired result from $C_0^\infty(\mathbf{R}^k)$ to $L^p(\mathbf{R}^k)$, since $C_0^\infty(\mathbf{R}^k)$ is dense in $L^p(\mathbf{R}^k)$. •

Corollary 4.1 *Let* $e_1, \cdots, e_k$ *denote the canonical basis of* $\mathbf{R}^k$, *let* $(f_n)_{n\in\mathbf{N}}$ *be a sequence in* W_1^p *such that*

(i) $||f_n - f||_p \to 0$ *as* $n \to \infty$,

(ii) for any $1 \le i \le k$ *the sequence* $(d_i f_n)_{n\in\mathbf{N}}$ *converges in* $L^p(\mathbf{R}^k)$.

Then $f \in W_1^p$ *and* $||f_n - f||_{1,p} \to 0$ *as* $n \to \infty$.

Proof

We have to show that f is weakly differentiable in direction e_i for $1 \le i \le k$, and $d_i f = \lim_{n\to\infty} d_i f_n \in L^p(\mathbf{R}^k)$. For this purpose let

$$u_i = \lim_{n\to\infty} d_i f_n,$$

which exists due to assumption (ii). Then by (i) for any $\phi \in C_0^\infty(\mathbf{R}^k), 1 \le i \le k$

$$\begin{aligned} \int f(x) d_i\phi(x)dx &= \lim_{n\to\infty} \int f_n(x) d_i\phi(x)dx \\ &= -\lim_{n\to\infty} \int d_i f_n(x)\phi(x)dx = -\int u_i(x)\phi(x)dx. \end{aligned}$$

This means that f possesses weak directional derivatives in direction e_i and $d_i f = u_i \in L^p(\mathbf{R}^k)$. Now Theorem 4.1 is applicable and finishes the proof. •

Corollary 4.2 *Let* $p \ge 1$. *Then* W_1^p *is a Banach space with respect to the norm* $||\cdot||_{1,p}$, *and for any* $a \in \mathbf{R}^k$ *the mapping* $d_a : W_1^p \to L^p(\mathbf{R}^k)$ *is continuous.*

Proof

We have to prove that W_1^p is complete with respect to $||\cdot||_{1,p}$. Let therefore $(f_n)_{n\in\mathbf{N}}$ be a Cauchy sequence in W_1^p. Then setting $f = \lim_{n\to\infty} f_n$ in $L^p(\mathbf{R}^k)$, we see that the hypotheses (i) and (ii) of Corollary 4.1 are satisfied, and it suffices to apply this Corollary. •

We finally need a local version of Sobolev spaces.

Definition 4.4 *For* $p \ge 1, s \in \mathbf{N}$ *let*

$$W_{s,loc}^p = \{f : f : \mathbf{R}^k \to \mathbf{R} \text{ measurable } f\phi \in W_s^p \text{ for } \phi \in C_0^\infty(\mathbf{R}^k).\}$$

*(*local Sobolev space of order (s, p)).

Theorem 4.2 *Let $p \geq 1, s \in \mathbf{N}$. Then $f \in W^p_{s,loc}$ iff for any $x_0 \in \mathbf{R}^k$ there exists an open neighborhood V_{x_0} of x_0 such that for any $\phi \in C_0^\infty(\mathbf{R}^k)$ with support in V_{x_0} we have $\phi f \in W^p_s$.*

Proof

We only need to prove the *only if* part of the claim. For any $x_0 \in \mathbf{R}^k$ let therefore V_{x_0} be given according to the statement of the assertion. Then $(V_{x_0})_{x_0 \in \mathbf{R}^k}$ is an open covering of $\mathbf{R}^k$. Then there exists a locally finite *partition of the unit* $(\phi_k)_{k \in \mathbf{N}} \subset C_0^\infty(\mathbf{R}^k)$ which is subordinate to the covering, i.e. such that

(i) $0 \leq \phi_n \leq 1$, for any $n \in \mathbf{N}$,

(ii) for any $n \in \mathbf{N}$ there exists $x_0(n)$ such that $\text{supp}(\phi_n) \subset V_{x_0(n)}$,

(iii) $\sum_{n \in \mathbf{N}} \phi_n = 1$,

(iv) for any compact set $K \subset \mathbf{R}^k$ the intersection of K and $\text{supp}(\phi_n)$ is non-empty for at most finitely many n.

Now let $\phi \in C_0^\infty(\mathbf{R}^k)$. Then for any $k \in \mathbf{N}$ (ii) gives $\text{supp}(\phi\phi_k) \subset V_{x_0(k)}$ and thus by assumption

$$\phi_k \phi f \in W^p_s, \quad k \in \mathbf{N}.$$

Since by (iv) the support of $\phi_k \phi$ is non-trivial for at most finitely many k, (iii) and linearity yield the desired

$$\phi f \in W^p_s.$$

•

We now turn to Gaussian Sobolev spaces. Our analysis will again be based on the differential operator we know from the above sketched classical calculus. Only the measure with respect to which we consider duality changes from the Lebesgue to the Gaussian measure. Since we thereby pass from an infinite to a finite measure, integrability properties for functions and therefore the domains of the dual operators change. This is why the notion of local Sobolev spaces is important. On these spaces, we can define our operators locally, without reference to integrability first. In fact, using Theorem 4.2, and for $s \in \mathbf{N}, p \geq 1, 1 \leq j_1, \cdots, j_s \leq k, f \in W^p_{s,loc}$ we can define

$$d_{j_1} d_{j_2} \cdots d_{j_s} f$$

locally on an open neighborhood V_{x_0} of an arbitrary point $x_0 \in \mathbf{R}^k$ by the corresponding generalized derivative of ϕf with $\phi \in C_0^\infty(\mathbf{R}^k)$ such that $\phi = 1$ on an open neighborhood $U_{x_0} \subset V_{x_0}$ of x_0. This gives a globally unique notion, since x_0 is arbitrary.

Definition 4.5 *Let $s \in \mathbf{N}, p \geq 1,\ 1 \leq j \leq k, f \in W^p_{s,loc}$, and denote by d_j the directional derivative in direction of the jth unit vector in $\mathbf{R}^k$ according to the preceding remark. Let then*

$$\nabla f = (d_1 f, \cdots, d_k f),$$

$$\delta_j f = -d_j f + x_j\, f,$$

$$Lf = \sum_{j=1}^k \delta_j d_j f = \sum_{j=1}^k [-d_j d_j f + x_j d_j f].$$

For any $1 \leq r \leq s$ we define more generally

$$\nabla^r f = (d_{j_1} d_{j_2} \cdots d_{j_r} f \,:\, 1 \leq j_1, j_2, \cdots, j_r \leq k).$$

This definition gives rise to the following notion of Gaussian Sobolev spaces.

Definition 4.6 *Let $p \geq 1, s \in \mathbf{N}$. Then let*

$$D^p_s(\mathbf{R}^k) = \{f \in W^p_{s,loc} : \sum_{r=0}^s ||\, |\nabla^r f|\, ||_p < \infty\},$$

$$||f||_{s,p} = \sum_{r=0}^s ||\, |\nabla^r f|\, ||_p$$

(k-dimensional Gaussian Sobolev space of order (s,p)).

Remark
$D^p_s(\mathbf{R}^k)$ is a Banach space. This is seen by arguments as for the proof of Corollary 4.2.

Since our calculus will be based mostly on the Hilbert case $p = 2$, we shall restrict our attention to this case whenever convenient. In this case, our ONB composed of k-dimensional Hermite polynomials as investigated in the previous chapter will play a central role, and adds structure to the setting. To get acquaintance with Gaussian Sobolev spaces, let us compute the operators defined on the series expansions with respect to this ONB.

For $f \in L^2(\mathbf{R}^k, \nu_k)$ we can write

$$f = \sum_{p \in E_k} \frac{c_p(f)}{p!} H_p$$

with coefficients $c_p(f) \in \mathbf{R}, p \in E_k$. Due to orthogonality, the Gaussian norm is given by

$$||f||_2 = \sum_{p \in E_k} \frac{c_p(f)^2}{p!^2} \langle H_p | H_p \rangle = \sum_{p \in E_k} \frac{c_p(f)^2}{p!}.$$

We also write $f \sim (c_p(f))$ to denote this series expansion. Denote by $\mathcal{P}$ the linear hull of the k-dimensional Hermite polynomials. Plainly, $\mathcal{P} \subset W^p_{s,loc}$ for any $s \in \mathbf{N}, p \geq 1$. According to chapter 3, $\mathcal{P}$ is dense in $L^2(\mathbf{R}^k, \nu_k)$. And for functions in $\mathcal{P}$, the generalized derivatives d_j are just identical to the usual partial derivatives in direction j, $1 \leq j \leq k$.

We first calculate the operators on Hermite polynomials. In fact, for $p \in E_k, 1 \leq j \leq k$ we have in the non-trivial cases

$$d_j H_p = p_j \prod_{i \neq j} H_{p_i} H_{p_j - 1}, \quad \delta_j H_p = \prod_{i \neq j} H_{p_i} H_{p_j + 1}, \quad L H_p = |p| H_p.$$

Hence for $f \sim (c_p(f)) \subset \mathcal{P}, 1 \leq j \leq k$ we may write

$$\begin{aligned} d_j f &= \sum_{p \in E_k} \frac{c_p(f)}{p!} p_j \prod_{i \neq j} H_{p_i} H_{p_j - 1}, \\ \delta_j f &= \sum_{p \in E_k} \frac{c_p(f)}{p!} \prod_{i \neq j} H_{p_i} H_{p_j + 1}, \\ L f &= \sum_{p \in E_k} \frac{c_p(f)}{p!} |p| H_p. \end{aligned}$$

According to Corollary 4.2 and the calculations just sketched, the natural domains of the operators extending ∇, δ_j and L beyond $\mathcal{P}$ must be those distributions in $\mathbf{R}^k$ for which the formulas just given generate convergent series in the L^2-norm with respect to ν_k. The most important domain is the one of ∇, the Sobolev space $D^2_1(\mathbf{R}^k)$. For $f \sim (c_p(f)) \in \mathcal{P}$

we have

$$
\begin{aligned}
||\,|\nabla f|\,||_2^2 &= \int_{\mathbf{R}^k} |\nabla f|^2(x)\nu_k(dx) \\
&= \sum_{j=1}^k \int_{\mathbf{R}^k} |d_j f|^2(x)\nu_k(dx) \\
&= \sum_{j=1}^k \sum_{p\in E_k} p_j^2 \frac{c_p(f)^2}{p!^2} \prod_{i\neq j} p_i!(p_j-1)! \\
&= \sum_{j=1}^k \sum_{p\in E_k} p_j \,\frac{c_p(f)^2}{p!} \\
&= \sum_{p\in E_k} |p| \frac{c_p(f)^2}{p!}.
\end{aligned}
$$

If in addition $f \in L^2(\mathbf{R}^k, \nu_k)$, we may write $f \sim (c_p(f))$ and approximate it by $f_n = \sum_{p\in E_k, |p|\leq n} \frac{c_p(f)}{p!} \in \mathcal{P}, n \in \mathbf{N}$. Hence, according to corollary 4.2, f belongs to $D_1^2(\mathbf{R}^k)$ if the following series converges

$$
\begin{aligned}
||\,|\nabla f|\,||_2^2 &= \lim_{n\to\infty} ||\,|\nabla f_n|\,||_2^2 \\
&= \lim_{n\to\infty} \sum_{p\in E_k, |p|\leq n} |p| \frac{c_p(f)^2}{p!} \\
&= \sum_{p\in E_k} |p| \frac{c_p(f)^2}{p!} < \infty.
\end{aligned}
$$

Along these lines, we now turn to describing Gaussian Sobolev spaces and the domains of our principal operators for $p = 2$ by means of Hermite expansions. We start with the case $k = 1$.

Theorem 4.3 *Let $r \in \mathbf{N}$, $f \sim (c_p(f)) \in L^2(\mathbf{R}, \nu_1) \cap W^2_{r,loc}$. Denote $f_p = \frac{c_p(f)}{p!} H_p, p \geq 0$. Then the following are equivalent:*
(i) $\nabla^r f \in L^2(\mathbf{R}, \nu_1)$,
(ii) $\sum_{p\geq 0} p^r ||f_p||_2^2 < \infty$,
(iii) $f \in D_r^2(\mathbf{R})$,
(iv) $\delta^r f \in L^2(\mathbf{R}, \nu_1)$.
In particular, $D_r^2(\mathbf{R})$ is the domain of ∇^r, δ^r in $L^2(\mathbf{R}, \nu_1)$. For $f, g \in D_1^2(\mathbf{R})$ we have

$$\langle \nabla f | g\rangle = \langle f | \delta g\rangle.$$

Proof

1. We prove equivalence of (i) and (ii). We have

$$\nabla f = \sum_{p\geq 1} \frac{p\, c_p(f)}{p!} H_{p-1} = \sum_{p\geq 0} \frac{c_{p+1}(f)}{p!} H_p,$$

and therefore by iteration

$$\nabla^r f = \sum_{p\geq 0} \frac{c_{p+r}(f)}{p!} H_{p-1}.$$

Therefore

$$||\nabla^r f||_2^2 = \sum_{p\geq 0} \frac{c_{p+r}(f)^2}{p!}, \quad ||f_p||_2^2 = \frac{c_p(f)^2}{p!},$$

and hence

$$||\nabla^r f||_2^2 = \sum_{p\geq 0} \frac{(p+r)!}{p!} ||f_{p+r}||_2^2 < \infty$$

if and only if

$$\sum_{p\geq 0} (p+r)^r ||f_{p+r}||_2^2 < \infty,$$

and this is the case if and only if

$$\sum_{p\geq 0} p^r ||f_p||_2^2 < \infty.$$

2. We next prove that (ii) and (iv) are equivalent. Note that

$$\delta f = \sum_{p\geq 0} \frac{c_p(f)}{p!} H_{p+1}, \quad \text{and therefore} \quad \delta^r f = \sum_{p\geq 0} \frac{c_p(f)}{p!} H_{p+r}.$$

This implies that

$$||\delta^r f||_2^2 = \sum_{p\geq 0} \frac{c_p(f)^2}{p!^2} (p+r)! = \sum_{p\geq 0} (p+r)\cdots(p+1) \frac{c_p(f)^2}{p!} < \infty$$

if and only if

$$\sum_{p\geq 0} p^r \frac{c_p(f)^2}{p!} = \sum_{p\geq 0} p^r ||f_p||_2^2 < \infty.$$

3. The equivalence of (i) and (iii) is contained in the definition.

4. Let $f \sim (c_p(f)), g \sim (c_p(g)) \in D_1^2(\mathbf{R})$. Then we have

$$\langle \nabla f | g \rangle = \sum_{p\geq 0} \frac{c_{p+1}(f)}{p!} \frac{c_p(g)}{p!} p!,$$

whereas

$$\langle f|\delta g\rangle = \sum_{p\geq 0} \frac{c_{p+1}(f)}{p!}\frac{c_p(g)}{p!}p!.$$

This completes the proof. •

The differential calculus on Gaussian spaces obeys similar rules as the classical differential calculus.

Theorem 4.4 *Let $g \in D_1^4(\mathbf{R}), \mu = \nu_1 \circ g^{-1}$. If $\phi \in L^4(\mathbf{R}, \mu)$, and $\nabla\phi \in L^4(\mathbf{R}, \mu)$, we have*

$$\nabla(\phi \circ g) = (\nabla\phi) \circ g \cdot \nabla g.$$

Proof

If ϕ and g belong to $C_0^\infty(\mathbf{R})$, the assertion is clear. To generalize, approximate in $\mathcal{P}$ and use Hölder's inequality. •

Theorem 4.5 *Let $f, g \in D_1^4(\mathbf{R})$. Then $f \cdot g \in D_1^2(\mathbf{R})$ and we have*

$$\nabla(f \cdot g) = f \cdot \nabla g + \nabla f \cdot g.$$

Proof

The assertion is clear for $f, g \in C_0^\infty(\mathbf{R})$. To generalize, approximate in $\mathcal{P}$ and use Hölder's inequality. •

We now turn to arbitrary finite dimension k, and interpret Gaussian Sobolev spaces by convergence properties of Hermite expansions as above.

Theorem 4.6 *Let $f \sim (c_p(f)) \in L^2(\mathbf{R}^k, \nu_k) \cap W^2_{1,loc}$. Denote $f_p = \frac{c_p(f)}{p!}H_p, p \in E_k$. Then the following are equivalent:*

(i) $|\nabla f| = [\sum_{j=1}^k (d_j f)^2]^{\frac{1}{2}} \in L^2(\mathbf{R}^k, \nu_k)$,

(ii) $\sum_{p\in E_k} |p| ||f_p||_2^2 < \infty$,

(iii) $f \in D_1^2(\mathbf{R}^k)$.

In particular, $D_1^2(\mathbf{R}^k)$ is the domain of ∇ in $L^2(\mathbf{R}^k, \nu_k)$. Analogous results hold for Sobolev spaces of order $(r, 2)$ with $r \in \mathbf{N}$.

Proof

Analogous to the proof of Theorem 4.2.•

5 Infinite dimensional Gaussian Sobolev spaces

To refer the infinite dimensional setting to the finite dimensional one, we use the following observation.

For $n \in \mathbf{N}$ recall

$$\pi^n : \mathbf{R}^{\mathbf{N}} \to \mathbf{R}^n, \quad (x_n)_{n\in\mathbf{N}} \mapsto (x_k)_{1\le k\le n}.$$

Let for $n \in \mathbf{N}$ let $\mathbf{C}_n = \sigma(\pi^n) = \sigma(\pi^1, \cdots, \pi^n)$. Then $(\mathbf{C}_n)_{n\in\mathbf{N}}$ is a filtration on $(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}})$.

Lemma 5.1 *Let $p \ge 1$, $f \in L^p(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu)$. Then*

$$\hat{f}_n = E(f|\mathbf{C}_n), \quad n \in \mathbf{N}$$

defines a martingale which converges ν-a.s. and in L^p to f.

Proof
This follows from a standard theorem of discrete martingale theory. •

Let in the following $f_n : \mathbf{R}^n \to \mathbf{R}$ the n-dimensional factorization of $\hat{f}_n$, related by

$$f_n \circ \pi^n = \hat{f}_n, n \in \mathbf{N}.$$

As a crucial observation for the definition of infinite dimensional Sobolev spaces, the martingale property is essentially not destroyed by the directional derivative operators.

Lemma 5.2 *Let $p > 1$, $f \in L^p(\mathbf{R}^{\mathbf{N}})$, $(f_n)_{n\in\mathbf{N}}$ the corresponding sequence according to the above remarks. Suppose that $\sup_{n\in\mathbf{N}} ||f_n||_{1,p} < \infty$. Then for any $j \in \mathbf{N}$ the sequence $(d_j f_n \circ \pi^n)_{n\in\mathbf{N}}$ converges in $L^p(\mathbf{R}^{\mathbf{N}})$, to a limit that we denote by $d_j f$. Corresponding statements hold true for higher order derivatives.*

Proof
Let $n \in \mathbf{N}, j \in \mathbf{N}$. Then for $n \ge j$ we have

$$E(d_j f_{n+1} \circ \pi^{n+1}|\mathbf{C}_n) = d_j f_n \circ \pi^n.$$

This means that $(d_j f_n \circ \pi^n)_{n\ge j}$ is a martingale with respect to $(\mathbf{C}_n)_{n\ge j}$ which, due to

$$\sup_{n\ge j} ||d_j f_n \circ \pi^n||_p \le \sup_{n\in\mathbf{N}} ||f_n||_{1,p} < \infty,$$

is bounded in $L^p(\mathbf{R^N})$ and hence converges in $L^p(\mathbf{R^N})$, due to $p > 1$. •

The preceding Lemmas give rise to the following definition of Sobolev spaces.

Definition 5.1 *Let* $p \geq 1, s \in \mathbf{N}$. *Then*

$$D_s^p(\mathbf{R^N}) = \{f \in L^p(\mathbf{R^N}, \nu) : f_n \in D_s^p(\mathbf{R}^n), n \in \mathbf{N}, \sup_{n \in \mathbf{N}} ||f_n||_{s,p} < \infty\},$$

*(*infinite dimensional Sobolev space of order (s,p)*), endowed with the norm*

$$||f||_{s,p} = \sup_{n \in \mathbf{N}} ||f_n||_{s,p}, \quad f \in D_s^p(\mathbf{R^N}).$$

This definition makes sense, for the following reasons.

Theorem 5.1 *Let* $p > 1, s \in \mathbf{N}$. *Then* $D_s^p(\mathbf{R^N})$ *is a Banach space with the norm* $||\cdot||_{s,p}$.

Proof

We prove the claim for $s = 1$. Let $(f^m)_{m \in \mathbf{N}}$ be a Cauchy sequence in $D_1^p(\mathbf{R^N})$, and $(f_n^m)_{n,m \in \mathbf{N}}$ the corresponding finite dimensional functions according to the remarks above. Then for $m, l \in \mathbf{N}$, $n \in \mathbf{N}$ Jensen's inequality and the martingale statement in the preceding proof give the following estimate

$$\limsup_{m,l \to \infty} ||f_n^m - f_n^l||_{1,p} \leq \lim_{m,l \to \infty} ||f^m - f^l||_{1,p} = 0.$$

$D_1^p(\mathbf{R}^n)$ being a Banach space for $n \in \mathbf{N}$, we know that

$$f_n = \lim_{m \to \infty} f_n^m \in D_1^p(\mathbf{R}^n)$$

exists. Let $\hat{f}_n = f_n \circ \pi^n$. Now let $f = \lim_{m \to \infty} f^m$ in $L^p(\mathbf{R^N})$. Then by uniform integrability

$$E(f|\mathbf{C}_n) = E(\lim_{m \to \infty} f^m|\mathbf{C}_n) = \lim_{m \to \infty} E(f^m|\mathbf{C}_n) = \lim_{m \to \infty} \hat{f}_n^m = \hat{f}_n.$$

Moreover

$$\sup_{n \in \mathbf{N}} ||f_n||_{1,p} \leq \sup_{m,n \in \mathbf{N}} ||f_n^m||_{1,p} \leq \sup_{m \in \mathbf{N}} ||f^m||_{1,p} < \infty.$$

Hence by definition $f \in D_1^p(\mathbf{R}^\mathbf{N})$, and by Fatou's lemma

$$||f - f^m||_{1,p} \leq \liminf_{l\to\infty} ||f^m - f^l||_{1,p} \to 0$$

as $m \to \infty$. •

According to Lemma 5.2, the gradient on the infinite dimensional Gaussian Sobolev spaces is defined as follows.

Definition 5.2 *Let* $p > 1$, $f \in D_1^p(\mathbf{R}^\mathbf{N})$. *Then let*

$$\nabla f = (d_j f)_{j\in\mathbf{N}}$$

*(*Malliavin gradient *or* Malliavin derivative*), where for any* $j \in \mathbf{N}$ *according to Lemma 5.2*

$$d_j f = \lim_{n\to\infty} d_j f_n \circ \pi^n.$$

Accordingly, for $s \in \mathbf{N}$ *we define* $\nabla^r f, 1 \leq r \leq s$, *for* $f \in D_s^p(\mathbf{R}^\mathbf{N})$.

Remark

The gradient ∇ being a continuous mapping from $D_1^p(\mathbf{R}^n)$ to $L^p(\mathbf{R}^n, \nu_n)$ for any finite dimension n, Lemma 5.2 and the definition of the Malliavin gradient imply, that ∇ is a continuous mapping from $D_1^p(\mathbf{R}^\mathbf{N})$ to $L^p(\mathbf{R}^\mathbf{N}, \nu)$.

Let us now again restrict our attention to $p = 2$ and describe Gaussian Sobolev spaces by means of the generalized Hermite polynomials. First of all, suppose $f = \sum_{p\in E} \frac{c_p(f)}{p!} \in L^2(\mathbf{R}^\mathbf{N}, \nu)$. We shall continue to use the notation $f \sim (c_p(f))$. Then for $n \in \mathbf{N}$, we have $f_n = \sum_{p\in E_n} \frac{c_{(p,0)}(f)}{p!} H_p$, where we put $(p, 0) = (p_1, \cdots, p_n, 0, 0, \cdots)$ for $p = (p_1, \cdots, p_n) \in E_n$. Therefore, we also have $\hat{f}_n = \sum_{p\in E_n} \frac{c_{(p,0)}(f)}{p!} H_{(p,0)}$. Let again $\mathcal{P}$ be the linear hull generated by all generalized Hermite polynomials.

As in the preceding chapter, we may calculate the gradient norms for $f \sim (c_p(f)) \in D_1^2(\mathbf{R}^\mathbf{N})$. In fact, we have for $j \in \mathbf{N}$

$$\begin{aligned} d_j f = \lim_{n\to\infty} d_j f_n \circ \pi^n &= \lim_{n\to\infty} \sum_{p\in E_n} \frac{c_{(p,0)}(f)}{p!} p_j \prod_{i\neq j} H_{p_i} H_{p_j - 1} \\ &= \sum_{p\in E} \frac{c_p(f)}{p!} p_j \prod_{i\neq j} H_{p_i} H_{p_j - 1}. \end{aligned}$$

Furthermore, for $f \in D_1^2(\mathbf{R}^{\mathbf{N}})$ let us compute the norm of $|\nabla f| = [\sum_{j\in\mathbf{N}}(d_j f)^2]^{\frac{1}{2}}$ in $L^2(\mathbf{R}^{\mathbf{N}}, \nu)$. In fact, we have, using the calculation of gradient norms in the preceding chapter,

$$\begin{aligned} \infty > \sup_{n\in\mathbf{N}} |||\nabla f_n \circ \pi^n|||_2^2 &= |||\nabla f|||_2^2 \\ &= \sup_{n\in\mathbf{N}} \sum_{p\in E_n} |p| \frac{c_{(p,0)}(f)^2}{p!} = \sum_{p\in E} |p| \frac{c_p(f)^2}{p!}. \end{aligned}$$

We therefore obtain the following main result about the description of the infinite dimensional Gaussian Sobolev spaces or order $(1, 2)$.

Theorem 5.2 *For $f \in L^2(\mathbf{R}^{\mathbf{N}}, \nu)$ the following are equivalent:*
(i) $f \in D_1^2(\mathbf{R}^{\mathbf{N}})$,
(ii) $\sum_{p\in E} |p| \frac{c_p(f)^2}{p!} < \infty$,
(iii) $|\nabla f_n| \circ \pi^n = [\sum_{j\in\mathbf{N}}(d_j f_n)^2 \circ \pi^n]^{\frac{1}{2}}$ converges in $L^2(\mathbf{R}^{\mathbf{N}}, \nu)$ to $|\nabla f|$.
Moreover, $D_1^2(\mathbf{R}^{\mathbf{N}})$ is a Hilbert space with respect to the scalar product

$$(f, g)_{1,2} = \langle f|g\rangle + \sum_{j\in\mathbf{N}} \langle d_j f|d_j g\rangle, \quad f, g \in D_1^2(\mathbf{R}^{\mathbf{N}}).$$

For $p \geq 2$ $\mathcal{P}$ is dense in $D_1^p(\mathbf{R}^{\mathbf{N}})$. Analogous results hold for Sobolev spaces of order $(s, 2)$ with $s \in \mathbf{N}$.

6 Absolute continuity in infinite dimensional Gaussian space

We are now in a position to discuss the main result of Malliavin's calculus in the framework of infinite dimensional Gaussian sequence spaces. The result is about the smoothness of laws of random variables defined on the Gaussian space. We start with a generalization of Lemma 1.1 to finite measures on $\mathbf{B}^d$ for $d \in \mathbf{N}$.

Lemma 6.1 *Let $\mu|\mathbf{B}^d$ be a finite measure. Assume there exists $c \in \mathbf{R}$ such that for all $\phi \in C^1(\mathbf{R}^d)$ with bounded partial derivatives, and any $1 \leq j \leq d$ we have*

$$|\int \frac{\partial}{\partial x_j} \phi(x)\, \mu(dx)| \leq c\, ||\phi||_\infty.$$

Then $\mu << \lambda^d$ (d-dimensional Lebesgue measure).

Proof
For simplicity, we argue for $d = 2$, and omit the superscript denoting 2-dimensional Lebesgue measure.

1. Assume that $\phi \in C^1(\mathbf{R}^2)$ possesses compact support. **We show:**

$$[\int |\phi|^2 d\lambda]^{\frac{1}{2}} \leq \frac{1}{2}[\int |\frac{\partial}{\partial x_1}\phi|^2 d\lambda + \int |\frac{\partial}{\partial x_2}\phi|^2 d\lambda].$$

In fact, we have

$$\begin{aligned}
[\int |\phi|^2 d\lambda]^{\frac{1}{2}} &\leq [\int \sup_{x_1 \in \mathbf{R}} |\phi(x_1, x_2)| dx_2 \int \sup_{x_2 \in \mathbf{R}} |\phi(x_1, x_2)| dx_1]^{\frac{1}{2}} \\
&\leq [\int |\frac{\partial}{\partial x_1}\phi(x_1, x_2)| dx_1\, dx_2 \int |\frac{\partial}{\partial x_2}\phi(x_1, x_2)| dx_2\, dx_1]^{\frac{1}{2}} \\
&\leq \frac{1}{2}[\int |\frac{\partial}{\partial x_1}\phi|^2 d\lambda + \int |\frac{\partial}{\partial x_2}\phi|^2 d\lambda].
\end{aligned}$$

2. Let $0 \leq u$ be continuous with compact support, and such that $\int u d\lambda = 1$, define for $\varepsilon > 0$ $u_\varepsilon = \frac{1}{\varepsilon^2} u(\frac{1}{\varepsilon}\cdot)$. Moreover, let

$$\psi_\varepsilon = \int u_\varepsilon(\cdot - y)\mu(dy)$$

be a smoothed version of μ. Then we obtain for h continuous with compact support, using Fubini's theorem,

$$\begin{aligned}
\int \psi_\varepsilon(x) h(x) dx &= \int [\int u_\varepsilon(x - y)\mu(dy)] h(x) dx \\
&= \int [\int u_\varepsilon(x - y) h(x) dx] \mu(dy) \\
&= \int [\int u(x) h(\varepsilon x + y) dx] \mu(dy) \\
&\to \int h(y)\mu(dy).
\end{aligned}$$

3. **We show:**

$$L^2(\mathbf{R}^2) \ni g \mapsto \int g d\mu \in \mathbf{R}$$

is a continuous linear functional.

In fact, let $\phi \in C^1(\mathbf{R}^2)$ have compact support, and let $\varepsilon > 0$. Then by hypothesis and smoothness of ψ_ε with a calculation as in 2.

$$\begin{aligned}
|\int \frac{\partial}{\partial x_i}\psi_\varepsilon(x)\phi(x) dx| &= |\int \psi_\varepsilon(x) \frac{\partial}{\partial x_i}\phi(x) dx| \\
&= |\int [\int u_\varepsilon(x - y) \frac{\partial}{\partial x_i}\phi(x) dx] \mu(dy)|
\end{aligned}$$

$$
\begin{aligned}
&= |\int[\int \frac{\partial}{\partial x_i} u_\varepsilon(x-y)\phi(x)dx]\mu(dy)| \\
&= |\int[\int \frac{\partial}{\partial y_i} u_\varepsilon(x-y)\phi(x)dx]\mu(dy)| \\
&\leq c||\int u_\varepsilon(x-\cdot)\phi(x)dx||_\infty \\
&\leq c||\phi||_\infty.
\end{aligned}
$$

Generalizing this inequality to bounded measurable ϕ, and then taking $\phi = \mathrm{sgn}(\psi_\varepsilon)$ yields the inequality

$$\int |\frac{\partial}{\partial x_i}\psi_\varepsilon| d\lambda \leq c$$

for any $\varepsilon > 0$. Now let $\varepsilon > 0, g \in L^2(\mathbf{R}^2)$ be given. Then, using 1. and the estimate above

$$
\begin{aligned}
|\int \psi_\varepsilon(x)g(x)dx| &\leq [\int |\psi_\varepsilon(x)|^2 dx \int |g(x)|^2 dx]^{\frac{1}{2}} \\
&\leq \frac{1}{2}[\int |\frac{\partial}{\partial x_1}\psi_\varepsilon| d\lambda + \int |\frac{\partial}{\partial x_2}\psi_\varepsilon| d\lambda]^{\frac{1}{2}} ||g||_2 \\
&\leq c||g||_2.
\end{aligned}
$$

Applying this inequality in the special case, in which g is continuous with compact support, and using 2. we get

$$|\int g(x)\mu(dx)| \leq c||g||_2.$$

Finally extend this inequality to $g \in L^2(\mathbf{R}^2)$ by approximating it with continuous functions of compact support. This yields the desired continuity of the linear functional.

4. It remains to apply Riesz' representation theorem to find a square integrable density for μ. •

We now consider a vector $f = (f^1, \cdots, f^d)$ with components in $L^2(\mathbf{R}^\mathbf{N}, \mathbf{B}^\mathbf{N}, \nu)$. Our aim is to study the absolute continuity with respect to λ^d of the law of f under ν, i.e. of the probability measure $\nu \circ f^{-1}$. For this purpose we plan to apply the criterion of Lemma 6.1. Let $\phi \in C^1(\mathbf{R}^d)$ possess bounded partial derivatives. Then, the integral transformation theorem gives

$$\int \frac{\partial}{\partial x_i}\phi d\nu \circ f^{-1} = \int \frac{\partial}{\partial x_i}\phi \circ f d\nu.$$

In case $d = 1$ at this place we use integration by parts hidden in the representations

$$d(\phi \circ f) = \phi'(f)df,$$
$$\phi'(f) = d(\phi \circ f)\frac{1}{df}.$$

Our infinite dimensional analogue of d is the Malliavin gradient ∇. Hence, we need a chain rule for ∇.

Theorem 6.1 *Let $p \geq 2$, $f \in D_1^p(\mathbf{R^N})^d, \phi \in C^1(\mathbf{R}^d)$ with bounded partial derivatives. Then*

$$\phi \circ f \in D_1^p(\mathbf{R^N}) \text{ and } \nabla[\phi \circ f] = \sum_{i=1}^d \frac{\partial}{\partial x_i}\phi(f) \cdot \nabla f^i.$$

Proof

Use Theorem 5.2 to choose a sequence $(f_n)_{n\in\mathbf{N}} \subset \mathcal{P}^d$ such that for any $1 \leq i \leq d$

$$||f_n^i - f^i||_{1,p} \to 0.$$

For each $n \in \mathbf{N}$ we have

$$\nabla[\phi \circ f_n] = \sum_{i=1}^d \frac{\partial}{\partial x_i}\phi(f_n) \cdot \nabla f_n^i.$$

Since ∇ is continuous on $D_1^p(\mathbf{R^N})$, and since the partial derivatives of ϕ are bounded, we furthermore obtain that

$$\nabla[\phi \circ f] = \lim_{n\to\infty} \nabla[\phi \circ f_n] = \sum_{i=1}^d \frac{\partial}{\partial x_i}\phi(f) \cdot \nabla f^i$$

in $L^2(\mathbf{R^N}, \nu)$. This completes the proof.•

We next present a calculation leading to the verification of the absolute continuity criterion of Lemma 6.1. We concentrate on the algebraic steps, and remark that their analytic background can be easily provided with the theory of chapter 5. The first aim of the calculations must be to isolate, for a given test function $\phi \in C^1(\mathbf{R}^d)$ with bounded partial derivatives, the expression $\frac{\partial}{\partial x_i}\phi(f), 1 \leq i \leq d$. Recall the notation

$$(x, y) = \sum_{i=1}^\infty x_i y_i, \quad x, y \in l^2.$$

For $1 \leq i, k \leq d$ let

$$\sigma_{ik} = (\nabla f^i, \nabla f^k).$$

Then we have for $1 \leq k \leq d$

$$\begin{aligned}(\nabla(\phi \circ f), \nabla f^k) &= \sum_{j=1}^{\infty} d_j(\phi \circ f) d_j f^k \\ &= \sum_{j=1}^{\infty} \sum_{1\leq i\leq d} \frac{\partial}{\partial x_i}\phi(f) d_j f^i d_j f^k \\ &= \sum_{1\leq i\leq d} \frac{\partial}{\partial x_i}\phi(f)\sigma_{ik}.\end{aligned}$$

We now assume that the matrix σ is (almost everywhere) invertible. Then, denoting its inverse by σ^{-1} we may write

$$\begin{aligned}\frac{\partial}{\partial x_i}\phi(f) &= \sum_{1\leq k\leq d} (\nabla(\phi \circ f), \nabla f^k \sigma^{-1}_{ki}) \\ &= \sum_{1\leq k\leq d} \sum_{j=1}^{\infty} d_j(\phi \circ f)\sigma^{-1}_{ki}\, d_j f^k.\end{aligned}$$

We next assume, that the dual operator δ_j of d_j, which is defined in the usual way on $\mathcal{P}$, is well defined and the series appearing is summable. Then we have

$$\begin{aligned}\int \frac{\partial}{\partial x_i}\phi(f) d\nu &= \int \sum_{1\leq k\leq d} \sum_{j=1}^{\infty} d_j(\phi \circ f)\sigma^{-1}_{ki}\, d_j f^k d\nu \\ &= \int \phi \circ f[\sum_{1\leq k\leq d} \sum_{j=1}^{\infty} \delta_j(d_j f^k \sigma^{-1}_{ki})] d\nu.\end{aligned}$$

The right hand side can be estimated by $c||\phi||_\infty$ with

$$c = ||\sum_{1\leq k\leq d} \sum_{j=1}^{\infty} \delta_j(d_j f^k \sigma^{-1}_{ki})||_2$$

in $L^2(\mathbf{R^N}, \nu)$. It can be seen (in analogy to Theorem 4.3) that this series makes sense under the hypotheses of the following main theorem.

Theorem 6.2 *Suppose that* $f = (f^1, \cdots, f^d) \in L^2(\mathbf{R^N}, \nu)$ *satisfies*

(i) $f^i \in D^4_2(\mathbf{R^N})$ *for* $1 \leq i \leq d$,

(ii) $\sigma_{ik} = (\nabla f^i, \nabla f^k)$, $1 \leq i, k \leq d$, *is* ν*-a.s. invertible and* $\sigma^{-1}_{ki} \in D^4_1(\mathbf{R^N})$ *for* $1 \leq i, k \leq d$.

Then we have $\nu \circ f^{-1} << \lambda^d$.

Proof
Approximate f by polynomials, use continuity properties of the operators.•

7 The canonical Wiener space: multiple integrals

We now return to the canonical Wiener space. The transfer between sequence and canonical space is provided by the isomorphism of chapter 1. We briefly recall it. Let $(g_i)_{i\in\mathbf{N}}$ be an orthonormal sequence in $L^2(\mathbf{R}_+)$. Then it is given by

$$T : L^p(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu) \to L^p(\Omega, \mathcal{F}, P), \quad f \mapsto f \circ ((W(g_i)_{i\in\mathbf{N}}),$$

where $W(g_i)$ is the Gaussian stochastic integral of g_i for $i \in \mathbf{N}$. It will be constructed in the following chapter. For simplicity we confine our attention to the canonical Wiener space in one dimension, i.e. $\Omega = C(\mathbf{R}_+, \mathbf{R})$, $\mathcal{F}$ the (completed) Borel σ-algebra on Ω generated by the topology of uniform convergence on compact sets in $\mathbf{R}_+$, P Wiener measure on $\mathcal{F}$.

In the approach of differential calculus on Gaussian sequence spaces, in the Hilbert space setting the most important tool proved to be the Hermite expansions of functions in $L^2(\mathbf{R}^{\mathbf{N}}, \nu)$. In the setting on the canonical space, they can be given a different interpretation which we shall now develop.

According to our isomorphism, the objects corresponding to generalized Hermite polynomials on the canonical space are given by

$$\prod_{i=1}^{\infty} H_{p_i}(W(g_i)),$$

$p \in E$. We shall interpret these objects as *iterated Itô integrals.* To do this, we use the abbreviation $\mathbf{B}^1_+$ for the Borel sets of $\mathbf{R}_+$.

Definition 7.1 *For* $m \in \mathbf{N}$ *let*

$$\mathcal{E}_m = \{f | f : \mathbf{R}^m_+ \to \mathbf{R}, f = \sum_{i_1,\cdots,i_m=1}^{n} a_{i_1\cdots i_m} 1_{A_{i_1}\times\cdots A_{i_m}},$$
$$(A_i)_{1\le i\le n} \subset \mathbf{B}^1_+ \text{ p.d.}, a_{i_1\cdots i_m} = 0 \text{ in case } i_k = i_l \text{ for some } k \neq l\}.$$

Remark
For $f \in L^2(\mathbf{R}_+)$ of the form

$$f = \sum_{i=1}^{n} a_i 1_{J_i}, \ (J_i)_{1\le i\le n} \text{ p.d. intervals in } \mathbf{R}_+$$

let

$$W(f) = \sum_{i=1}^{n} a_i W(J_i) = \sum_{i=1}^{n} a_i (W_{t_i} - W_{s_i}),$$

if $J_i =]s_i, t_i], 1 \leq i \leq n$. Then by Itô's isometry we have

$$||W(f)||_2^2 = ||f||_2^2.$$

Since $\mathcal{E}_1$ is dense in $L^2(\mathbf{R}_+)$, we can extend the linear mapping $f \mapsto W(f)$ to $L^2(\mathbf{R}_+)$. Therefore, in particular for $A \in \mathbf{B}_+^1$ with finite Lebesgue measure, $W(A) = W(1_A)$ is defined. It will be used for the definition of the following multiple stochastic integrals. In this chapter, the scalar product $\langle \cdot, \cdot \rangle$ will be with respect to Lebesgue measure on $\mathbf{R}_+^m$ with unspecified integer m.

Definition 7.2 *For* $f = \sum_{i_1,\cdots,i_m=1}^{n} a_{i_1\cdots i_m} 1_{A_{i_1} \times \cdots A_{i_m}} \in \mathcal{E}_m$ *let*

$$I_m(f) = \sum_{i_1,\cdots,i_m=1}^{n} a_{i_1\cdots i_m} W(A_{i_1}) \cdots W(A_{i_m}).$$

The additivity of $\mathbf{B}_+^1 \ni A \mapsto W(A) \in \mathbf{R}$ *implies that* I_m *is well defined.*

We state some elementary properties of I_m. Denote by $\mathcal{S}_m$ the set of all permutations of the numbers $1, \cdots, m$.

Lemma 7.1 *Let* $m, q \in \mathbf{N}, f \in \mathcal{E}_m, g \in \mathcal{E}_q$.
(i) $I_m | \mathcal{E}_m$ *is linear,*
(ii) if $\tilde{f}(t_1, \cdots, t_m) = \frac{1}{m!} \sum_{\sigma \in \mathcal{S}_m} f(t_{\sigma(1)}, \cdots, t_{\sigma(m)})$ *(*symmetrization of f*), then*

$$I_m(f) = I_m(\tilde{f}),$$

(iii) $E(I_m(f) I_q(g)) = m! \langle \tilde{f}, \tilde{g} \rangle$, *if* $m = q$, *and 0 otherwise.*

Proof

1. (i) follows from the additivity of the map $A \mapsto W(A)$.

2. (ii) is a direct consequence of the fact that in the definition of I_m the product $W(A_{i_1}) \cdots W(A_{i_m})$ is invariant under permutations of the factors.

3. By (ii), we may assume that f, g are symmetric. By choosing common subdivisions, we may further assume that

$$\begin{aligned} f &= \sum_{i_1,\cdots,i_m=1}^{n} a_{i_1\cdots i_m} 1_{A_{i_1}\times\cdots A_{i_m}}, \\ g &= \sum_{i_1,\cdots,i_q=1}^{n} b_{i_1\cdots i_q} 1_{A_{i_1}\times\cdots A_{i_q}}. \end{aligned}$$

Now if $m \neq q$, by the assumptions that $(A_i)_{1\leq i\leq n}$ consists of p.d. Borel sets, and that coefficients vanish if two of the indices coincide, $E(I_m(f)I_q(g)) = 0$ is evident. Assume $m = q$. Then, again by these two assumptions and symmetry

$$\begin{aligned} E(I_m(f)I_m(g)) &= \sum_{i_1,\cdots,i_m=1}^{n} \sum_{j_1,\cdots,j_m=1}^{n} a_{i_1\cdots i_m} b_{j_1\cdots j_m} E(\prod_{p=1}^{m} W(A_{i_p})W(A_{j_p})) \\ &= m!^2 \sum_{i_1<\cdots<i_m} \sum_{j_1<\cdots,j_m} a_{i_1\cdots i_m} b_{j_1\cdots j_m} E(\prod_{p=1}^{m} W(A_{i_p})W(A_{j_p})) \\ &= m!^2 \sum_{i_1<\cdots<i_m} a_{i_1\cdots i_m} b_{i_1\cdots i_m} \prod_{p=1}^{m} \lambda(A_{i_p}) \\ &= m!\langle f, g\rangle. \end{aligned}$$

•

To extend I_m beyond the space $\mathcal{E}_m$ of elementary functions, we proceed as for $m = 1$.

Lemma 7.2 *$\mathcal{E}_m$ is dense in $L^2(\mathbf{R}_+^m)$ for any $m \in \mathbf{N}$.*

Proof

We may assume $m \geq 2$, the assertion being known for $m = 1$. Due to standard results of measure theory, it is enough **to show:** for $A_1, \cdots, A_m \in \mathbf{B}_+^1$ with finite Lebesgue measure, and $\varepsilon > 0$, there exists $f \in \mathcal{E}_m$ such that

$$||1_{A_1\times\cdots\times A_m} - f||_2 < \varepsilon.$$

Let $\delta > 0$ to be determined later. Choose $B_1, \cdots, B_n \in \mathbf{B}_+^1$ with finite Lebesgue measure, pairwise disjoint, and such that for any $1 \leq j \leq n$ we have $\lambda(B_j) < \delta$, and such that any A_i can be represented by a finite union of some of the B_j. Then we have

$$1_{A_1\times\cdots\times A_m} = \sum_{i_1,\cdots,i_m=1}^{n} b_{i_1\cdots i_m} 1_{B_{i_1}\times\cdots\times B_{i_m}},$$

where $b_{i_1\cdots i_m} = 1$ if $B_{i_1} \times \cdots \times B_{i_m} \subset A_1 \times \cdots \times A_m$, and 0, if not. Let $I = \{(i_1, \cdots, i_m) : i_k \neq i_l \text{ for } k \neq l\}$, and $J = \{1, \cdots, n\}^m \setminus I$. Then by definition

$$f = \sum_{(i_1,\cdots,i_m)\in I} b_{i_1\cdots i_m} 1_{B_{i_1}\times\cdots\times B_{i_m}} \in \mathcal{E}_m$$

and we have

$$\begin{aligned}
||1_{A_1\times\cdots\times A_m} - f||_2^2 &= \sum_{(i_1,\cdots,i_m)\in J} b^2_{i_1\cdots i_m} \prod_{p=1}^{m} \lambda(B_{i_p}) \\
&\leq \frac{m(m-1)}{2} \sum_{i=1}^{n} \lambda(B_i)^2 (\sum_{i=1}^{n} \lambda(B_i))^{m-2} \\
&\leq \frac{m(m-1)}{2} \delta (\sum_{i=1}^{n} \lambda(B_i))^{m-1}
\end{aligned}$$

Finally, we have to choose δ small enough. •

Using Lemma 7.2, we may now extend I_m to $L^2(\mathbf{R}_+^m)$.

Definition 7.3 *The linear and continuous extension of* $I_m|\mathcal{E}_m$ *to* $L^2(\mathbf{R}_+^m)$ *which exists according to Lemma 7.2 is called* multiple Wiener-Itô integral *of degree* m *and also denoted by* I_m.

Properties of the elementary integral are transferred in a straightforward way.

Theorem 7.1 *Let* $m, q \in \mathbf{N}, f \in L^2(\mathbf{R}_+^m), g \in L^2(\mathbf{R}_+^q)$. *Then*
(i) $I_m|L^2(\mathbf{R}_+^m)$ *is linear,*
(ii) we have

$$I_m(f) = I_m(\tilde{f}),$$

(iii) $E(I_m(f)I_q(g)) = m!\langle \tilde{f}, \tilde{g}\rangle$, *if* $m = q$, *and 0 otherwise,*
(iv) $I_1(f) = W(f)$, $f \in L^2(\mathbf{R}_+)$.

Notation

We write

$$\begin{aligned}
I_m(f) &= \int_{\mathbf{R}_+^m} f(t_1, \cdots, t_m) dW_{t_1} \cdots dW_{t_m} \\
&= \int_{\mathbf{R}_+^m} f(t_1, \cdots, t_m) W(dt_1) \cdots W(dt_m).
\end{aligned}$$

We next aim at explaining the relationship between generalized Hermite polynomials and multiple Wiener-Itô integrals. For this purpose we will need a recursive relationship between Hermite polynomials of different degrees.

Remark

Recall the definition of Hermite polynomials in one variable, given by

$$H_n = \delta^n 1.$$

Moreover, we may compute for $n \in \mathbf{N}$

$$xH_n = (d + \delta)H_n = nH_{n-1} + H_{n+1}, \text{ or } H_{n+1} = xH_n - nH_{n-1}.$$

For technical reasons, we need the following *operation of contraction.*

Definition 7.4 *Let $m \in \mathbf{N}$. Suppose $f \in L^2(\mathbf{R}_+^m), g \in L^2(\mathbf{R}_+)$. Then for $t_1, \cdots, t_m, t \in \mathbf{R}_+$*

$$f \otimes g(t_1, \cdots, t_m, t) = f(t_1, \cdots, t_m) \cdot g(t) \text{ (tensor product)},$$

$$f \otimes_1 g(t_1, \cdots, t_{m-1}) = \int_{\mathbf{R}_+} \tilde{f}(t_1, \cdots, t_m) g(t_m)\, dt_m \text{ (contraction)}.$$

The recursion relation for Hermite polynomials will emerge from the recursion relation between Wiener-Itô integrals stated in the following Lemma.

Lemma 7.3 *Let $m \in \mathbf{N}, f \in L^2(\mathbf{R}_+^m), g \in L^2(\mathbf{R}_+)$. Then we have*

$$I_m(f)I_1(g) = I_{m+1}(f \otimes g) + mI_{m-1}(f \otimes_1 g).$$

Proof

1. By linearity and density of $\mathcal{E}_m$ in $L^2(\mathbf{R}_+^m)$ we may assume that

$$f = 1_{A_1 \times \cdots \times A_m}, \quad g = 1_{A_0} \text{ or } g = 1_{A_1},$$

where $(A_i)_{0 \le i \le m} \subset \mathbf{B}_+^1$ is a collection of p.d. Borel sets with finite Lebesgue measure.

2. The case $g = 1_{A_0}$ is trivial. Then the second term on the right hand side of the claimed formula vanishes, and the other two terms are obviously identical by definition of the elementary integral.

3. Let now $g = 1_{A_1}$. For $\varepsilon > 0$ choose a collection of p.d. sets $B_1, \cdots, B_n \in \mathbf{B}^1_+$ such that $A_1 = \cup_{i=1}^n B_i$, and for any $1 \leq i \leq n$ we have $\lambda(B_i) < \varepsilon$. Then

$$\begin{aligned} I_m(f)I_1(g) &= W(A_1)^2 W(A_2)\cdots W(A_m) \\ &= \sum_{i\neq j} W(B_i)W(B_j)W(A_2)\cdots W(A_m) \\ &+ \sum_{1\leq i\leq n} [W(B_i)^2 - \lambda(B_i)]W(A_2)\cdots W(A_m) \\ &+ \lambda(A_1)W(A_2)\cdots W(A_m). \end{aligned}$$

a) We now prove that the first term on the right hand side of our formula is close to $I_{m+1}(f \otimes g)$. In fact, let

$$h_\varepsilon = \sum_{i\neq j} 1_{B_i\times B_j\times A_2\times\cdots\times A_m} \in \mathcal{E}_{m+1}.$$

Then

$$\begin{aligned} ||h_\varepsilon - f\otimes g||_2^2 &\leq \sum_{i=1}^n \lambda(B_i)^2\lambda(A_2)\cdots\lambda(A_m) \\ &\leq \varepsilon\lambda(A_1)\cdots\lambda(A_m). \end{aligned}$$

b) Let us next prove that the second term on the right hand side is negligeable in the limit $\varepsilon \to 0$. In fact, denote it by R_ε. Then, since for $1 \leq i \leq n$ the variance of $W^2(B_i) - \lambda(B_i)$ is given by $c\lambda(B_i)^2$ with some constant c, we obtain

$$E(R_\varepsilon^2) \leq c\sum_{i=1}^n \lambda(B_i)^2\lambda(A_2)\cdots\lambda(A_m) \leq \epsilon\, c\lambda(A_1)\lambda(A_2)\cdots\lambda(A_m).$$

c) To evaluate the last term, note that

$$1_{\widetilde{A_1\times\cdots\times A_m}} \otimes_1 1_{A_1} = \frac{1}{m} 1_{\widetilde{A_2\times\cdots\times A_m}} \cdot \lambda(A_1).$$

Therefore

$$\lambda(A_1)W(A_2)\cdots W(A_m) = mI_{m-1}(1_{\widetilde{A_1\times\cdots\times A_m}} \otimes_1 1_{A_1}),$$

and we obtain the desired recursion formula. •

This finally puts us in a position to derive the relationship between Hermite polynomials and iterated stochastic integrals.

Theorem 7.2 *Let $m \in \mathbf{N}$, $h \in L^2(\mathbf{R}_+)$ be such that $||h||_2 = 1$. Denote by $h^{\otimes m}$ the m-fold tensor product of h with itself. Then we have*

$$H_m(W(h)) = I_m(h^{\otimes m}).$$

Let $\mathcal{H}_0 = \mathbf{R}$, $\mathcal{H}_m = I_m(L^2(\mathbf{R}_+^m)), m \in \mathbf{N}$. Then: $(\mathcal{H}_m)_{m \in \mathbf{N}}$ is a sequence of pairwise orthogonal closed linear subspaces of $L^2(\Omega, \mathcal{F}, P)$ and we have

$$L^2(\Omega, \mathcal{F}, P) = \oplus_{m=0}^{\infty} \mathcal{H}_m.$$

In particular, for any $F \in L^2(\Omega, \mathcal{F}, P)$ there exists a sequence $(f_m)_{m \geq 0}$ of functions $f_m \in L^2(\mathbf{R}_+^m)$ such that

$$F = \sum_{m=0}^{\infty} I_m(f_m).$$

The representation with symmetric f_m is λ^m-a.e. unique, $m \in \mathbf{N}$.

Proof

1. We first have **to prove:**

$$H_m(W(h)) = I_m(h^{\otimes m}).$$

This is done by induction on m. For $m = 1$, the formula is clear from $H_1 = x, I_1(h) = W(h)$. Now assume it is known for m. Then Lemma 7.3 and the recursion formula for Hermite polynomials given above combine to yield, remembering that $||h||_2 = 1$,

$$\begin{aligned} I_{m+1}(h^{\otimes m+1}) &= I_m(h^{\otimes m}) I_1(h) - m I_{m-1}(h^{\otimes m} \otimes_1 h) \\ &= I_m(h^{\otimes m}) I_1(h) - m I_{m-1}(h^{\otimes m-1}) \\ &= H_m(W(h)) H_1(W(h)) - m H_{m-1}(W(h)) \\ &= H_{m+1}(W(h)). \end{aligned}$$

2. Let $L_s^2(\mathbf{R}_+^m)$ be the linear space of symmetric functions in $L^2(\mathbf{R}_+^m)$. Then by Theorem 7.1

$$||I_m(\tilde{f})||_2^2 = m!||\tilde{f}||_2^2,$$

hence $\mathcal{H}_m = I_m(L_s^2(\mathbf{R}_+^m))$ is closed. Orthogonality is also a consequence of Theorem 7.1.

3. Let $(g_i)_{i\in\mathbf{N}}$ be an orthonormal basis of $L^2(\mathbf{R}_+)$, $F \in L^2(\Omega, \mathcal{F}, P)$. Let $f \in L^2(\mathbf{R}^\mathbf{N}, \mathbf{B}^\mathbf{N}, \nu)$ be such that $T(f) = F$. Assume $f \sim (c_p(f))$, according to the notation of chapter 3. For $m \geq 0$, define

$$f_m = \sum_{p\in E, |p|=m} \frac{c_p(f)}{p!} \prod_{i\in\mathbf{N}} g_i^{\otimes p_i},$$

considered as a function of m variables. Then $f_m \in L^2(\mathbf{R}_+^m)$ and with the help of Lemma 7.3 we see

$$\begin{aligned} I_m(f_m) &= \sum_{p\in E, |p|=m} \frac{c_p(f)}{p!} I_m(\prod_{i\in\mathbf{N}} g_i^{\otimes p_i}) \\ &= \sum_{p\in E, |p|=m} \frac{c_p(f)}{p!} \prod_{i\in\mathbf{N}} I_{p_i}(g_i^{\otimes p_i}) \\ &= \sum_{p\in E, |p|=m} \frac{c_p(f)}{p!} \prod_{i\in\mathbf{N}} H_{p_i}(W(g_i)). \end{aligned}$$

Summing this expression over m yields the desired

$$F = \sum_{m=0}^{\infty} I_m(f_m).$$

The remaining claims are obvious.•

8 The canonical Wiener space: Malliavin's derivative

In this chapter we shall investigate the analogue of the gradient we encountered in the differential calculus on the sequence space. Fix again an orthonormal basis $(g_i)_{i\in\mathbf{N}}$ of $L^2(\mathbf{R}_+)$, and recall the isomorphism

$$T : L^2(\mathbf{R}^\mathbf{N}, \mathbf{B}^\mathbf{N}, \nu) \to L^2(\Omega, \mathcal{F}, P), \quad f \mapsto f((W(g_i)_{i\in\mathbf{N}}).$$

Of course, every permutation of the orthonormal basis functions gives another orthonormal basis. So here we encounter a problem of coordinate dependence of our objects of study. How can we define Malliavin's derivative on the canonical space in a both consistent and basis independent way? According to Theorem 5.2

$$\nabla f = (d_j f)_{j\in\mathbf{N}}$$

takes values in l^2. The corresponding object on the side of the canonical space is $L^2(\mathbf{R}_+)$. It is therefore plausible if we set

Definition 8.1 *For $n \in \mathbf{N}$ let $C_p^\infty(\mathbf{R}^n)$ denote the set of smooth functions the partial derivatives of which possess polynomial growth. Let*

$$\begin{aligned}\mathcal{S} &= \{F|F = f(W(h_1), \cdots, W(h_n)),\ h_1, \cdots, h_n \in L^2(\mathbf{R}_+), \\ &\quad f \in C_p^\infty(\mathbf{R}^n), n \in \mathbf{N}\}.\end{aligned}$$

For $F = f(W(h_1), \cdots, W(h_n)) \in \mathcal{S}, t \geq 0$ let

$$D_t F = \sum_{i=1}^n \frac{\partial}{\partial x_i} f(W(h_1), \cdots, W(h_n))\, h_i(t).$$

To see if this is a good candidate for the definition of the Malliavin gradient in the setting of the canonical space, let us verify in detail the independence on the specific representation of functionals. Let $h_1, \cdots, h_n \in L^2(\mathbf{R}_+)$, and $g_1, \cdots, g_m \in L^2(\mathbf{R}_+)$ orthonormal, such that the linear hulls of the two systems are identical, and such that with $f \in C_p^\infty(\mathbf{R}^n), g \in C_p^\infty(\mathbf{R}^m)$ we have

$$f(W(h_1), \cdots, W(h_n)) = g(W(g_1), \cdots, W(g_m)).$$

For $1 \leq i \leq n$ write

$$h_i = \sum_{j=1}^m \langle h_i, g_j \rangle g_j.$$

Then, denoting

$$\Gamma = (\langle h_i, g_j \rangle)_{1 \leq i \leq n, 1 \leq j \leq m},$$

we obviously have $f \circ \Gamma = g$. Therefore

$$\begin{aligned}\sum_{j=1}^m \frac{\partial}{\partial x_j} (f \circ \Gamma)(W(g_1), \cdots, W(g_m)) g_j &= \sum_{j=1}^m \sum_{i=1}^n \frac{\partial}{\partial x_i} f(W(h_1), \cdots, W(h_n)) \\ &\quad \cdot \langle h_i, g_j \rangle g_j \\ &= \sum_{i=1}^n \frac{\partial}{\partial x_i} f(W(h_1), \cdots, W(h_n)) h_i.\end{aligned}$$

This proves that the definition of D is independent of the representation of functionals in $\mathcal{S}$.

If $h \in L^2(\mathbf{R}_+)$ is another function, we have by definition

$$\langle D.F, h \rangle = \sum_{i=1}^m \frac{\partial}{\partial x_i} f(W(g_1), \cdots, W(g_m)) \langle g_i, h \rangle,$$

in particular for $i \in \mathbf{N}$

$$\langle D.F, g_i \rangle = \frac{\partial}{\partial x_i} f(W(g_1), \cdots, W(g_m)) = d_i f(W(g_1), \cdots, W(g_m)).$$

We therefore may interpret $\langle D.F, g_i \rangle$ as directional derivative in direction of g_i, and we have by Parseval's identity

$$\begin{aligned} \langle DF, DF \rangle = \sum_{i=1}^m \langle DF, g_i \rangle^2 &= \sum_{i=1}^m (d_i f)^2 (W(g_1), \cdots, W(g_m)) \\ &= |\nabla f|^2 (W(g_1), \cdots, W(g_m)). \end{aligned}$$

Analogously, higher derivatives are related to each other. So we see that the isomorphism T also maps $\langle DF, DF \rangle$ to $|\nabla f|^2$. Consequently, we can just transfer the definitions of Gaussian Sobolev spaces to the setting of the canonical Wiener space.

Definition 8.2 *Let $p \geq 2, s \in \mathbf{N}$. For $F \in L^2(\Omega, \mathcal{F}, P)$ denote by $f \in L^2(\mathbf{R}^\mathbf{N}, \mathbf{B}^\mathbf{N}, \nu)$ the function for which we have $F = f((W(g_i)_{i \in \mathbf{N}})$. Then let*

$$D_s^p = \{F | f \in D_s^p(\mathbf{R}^\mathbf{N})\}$$

*(*canonical Gaussian Sobolev space of order (s,p)*), with the norm*

$$||F||_{s,p} = ||f||_{s,p}.$$

For $F = T(f) \in D_s^p$, $1 \leq r \leq s$, let

$$D^r F = \sum_{j_1, \cdots, j_r = 1}^{\infty} d_{j_1} \cdots d_{j_r} f(W(g_i)_{i \in \mathbf{N}}) g_{j_1} \otimes \cdots \otimes g_{j_r}$$

*(*canonical Malliavin derivative of order r*).*

Remark

From our knowledge of sequence spaces we can easily derive that D_s^p is a Banach space with respect to the norm $|| \cdot ||_{s,p}$ for $p \geq 2, s \in \mathbf{N}$, and that for $F \in D_s^p$ we have

$$||F||_{s,p} = \sum_{r=0}^s ||\langle D^r F, D^r F \rangle||_p.$$

We know that D_s^p is the closure of $\mathcal{S}$ with respect to the norm $||\cdot||_{s,p}$.

Turning to $p=2$, we know that D_1^2 is a Hilbert space with respect to the scalar product

$$(F,G)_{1,2} = E(FG) + E(\langle DF, DG\rangle), \quad F,G \in D_1^2.$$

Moreover, we know that D is a closed operator, defined on D_1^2, which is continuous as a mapping from D_1^2 to $L^2(\Omega,\mathcal{F},P)$.

Let us now investigate how D operates on the decomposition into Wiener-Itô integrals.

Theorem 8.1 *Let $F=\sum_{m=0}^\infty I_m(f_m) \in L^2(\Omega,\mathcal{F},P)$ be given, f_m symmetric for any $m\ge 0$. Then we have*

$$F\in D_1^2 \text{ if and only if } \sum_{m=1}^\infty m\, m!\, ||f_m||_2^2 < \infty.$$

In this case we have

$$D_tF = \sum_{m=1}^\infty m\, I_{m-1}(f_m(\cdot,t))$$

(for $P\otimes\lambda$-a.e. $(\omega,t)\in\Omega\times\mathbf{R}_+$).

Proof

1. Suppose that with respect to an orthonormal basis $(g_i)_{i\in\mathbf{N}}$ of $L^2(\mathbf{R}_+)$ we have $F=T(f)$ with $f\sim(c_p(f))$. As before, for $m\ge 0$ let

$$f_m = \sum_{p\in E,|p|=m} \frac{c_p(f)}{p!} \prod_{i\in\mathbf{N}} g_i^{\otimes p_i}.$$

We interpret $\prod_{i\in\mathbf{N}} g_i^{\otimes p_i}$ as $\prod_{j=1}^k g_{i_j}^{\otimes p_{i_j}}$, if $i_1,\cdots,i_k$ are precisely those indices for which $p_{i_1},\cdots,p_{i_k}>0$.

2. Let now $p\in E$ such that $|p|=m$, and let $t\ge 0$. Then

$$\begin{aligned} D_tI_m(\prod_{i\in\mathbf{N}} g_i^{\otimes p_i}) &= D_t\prod_{i\in\mathbf{N}} I_{p_i}(g_i^{\otimes p_i}) \\ &= D_tH_p((W(g_i)_{i\in\mathbf{N}}) \\ &= \sum_{i\in\mathbf{N}} p_i \prod_{j\neq i} H_{p_j}(W(g_j))H_{p_i-1}(W(g_i))g_i(t) \\ &= I_{m-1}(\sum_{i\in\mathbf{N}} p_i \prod_{j\neq i} g_j^{\otimes p_j} g_i^{\otimes p_i-1} g_i(t)). \end{aligned}$$

Hence by closedness of D, symmetry of f_m and $|p| = m$, the desired formula

$$D_t I_m(f_m) = m I_{m-1}(f_m(\cdot, t))$$

follows.

3. For $n \in \mathbf{N}$ let now

$$F_n = \sum_{m=0}^{n} I_m(f_m).$$

By the closedness of the operator D and the remarks above, we know that

$$F \in D_1^2 \text{ if and only if } (F_n)_{n\in\mathbf{N}} \text{ is Cauchy in } D_1^2.$$

Now we know from the first part of the proof that

$$DF_n = \sum_{m=1}^{n} m\, I_{m-1}(f_m(\cdot, \cdot)).$$

Let $n, m \in \mathbf{N}, n \geq m$ be given. Then

$$\begin{aligned} E(\langle D(F_n - F_m), D(F_n - F_m)\rangle) &= \sum_{k=m+1}^{n} k^2 \int_{\mathbf{R}_+} (k-1)! \langle f_k(\cdot, t), f_k(\cdot, t)\rangle dt \\ &= \sum_{k=m+1}^{n} k^2 (k-1)! ||f_k||_2^2. \end{aligned}$$

Hence $(DF_n)_{n\in\mathbf{N}}$ is a Cauchy sequence in D_1^2 if and only if $\sum_{k=0}^{\infty} k k! ||f_k||_2^2 < \infty$. In this case, the series with the desired representation converges. •

We need some rules to be able to calculate with the Malliavin gradient D.

Theorem 8.2 *Let $p \geq 2, d \in \mathbf{N}$, $\phi \in C^1(\mathbf{R}^d)$ with bounded partial derivatives, let $F = (F^1, \cdots, F^d) \in (D_1^p)^d$. Then $\phi \circ F \in D_1^p$ and we have*

$$D\phi \circ F = \sum_{i=1}^{d} \frac{\partial}{\partial x_i} \phi(F) DF^i.$$

Proof

The proof of Theorem 6.1 translates. •

With the following properties we prepare a study of the dual operator of D.

Theorem 8.3 *Let $F \in \mathcal{S}, h \in L^2(\mathbf{R}_+)$. Then we have*

$$E(\langle DF, h\rangle) = E(FW(h)).$$

Proof

We may assume that $F = f(W(g_1), \cdots, W(g_n)), h = g_1$ with respect to an orthonormal system $g_1, \cdots, g_n \in L^2(\mathbf{R}_+)$. In this case we have by duality of d_1 and δ_1

$$\begin{aligned} E(\langle DF, h\rangle) &= E(\frac{\partial}{\partial x_1} f((W(g_1), \cdots, W(g_n))) \\ &= \langle \nabla f | (1, 0, \cdots, 0)\rangle \\ &= \langle f | \delta_1 1\rangle = \langle f | H_{(1,0,\cdots)}\rangle \\ &= E(f(W(g_1), \cdots, W(g_n))W(g_1)) \\ &= E(FW(h)). \end{aligned}$$

This completes the proof. •

Theorem 8.4 *Let $F, G \in \mathcal{S}, h \in L^2(\mathbf{R}_+)$. Then we have*

$$E(G\langle DF, h\rangle) = E(FGW(h) - F\langle DG, h\rangle).$$

Proof

Apply Theorem 8.3 to the function FG. •

9 The canonical Wiener space: Skorokhod's integral

In this chapter we dedicate a more careful study to the dual operator (in the sense of Hilbert space theory) of the Malliavin gradient than in the Gaussian sequence spaces. In the setting of the canonical Wiener space, this operator turns out to be a stochastic integral.

So far we know that

$$D : D_1^2 \to L^2(\Omega \times \mathbf{R}_+)$$

is densely defined and linear.

Definition 9.1 *Let*

$$dom(\delta) = \{u \in L^2(\Omega \times \mathbf{R}_+) \ : \ \text{there is } c \in \mathbf{R} \text{ such that for any } F \in D_1^2 \text{ we have } E(\langle DF, u\rangle) \leq c||F||_2\}.$$

For $u \in \mathrm{dom}(\delta)$ the mapping $F \mapsto E(\langle DF, u\rangle)$ can be extended to a continuous linear functional. Hence by Riesz' representation we may find $\delta(u) \in L^2(\Omega)$ such that

$$E(\langle DF, u\rangle) = E(F \cdot \delta(u)), \quad F \in D_1^2.$$

Since D is densely defined, $\delta(u)$ is unique for any $u \in \mathrm{dom}(\delta)$.

Definition 9.2 *For $u \in dom(\delta)$ the uniquely determined random variable $\delta(u) \in L^2(\Omega)$ is called* Skorokhod integral *of u.*

Notation
We write $\delta(u) = \int_{\mathbf{R}_+} u_t \delta W_t$.

Why is this operator called *integral*? To answer this question, we first ask how elementary processes are 'integrated'.

Definition 9.3 *Let*

$$\mathcal{S}_{L^2(\mathbf{R}_+)} = \{u | u = \sum_{i=1}^n F_i h_i, F_i \in \mathcal{S}, h_i \in L^2(\mathbf{R}_+), n \in \mathbf{N}\}.$$

Lemma 9.1 *Let $u = \sum_{i=1}^n F_i h_i \in \mathcal{S}_{L^2(\mathbf{R}_+)}$. Then we have*

$$\delta(u) = \sum_{i=1}^n [F_i W(h_i) - \langle DF_i, h_i\rangle].$$

Proof
By linearity of δ we may assume that $u = Fh$ with $F \in \mathcal{S}$ and $h \in L^2(\mathbf{R}_+)$. Then for $G \in \mathcal{S}$ by means of Theorem 8.4

$$\begin{aligned} E(\langle u, DG\rangle) &= E(F\langle h, DG\rangle) \\ &= E(FGW(h) - G\langle h, DF\rangle) \\ &= E(G[FW(h) - \langle DF, h\rangle]). \end{aligned}$$

Hence we have

$$\delta(u) = FW(h) - \langle h, DF\rangle.$$

This completes the proof. •

Recall now the standard Wiener filtration $(\mathcal{F}_t)_{t\geq 0}$, which for $t \geq 0$ is given by the P-completion $\mathcal{F}_t$ of $\sigma(W_s : s \leq t)$. Lemma 9.1 yields the elementary Itô integral, if F_i is $\mathcal{F}_{t_i}$-measurable, $h_i = 1_{]t_i,t_{i+1}]}$, where $0 = t_0 < t_1 < \cdots < t_n$, if $\langle DF_i, h_i\rangle = 0, 1 \leq i \leq n-1$. This is indeed the case, as we will show now.

Lemma 9.2 *Let $F \in D_1^2, A \in \mathbf{B}_+^1, \mathcal{F}_A = \sigma(W(1_B) : B \subset A, \lambda(B) < \infty)$. Then we have*

$$E(F|\mathcal{F}_A) \in D_1^2$$

and

$$D_t E(F|\mathcal{F}_A) = E(D_t F|\mathcal{F}_A) 1_A(t)$$

(in $L^2(\Omega \times \mathbf{R}_+)$).

Proof

1. We first consider $F = f(W(h_1), \cdots, W(h_n)) \in \mathcal{S}$. By setting $g(x_1, \cdots, x_n, y_1, \cdots, y_n) = f(x_1 + y_1, \cdots, x_n + y_n), x_1, \cdots, y_n \in \mathbf{R}$, we can write

$$F = g(W(h_1 1_A), \cdots, W(h_n 1_A), W(h_1 1_{A^c}), \cdots, W(h_n 1_{A^c})).$$

Let

$$Q = P \circ (W(h_1 1_{A^c}), \cdots, W(h_n 1_{A^c}))^{-1}.$$

Then by independence of $\mathcal{F}_A$ and the vector $(W(h_1 1_{A^c}), \cdots, W(h_n 1_{A^c}))$ we have

$$E(F|\mathcal{F}_A) = \int g(W(h_1 1_A), \cdots, W(h_n 1_A), y_1, \cdots, y_n) dQ(y_1, \cdots, y_n).$$

Hence $E(F|\mathcal{F}_A) \in \mathcal{S}$ and

$$\begin{aligned} D_t(E(F|\mathcal{F}_A) &= \sum_{i=1}^n \int \frac{\partial}{\partial x_i} g(W(h_1 1_A), \cdots, W(h_n 1_A), y_1, \cdots, y_n) \\ &\quad dQ(y_1, \cdots, y_n) h_i(t) 1_A(t) \\ &= E(D_t F|\mathcal{F}_A) 1_A(t). \end{aligned}$$

2. It remains to approximate $F \in D_1^2$ by standard arguments. •

Theorem 9.1 *Let $u \in L^2(\Omega \times \mathbf{R}_+)$ be $(\mathcal{F}_t)$-adapted. Then*

$$u \in dom(\delta) \text{ and } \delta(u) = \int_{\mathbf{R}_+} u_t dW_t \text{ (Itô integral).}$$

Proof

1. Let $0 \leq s < t, F \in L^2(\Omega, \mathcal{F}_s, P)$. **We prove:**

$$u = F1_{]s,t]} \in \text{dom}(\delta) \text{ and } \delta(u) = F(W_t - W_s).$$

a) Let first $F \in \mathcal{S}$ in addition. Then by Lemma 9.1 and 9.2 we may write

$$\begin{aligned} \delta(u) &= F(W_t - W_s) - \langle DF, 1_{]s,t]} \rangle \\ &= F(W_t - W_s) - \langle DF 1_{[0,s]}, 1_{]s,t]} \rangle \\ &= F(W_t - W_s). \end{aligned}$$

b) For $F \in L^2(\Omega, \mathcal{F}_s, P)$ let $(F^n)_{n \in \mathbf{N}} \subset \mathcal{S}$ such that $F^n \to F$ in $L^2(\Omega, \mathcal{F}, P)$. Then also $\mathcal{S} \ni G^n = E(F^n | \mathcal{F}_s) \to F$ in $L^2(\Omega, \mathcal{F}, P)$. Hence by a) for any $n \in \mathbf{N}$

$$\delta(G^n 1_{]s,t]}) = G^n (W_t - W_s).$$

Moreover, this sequence is a Cauchy sequence in $L^2(\Omega, \mathcal{F}, P)$. Since δ is a closed operator (as a dual operator), we obtain that

$$F1_{]s,t]} \in \text{dom}(\delta) \text{ and } \delta(F1_{]s,t]}) = F(W_t - W_s).$$

2. a) Let now $u = \sum_{j=1}^n F_j 1_{]s_j,t_j]} \in L^2(\Omega \times \mathbf{R}_+)$, where $s_j \leq t_j, F_j \mathcal{F}_{s_j}$-measurable, $1 \leq j \leq n$. Then by linearity

$$u \in \text{dom}(\delta) \text{ and } \delta(u) = \sum_{j=1}^n F_j (W_{t_j} - W_{s_j}).$$

b) Now given u as in the claim, choose a sequence $(u^n)_{n \in \mathbf{N}}$ of simple adapted processes as in a) such that $||u^n - u||_2 \to 0$ in $L^2(\Omega \times \mathbf{R}_+)$. Then use the closedness of δ and the definition of the Itô integral to obtain that

$$\delta(u) = \lim_{n \to \infty} \delta(u^n) = \lim_{n \to \infty} \int_{\mathbf{R}_+} u_t^n dW_t = \int_{\mathbf{R}_+} u_t dW_t.$$

This completes the proof. •

We next ask the question how the Skorokhod integral operates on the decomposition into multiple Wiener-Itô integrals.

Lemma 9.3 *Let $u \in L^2(\Omega \times \mathbf{R}_+)$. Then for $m \geq 0$ there exist functions $f_m \in L^2(\mathbf{R}_+^{m+1})$ such that f_m is symmetric in its first m variables, and such that*

$$u_t = \sum_{m=0}^{\infty} I_m(f_m(\cdot, t)) \text{ in } L^2(\Omega \times \mathbf{R}_+).$$

We have

$$E(\int_{\mathbf{R}_+} u_s^2 ds) = \sum_{m=0}^{\infty} m! ||f_m||_2^2.$$

Proof

Choose a sequence of elementary processes $(u^n)_{n \in \mathbf{N}} \subset L^2(\Omega \times \mathbf{R}_+)$ such that

$$||u^n - u||_2 \to 0 \quad (n \to \infty).$$

Suppose $F_k^n \in L^2(\Omega), g_k^n \in L^2(\mathbf{R}_+), 1 \leq k \leq m_n, n \in \mathbf{N}$ are given such that

$$u_t^n = \sum_{k=1}^{m_n} F_k^n g_k^n(t).$$

For $n \in \mathbf{N}, 1 \leq k \leq m_n$ let

$$F_k^n = \sum_{m=0}^{\infty} I_m(f_m^{k,n}), \quad f_m^{k,n} \in L^2(\mathbf{R}_+^m) \text{ symmetric.}$$

Then we have

$$u_t^n = \sum_{m=0}^{\infty} I_m(\sum_{k=1}^{m_n} f_m^{k,n} g_k^n(t)), \quad t \in \mathbf{R}_+.$$

Define

$$f_m^n = \sum_{k=1}^{m_n} f_m^{k,n} g_k^n, \quad m \geq 0, n \in \mathbf{N}.$$

Then $f_m^n \in L^2(\mathbf{R}_+^{m+1})$, f_m^n is symmetric in its first m variables, and due to orthogonality and symmetry we have for $l, n \in \mathbf{N}$

$$||u^n - u^l||_2^2 = \sum_{m=0}^{\infty} m! ||f_m^n - f_m^l||_2^2.$$

Hence for any $m \geq 0$ $(f_m^n)_{n \in \mathbf{N}}$ is a Cauchy sequence in $L^2(\mathbf{R}_+^{m+1})$ which converges to a function f_m which is also symmetric in the first m variables. We obtain for $u^{n,M} = \sum_{m=0}^{M} \sum_{k=1}^{m_n} F_k^n g_k^n, n, M \in \mathbf{N}$

$$\begin{aligned} \infty > ||u||_2^2 &= \lim_{n \to \infty} ||u^n||_2^2 = \sup_{M \in \mathbf{N}} \lim_{n \to \infty} ||u^{n,M}||_2^2 \\ &= \sup_{M \in \mathbf{N}} \sum_{m=0}^{M} m! ||f_m||_2^2 = \sum_{m=0}^{\infty} m! ||f_m||_2^2. \end{aligned}$$

By a similar argument and by definition we must have

$$u_t = \sum_{m=0}^{\infty} I_m(f_m(\cdot, t)), \quad t \geq 0.$$

This completes the proof. •

Theorem 9.2 *Let* $u \in L^2(\Omega \times \mathbf{R}_+)$, $u = \sum_{m=0}^{\infty} I_m(f_m(\cdot, \cdot))$ *according to Lemma 9.3. Then we have*

$$u \in dom(\delta) \text{ if and only if } \sum_{m=0}^{\infty} (m+1)! ||\tilde{f_m}||_2^2 < \infty.$$

In this case

$$\delta(u) = \sum_{m=0}^{\infty} I_{m+1}(f_m).$$

Proof

1. Let $n \in \mathbf{N}, g \in L^2(\mathbf{R}_+^n), G = I_n(g)$. **We show:**

$$E(\langle u, DG\rangle) = E(I_n(f_{n-1})G).$$

In fact, by Theorem 8.1

$$\begin{aligned} E(\langle u, DG\rangle) &= E(\langle u, nI_{n-1}(g(\cdot, \cdot))\rangle) \\ &= n \int_{\mathbf{R}_+} E(I_{n-1}(f_{n-1}(\cdot, t))I_{n-1}(g(\cdot, t)))dt \\ &= n!\langle f_{n-1}, g\rangle = E(I_n(f_{n-1})G). \end{aligned}$$

2. Let us now prove the *if* part of the claim. For this purpose, let $u \in \mathrm{dom}(\delta), G = I_n(g) \in \mathcal{H}_n, n \in \mathbf{N}$. Then by the first part and by duality

$$E(\delta(u)G) = E(I_n(f_{n-1})G).$$

By extending G to other components of $L^2(\Omega)$ and linearity we obtain

$$L^2(\Omega) \ni \delta(u) = \sum_{n=0}^{\infty} I_{n+1}(f_n), \text{ and } \sum_{n=0}^{\infty} (n+1)! ||\tilde{f_n}||_2^2 < \infty.$$

3. Let us now establish the *only if* part of the claim. Assume that $\sum_{n=0}^{\infty} (n+1)! ||\tilde{f_n}||_2^2 < \infty$. Let $V = \sum_{n=0}^{\infty} I_{n+1}(f_n)$ which is well defined, and $G = \sum_{n=0}^{\infty} I_n(g_n) \in L^2(\Omega)$ with $g_n \in L^2(\mathbf{R}_+^n), n \in \mathbf{N}$, symmetric, and

finally let $G^n = \sum_{k=0}^n I_k(g_k), n \in \mathbf{N}$. Then we may again appeal to the first part of the proof to get

$$E(\langle u, DG^n\rangle) = E(VG^n), \quad n \in \mathbf{N}.$$

By approximation, this equation extends to $G \in D_1^2$. Since $V \in L^2(\Omega)$, we obtain $u \in \text{dom}(\delta)$ and $\delta(u) = V$. This completes the proof. •

For practical purposes it is not easy to deal with $\text{dom}(\delta)$ when discussing the Skorokhod integral. The space is analytically hardly accessible. For having a simpler treatment of questions related to the calculus of Skorokhod's integral it is preferable to work on the following subspace.

Definition 9.4 *Let*

$$\begin{aligned} \mathbf{L}_1^2 &= \{u|u \in L^2(\Omega \times \mathbf{R}_+), u_t \in D_1^2 \textit{ for } \lambda - \textit{ a.a. } t \geq 0, \\ &\quad \textit{for some measurable version of } (s,t) \mapsto D_s u_t \textit{ we have} \\ &\quad E(\int_{\mathbf{R}_+}\int_{\mathbf{R}_+} |D_s u_t|^2 ds dt) < \infty\}. \end{aligned}$$

Remark

$\mathbf{L}_1^2$ is a Hilbert space with the norm $||u||_{1,2}^2 = ||u||_2^2 + ||Du||_2^2$.

How can $\mathbf{L}_1^2$ be described in terms of Wiener-Itô decompositions?

Remark

Let $u \in L^2(\Omega \times \mathbf{R}_+)$, $u = \sum_{m=0}^\infty I_m(f_m(\cdot,\cdot))$ according to Lemma 9.3. Let us formulate in these terms the conditions of the definition of $\mathbf{L}_1^2$. First of all, for $t \geq 0$ according to Theorem 8.1 $u_t \in D_1^2$ means that

$$\sum_{m=0}^\infty m\, m! ||f_m(\cdot, t)||_2^2 < \infty.$$

In the same terms, $E(\int_{\mathbf{R}_+}\int_{\mathbf{R}_+} |D_s u_t|^2 ds dt) < \infty$ then means that

$$\sum_{m=0}^\infty m\, m! \int_{\mathbf{R}_+} ||f_m(\cdot,t)||_2^2 dt = \sum_{m=0}^\infty m\, m! ||f_m||_2^2 < \infty.$$

The latter is the case iff

$$\sum_{m=0}^\infty (m+1)! ||f_m||_2^2 < \infty.$$

Compare this with the condition we obtained in Theorem 9.2. Since for $m \geq 0$ we have $||\tilde{f}_m||_2 \leq ||f_m||_2$, we obviously have

$$\mathbf{L}_1^2 \subset \text{dom}(\delta).$$

How is Itô's isometry transferred to the Skorokhod integral?

Theorem 9.3 *Let* $u, v \in \mathbf{L}_1^2$. *Then*

$$E(\delta(u)\delta(v)) = E(\int_{\mathbf{R}_+} u_t v_t dt) + E(\int_{\mathbf{R}_+^2} D_t u_s \, D_s v_t \, dsdt).$$

Proof

1. Let first $u \in \mathcal{S}_{L^2(\mathbf{R}_+)}$. **We show:**

$$D_t \delta(u) = u + \delta(D_t u) \text{ (in } L^2(\Omega \times \mathbf{R}_+)).$$

By linearity, we may further assume that $u = Fh$, where $F \in \mathcal{S}, h \in L^2(\mathbf{R}_+)$. Then according to Lemma 9.1 we may write $\delta(u) = FW(h) - \langle DF, h\rangle$, and therefore

$$\begin{aligned} D_t\delta(u) &= D_t F W(h) + Fh - \langle D_t DF, h\rangle \\ &= u + \delta(D_t F h) \\ &= u + \delta(D_t u). \end{aligned}$$

2. Let still $u \in \mathcal{S}_{L^2(\mathbf{R}_+)}$. Then duality, the calculation just obtained and the fact that $v \in \mathbf{L}_1^2$ lead to

$$\begin{aligned} E(\delta(u)\delta(v)) &= E(\langle D\delta(u), v\rangle) \\ &= E(\langle u, v\rangle + \langle \delta(D.u), v\rangle) \\ &= E(\langle u, v\rangle) + E(\int_{\mathbf{R}_+^2} D_t u_s \, D_s v_t \, dsdt). \end{aligned}$$

3. It remains to do an approximation of u by functions in $\mathcal{S}_{L^2(\mathbf{R}_+)}$, and to use that convergence in L_1^2 implies convergence for all three terms in the formula. •

We finally give a rule for the Malliavin differentiation of Itô integrals which will be of use in the applications of Malliavin's calculus to stochastic analysis to be discussed.

Theorem 9.4 *Let* $u \in L^2(\Omega \times [0,1])$ *be adapted,* $X_t = \int_0^t u_s dW_s, 0 \le t \le 1$, *its Itô integral process. Then we have*

$$u \in \mathbf{L}_1^2 \text{ if and only if } X_T \in D_1^2 \text{ for all } T \in [0,1].$$

In this case $X \in \mathbf{L}_1^2$ *and for* $0 \le t \le T \le 1$ *we have*

$$D_t X_T = u_t 1_{[0,T]}(t) + \int_t^T D_t u_r dW_r,$$

and

$$\int_0^T E(|D_t X_T|^2) dt = \int_0^T E(u_t)^2 dt + \int_0^T \int_t^T E(|D_t u_r|^2) drdt.$$

Proof

For simplicity let $T = 1$. This time we use Wiener-Itô decompositions for approximations.

1. We will use an extension of the representation of Lemma 9.3 to adapted u. Let $u = \sum_{m=0}^{\infty} I_m(f_m(\cdot, \cdot))$ be the representation according to Lemma 9.3. Here for any $m \geq 0$ the function $f_m \in L^2(\mathbf{R}_+^{m+1})$ is symmetric in its first m variables. **We show:**

$$\begin{aligned} f_m(\cdot, t) &= f_m(\cdot, t) 1_{[0,t]^m}(\cdot) \quad (\text{in } L^2(\mathbf{R}_+^{m+1})), \quad \text{and thus} \\ ||\tilde{f_m}||_2^2 &= \frac{1}{m+1} ||f_m||_2^2. \end{aligned}$$

In fact, resume the notation of the proof of Lemma 9.3, to specialize to the adapted case. We approximate u in $L^2(\Omega \times [0, 1])$ by functions

$$u^n = \sum_{k=1}^{m_n} F_k^n 1_{]t_k^n, t_{k+1}^n]}, \quad n \in \mathbf{N},$$

where $0 = t_1^n < \cdots < t_{m_n}^n = 1, F_k^n$ is $\mathcal{F}_{t_k^n}$-measurable. In the Wiener-Itô decomposition

$$F_k^n = \sum_{m=0}^{\infty} I_m(f_m^{k,n})$$

of F_k^n, due to its measurability properties and Lemma 9.2, we have for $t_1, \cdots, t_m \in \mathbf{R}_+$

$$\begin{aligned} f_m^{k,n}(t_1, \cdots, t_m) &= m! D_{t_1} \cdots D_{t_m} I_m(f_m^{k,n}) \\ &= m! D_{t_1} \cdots D_{t_m} I_m(f_m^{k,n}) 1_{[0,t_k^n]^m}(t_1, \cdots, t_m) \\ &= f_m^{k,n}(t_1, \cdots, t_m) \, 1_{[0,t_k^n]^m}(t_1, \cdots, t_m). \end{aligned}$$

Hence the functions

$$f_m^n = \sum_{k=1}^{m_n} F_m^{k,n} 1_{]t_k^n, t_{k+1}^n]}, \quad m \geq 0, n \in \mathbf{N},$$

possess the property

$$f_m^n(\cdot, t) = f_m^n(\cdot, t) 1_{[0,t]^m}(\cdot) \quad \text{for any} \quad t \geq 0.$$

Now use a diagonal sequence argument to select a subsequence $(v^n)_{n \in \mathbf{N}}$ of $(u^n)_{n \in \mathbf{N}}$ with corresponding Wiener-Itô functions $g_m^n \in L^2(\mathbf{R}_+^{m+1})$, symmetric in their first m variables, and such that for any $m \geq 0$ $g_m^n \to f_m$ $P \otimes \lambda$-a.e. Hence we have for $m \geq 0$

$$f_m(\cdot, t) = f_m(\cdot, t)1_{[0,t]^m}(\cdot) \quad (\text{in } L^2(\mathbf{R}_+^{m+1})).$$

To prove the second assertion, note that due to the validity of the first

$$\tilde{f}_m = \frac{1}{m+1} \sum_{i=1}^{m+1} h_m^i,$$

where the h_m^i have disjoint support, and their norms in $L^2(\mathbf{R}_+^{m+1})$ are identical to the one of f_m. Hence

$$||\tilde{f}_m||_2^2 = \frac{1}{(m+1)^2} \sum_{i=1}^{m+1} ||h_m^i||_2^2 = \frac{1}{m+1}||f_m||_2^2,$$

as claimed.

2. Now as was shown in the remark above, $u \in \mathbf{L}_1^2$ translates into

$$\sum_{m=0}^{\infty} (m+1)!||f_m||_2^2 < \infty.$$

Moreover, we know that $X_1 = \delta(u) = \sum_{m=0}^{\infty} I_{m+1}(f_m)$, and by Theorem 8.1 that $X_1 \in D_1^2$ if and only if

$$\sum_{m=0}^{\infty} (m+1)(m+1)!||\tilde{f}_m||_2^2 < \infty.$$

But according to the first part of the proof

$$\sum_{m=1}^{\infty} (m+1)(m+1)!||\tilde{f}_m||_2^2 = \sum_{m=1}^{\infty} (m+1)!||f_m||_2^2.$$

This proves the first claim of the Theorem.

3. Let us next prove the differentiation formula. Note first that due to Lemma 9.2

$$D_t u_s = D_t u_s 1_{[t,1]}(s) \quad (\text{in } L^2([0,1]^2)).$$

Moreover, for $t \leq s \leq 1$ we have, according to Theorem 8.1

$$\begin{aligned} D_t u_s &= D_t \sum_{m=0}^{\infty} I_m(f_m(\cdot, s)) \\ &= \sum_{m=1}^{\infty} m I_{m-1}(f_m(\cdots, t, s)). \end{aligned}$$

Since $Du \in L^2(\Omega \times [0,1]^2)$ according to the definition of $\mathbf{L}_1^2$, and since for fixed t the process $(D_t u_s)_{t \le s \le 1}$ is adapted, we know from Theorem 9.1 that $(D_t u_s)_{t \le s \le 1}$ is Itô integrable and that its Itô and Skorokhod integral are identical. More formally,

$$\begin{aligned}\int_t^1 D_t u_s dW_s &= \int_0^1 D_t u_s dW_s = \delta(D_t u) \\ &= \sum_{m=0}^{\infty} m I_m(f_m(\cdots, t, \cdot)).\end{aligned}$$

Finally, we know that $X_1 = \delta(u) \in D_1^2$. We can compute the Malliavin derivative for $t \in [0,1]$ (in the usual sense of equality in $L^2(\Omega \times [0,1])$)

$$\begin{aligned}D_t \delta(u) &= D_t \sum_{m=0}^{\infty} I_{m+1}(f_m) \\ &= D_t \sum_{m=0}^{\infty} I_{m+1}(\tilde{f_m}) \\ &= \sum_{m=1}^{\infty} (m+1) I_m(\tilde{f_m}(\cdot, t)) \\ &= \sum_{m=1}^{\infty} I_m(f_m(\cdot, t)) + \sum_{m=1}^{\infty} m I_m(f_m(\cdots, t, \cdot)).\end{aligned}$$

The last line is now identified with $u_t + \int_t^1 D_t u_s dW_s$ by using the expansions of these two expressions given above. The norm equation follows from Itô's isometry. •

10 Backward stochastic differential equations

Backward stochastic differential equations (BSDE) constitute a very successful and active tool for stochastic finance and insurance, and more generally serve as a central method of stochastic control theory. In this chapter we shall establish the basic existence and uniqueness theory for these equations in case the coefficients are Lipschitz continuous.

We fix for the sequel a finite time horizon $T > 0$, and a dimension $m \in \mathbf{N}$. We start by explaining some notation. Let $(\Omega, \mathcal{F}, P)$ be the canonical n-dimensional Wiener space, with canonical Wiener process $W = (W^1, \cdots, W^n)$. Denote by $(\mathcal{F}_t)_{t \ge 0}$ the filtration of the canonical space, i.e. the natural filtration completed by sets of P-measure 0.

Let $L^2(\mathbf{R}^m)$ be the linear space of $\mathbf{R}^m$-valued $\mathcal{F}_T$-measurable random variables, endowed with norm $E(|X|^2)^{\frac{1}{2}}$. Let $H^2(\mathbf{R}^m)$ denote the linear space of $(\mathcal{F}_t)_{0\le t\le T}$-adapted measurable processes $X : \Omega \times [0,T] \to \mathbf{R}^m$ endowed with the norm $||X||_2 = E(\int_0^T |X_t|^2 dt)^{\frac{1}{2}}$. Further let $H^1(\mathbf{R}^m)$ denote the space of $(\mathcal{F}_t)_{0\le t\le T}$-adapted measurable processes $X : \Omega \times [0,T] \to \mathbf{R}^m$ with the norm $||X||_1 = E([\int_0^T |X_t|^2 dt]^{\frac{1}{2}})$. Finally, for $\beta > 0$ and $X \in H^2(\mathbf{R}^m)$ let

$$||X||_{2,\beta}^2 = E(\int_0^T e^{\beta t}|X_t|^2 dt),$$

and $H^{2,\beta}(\mathbf{R}^m)$ the space $H^2(\mathbf{R}^m)$ endowed with the norm $||\cdot||_{2,\beta}$.

We next describe the general hypotheses we want to require for the parameters of our BSDE. The terminal condition ξ will be supposed to belong to $L^2(\mathbf{R}^m)$. The *generator* will be a function

$$f : \Omega \times \mathbf{R}_+ \times \mathbf{R}^m \times \mathbf{R}^{n\times m} \to \mathbf{R}^m,$$

which is product measurable, adapted in the time parameter, and which fulfills

$$\text{(H1)} \quad f(\cdot, 0, 0) \in H^2(\mathbf{R}^m),$$

f is *uniformly Lipschitz*, i.e. there exists $C \in \mathbf{R}$ such that for any $(y_1, z_1), (y_2, z_2) \in \mathbf{R}^m \times \mathbf{R}^{n\times m}$, $P \otimes \lambda$-a.e. $(\omega, t) \in \Omega \times \mathbf{R}_+$

$$\text{(H2)} \quad |f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)| \le C[|y_1 - y_2| + |z_1 - z_2|].$$

Here for $z \in \mathbf{R}^{n\times m}$ we denote $|z| = (tr(zz^*))^{\frac{1}{2}}$.

Definition 10.1 *A pair of functions (f, ξ) fulfilling, besides the mentioned measurement requirements, hypotheses (H1), (H2), is said to be a* standard parameter.

Given standard parameters, we shall solve the problem of finding a pair of $(\mathcal{F}_t)_{0\le t\le T}$-adapted processes $(Y_t, Z_t)_{0\le t\le T}$ such that the *backward stochastic differential equation (BSDE)*

$$(*) \quad dY_t = Z_t^* dW_t - f(\cdot, t, Y_t, Z_t)dt, \quad Y_T = \xi,$$

is satisfied.

In order to construct a solution, a contraction argument on suitable Banach spaces will be used. For its derivation we shall need the following *a priori inequalities.*

Lemma 10.1 *For $i = 1, 2$ let (f^i, ξ^i) be standard parameters, $(Y^i, Z^i) \in H^2(\mathbf{R}^m) \times H^2(\mathbf{R}^{n\times m})$ solutions of (*) with corresponding standard parameters. Let C be a Lipschitz constant for f^1. Define for $0 \le t \le T$*

$$\begin{aligned} \delta Y_t &= Y_t^1 - Y_t^2, \\ \delta_2 f_t &= f^1(\cdot, t, Y_t^2, Z_t^2) - f^2(\cdot, t, Y_t^2, Z_t^2). \end{aligned}$$

Then for any triple (λ, μ, β) with $\lambda > 0, \lambda^2 > C, \beta \ge C(2 + \lambda^2) + \mu^2$ we have

$$\begin{aligned} ||\delta Y||_{2,\beta}^2 &\le T[e^{\beta T} E(|\delta Y_T|^2) + \frac{1}{\mu^2}||\delta_2 f||_{2,\beta}^2], \\ ||\delta Z||_{2,\beta}^2 &\le \frac{\lambda^2}{\lambda^2 - C}[e^{\beta T} E(|\delta Y_T|^2) + \frac{1}{\mu^2}||\delta_2 f||_{2,\beta}^2]. \end{aligned}$$

Proof

1. Let $(Y, Z) \in H^2(\mathbf{R}^m) \times H^2(\mathbf{R}^{n\times m})$ be a solution of (*) with standard parameters (f, ξ). This means that we may write for $0 \le t \le T$

$$(*) \quad Y_t = \xi - \int_t^T Z_s^* dW_s + \int_t^T f(\cdot, s, Y_s, Z_s) ds.$$

We show:

$$\sup_{0\le t\le T} |Y_t| \in L^2(\mathbf{R}^m).$$

In fact, due to (*) we have

$$\sup_{0\le t\le T} |Y_t| \le |\xi| + \int_0^T |f[\cdot, s, Y_s, Z_s)| ds + \sup_{0\le t\le T} |\int_t^T Z_s^* dW_s|,$$

and, with the help of Doob's inequality

$$E(\sup_{0\le t\le T} |\int_t^T Z_s^* dW_s|^2) \le 4E(\sup_{0\le t\le T} |\int_0^t Z_s^* dW_s|^2) \le 8E(\int_0^T |Z_s|^2 ds).$$

Since in addition (H1) and (H2) guarantee that $|\xi| + \int_0^T |f(\cdot, s, Y_s, Z_s)| ds \in L^2(\mathbf{R})$, we obtain the desired

$$E(\sup_{0\le t\le T} |Y_t|^2) < \infty.$$

2. Now we derive a preliminary bound. Apply Itô's formula to the semimartingale $(e^{\beta s}|\delta Y_s|^2)_{0\le s\le T}$ to obtain for $0 \le t \le T$

$$\begin{aligned} &e^{\beta T}|\delta Y_T|^2 - e^{\beta t}|\delta Y_t|^2 \\ &\quad = \beta \int_t^T e^{\beta s}|\delta Y_s|^2 ds + 2\int_t^T e^{\beta s}\langle \delta Y_s, f^1(\cdot, s, Y_s^1, Z_s^1) \\ &\qquad - f^2(\cdot, s, Y_s^2, Z_s^2)\rangle ds - 2\int_t^T e^{\beta s}\langle \delta Y_s, \delta Z_s^* dW_s\rangle + \int_t^T e^{\beta s}|\delta Z_s|^2 ds. \end{aligned}$$

By reordering the terms in the equation we obtain

$$\begin{aligned} e^{\beta t}|\delta Y_t|^2 &+ \beta \int_t^T e^{\beta s}|\delta Y_s|^2 ds + \int_t^T e^{\beta s}|\delta Z_s|^2 ds \\ &= e^{\beta T}|\delta Y_T|^2 + 2\int_t^T e^{\beta s}\langle \delta Y_s, \delta Z_s^* dW_s\rangle \\ &\quad -2\int_t^T e^{\beta s}\langle \delta Y_s, f^1(\cdot, s, Y_s^1, Z_s^1) - f^2(\cdot, s, Y_s^2, Z_s^2)\rangle ds. \end{aligned}$$

3. **We prove** for $0 \le t \le T$:

$$E(e^{\beta t}|\delta Y_t|^2) \le E(e^{\beta T}|\delta Y_T|^2) + \frac{1}{\mu^2} E(\int_t^T e^{\beta s}|\delta_2 f_s|^2 ds).$$

To prove this, first take expectations on both sides of the inequality obtained in 2., with the result

$$\begin{aligned} E(e^{\beta t}|\delta Y_t|^2) &+ \beta E(\int_t^T e^{\beta s}|\delta Y_s|^2 ds) + E\int_t^T e^{\beta s}|\delta Z_s|^2 ds) \\ &\le E(e^{\beta T}|\delta Y_T|^2) \\ &\quad +2E(\int_t^T e^{\beta s}\langle \delta Y_s, f^1(\cdot, s, Y_s^1, Z_s^1) - f^2(\cdot, s, Y_s^2, Z_s^2)\rangle ds). \end{aligned}$$

Now by our assumptions for $0 \le s \le T$

$$\begin{aligned} |f^1(\cdot, s, Y_s^1, Z_s^1) - f^2(\cdot, s, Y_s^2, Z_s^2)| &\le |f^1(\cdot, s, Y_s^1, Z_s^1) - f^1(\cdot, s, Y_s^2, Z_s^2)| \\ &\quad +|\delta_2 f_s| \\ &\le C[|\delta_s Y| + |\delta_s Z|] + |\delta_2 f_s|. \end{aligned}$$

The latter implies

$$\begin{aligned} &\int_t^T E(2e^{\beta s}|\langle \delta Y_s, f^1(\cdot, s, Y_s^1, Z_s^1) - f^2(\cdot, s, Y_s^2, Z_s^2)\rangle|ds \\ &\le \int_t^T 2e^{\beta s} E(|\delta Y_s|[C(|\delta_s Y| + |\delta_s Z|) + |\delta_2 f_s|]ds \\ &= \int_t^T 2e^{\beta s}[CE(|\delta Y_s|^2) + E(|\delta_s Y|(C|\delta_s Z|) + |\delta_2 f_s|)]ds. \end{aligned}$$

Now for $C, y, z, t > 0$ with $\mu, \lambda > 0$

$$\begin{aligned} 2y(Cz + t) &= 2Cyz + 2yt \\ &\le C[(y\lambda)^2 + (\frac{z}{\lambda})^2] + (y\mu)^2 + (\frac{t}{\mu})^2 \\ &= C(\frac{z}{\lambda})^2 + (\frac{t}{\mu})^2 + y^2(\mu^2 + C\lambda^2). \end{aligned}$$

With this we can estimate the last term in our inequality further:

$$\begin{aligned}\int_t^T 2e^{\beta s}[CE(|\delta Y_s|^2) &+ E(|\delta_s Y|(C|\delta_s Z| + |\delta_2 f_s|))]ds \\ &\leq \int_t^T e^{\beta s}[2CE(|\delta Y_s|^2) + \frac{C}{\lambda^2}E(|\delta_s Z|^2) \\ &\qquad + \frac{1}{\mu^2}E(|\delta_2 f_s|^2) + (\mu^2 + C\lambda^2)E(|\delta_s Y|^2]ds \\ &= \int_t^T e^{\beta s}[(\mu^2 + C(2+\lambda^2))E(|\delta Y_s|^2) \\ &\qquad + \frac{C}{\lambda^2}E(|\delta_s Z|^2) + \frac{1}{\mu^2}E(|\delta_2 f_s|^2)]ds.\end{aligned}$$

Summarizing, we obtain, using our assumptions on the parameters

$$\begin{aligned}(**) \quad E(e^{\beta t}|\delta Y_t|^2) &\leq E(\int_t^T e^{\beta s}|\delta Y_s|^2 ds)[-\beta + C(2+\lambda^2) + \mu^2] \\ &\quad + E(\int_t^T e^{\beta s}|\delta Z_s|^2 ds)[\frac{C}{\lambda^2} - 1] + E(e^{\beta T}|\delta Y_T|^2) \\ &\quad + \frac{1}{\mu^2}E(\int_t^T e^{\beta s}|\delta_2 f_s|^2 ds) + E(e^{\beta T}|\delta Y_T|^2) \\ &\leq E(e^{\beta T}|\delta Y_T|^2) + \frac{1}{\mu^2}E(\int_t^T e^{\beta s}|\delta_2 f_s|^2 ds).\end{aligned}$$

This is the claimed inequality.

4. In order to obtain the first inequality in the assertion, it remains to integrate the inequality resulting from 3. in $t \in [0, T]$.

5. The second inequality in the assertion follows from (**) by taking the second term from the right hand side to the left. This completes the proof. •

We are in a position to state existence and uniqueness results for our BSDE (*).

Theorem 10.1 *Let (ξ, f) be standard parameters. Then there exists a uniquely determined pair $(Y, Z) \in H^2(\mathbf{R}^m) \times H^2(\mathbf{R}^{n\times m})$ with the property*

$$\textit{(BSDE)} \quad Y_t = \xi - \int_t^T Z_s^* dW_s + \int_t^T f(\cdot, s, Y_s, Z_s)ds, \quad 0 \leq t \leq T.$$

Proof

Consider

$$\Gamma : H^{2,\beta}(\mathbf{R}^m) \times H^{2,\beta}(\mathbf{R}^{n\times m}) \to H^{2,\beta}(\mathbf{R}^m) \times H^{2,\beta}(\mathbf{R}^{n\times m}), (y, z) \mapsto (Y, Z),$$

where (Y, Z) is a solution of the BSDE

$$(*)\quad Y_t = \xi - \int_t^T Z_s^* dW_s + \int_t^T f(\cdot, s, y_s, z_s) ds, \quad 0 \le t \le T.$$

1. **We prove:** (Y, Z) is well defined. First of all, our assumptions yield

$$\xi + \int_t^T f(\cdot, s, y_s, z_s) ds \in L^2(\Omega), \quad 0 \le t \le T.$$

Therefore

$$M_t = E(\xi + \int_0^T f(\cdot, s, y_s, z_s) ds | \mathcal{F}_t), \quad 0 \le t \le T,$$

is a well defined martingale. M possesses a continuous version, due to the fact that we are working in a Wiener filtration. M is square integrable. Hence we may apply the martingale representation theorem, which provides (a unique) $Z \in H^2(\mathbf{R}^{n\times m})$ such that

$$M_t = M_0 + \int_0^t Z_s^* dW_s, \quad 0 \le t \le T.$$

Let now

$$Y_t = M_t - \int_0^t f(\cdot, s, y_s, z_s) ds.$$

Then Y is square integrable, and we have

$$Y_t = E(\xi + \int_t^T f(\cdot, s, y_s, z_s) ds | \mathcal{F}_t), 0 \le t \le T.$$

Hence

$$Y_T = \xi = M_0 + \int_0^T Z_s^* dW_s - \int_0^T f(\cdot, s, y_s, z_s) ds,$$

and thus for $0 \le t \le T$

$$\begin{aligned} Y_t &= \xi - M_0 - \int_0^T Z_s^* dW_s + \int_0^T f(\cdot, s, y_s, z_s) ds \\ &+ M_0 + \int_0^t Z_s^* dW_s - \int_0^t f(\cdot, s, y_s, z_s) ds \\ &= \xi - \int_t^T Z_s^* dW_s + \int_t^T f(\cdot, s, y_s, z_s) ds. \end{aligned}$$

2. **We prove:** For $\beta > 2(1+T)C$ the mapping Γ is a contraction. For this purpose, let $(y^1, z^1), (y^2, z^2) \in H^{2,\beta}(\mathbf{R}^m) \times H^{2,\beta}(\mathbf{R}^{n\times m})$, $(Y^1, Z^1), (Y^2, Z^2)$ corresponding solutions of (*) according to 1. We apply Lemma 10.1 with $C = 0, \beta = \mu^2$, and $f^i = f(\cdot, y^i, z^i)$. With this

choice we obtain

$$\begin{aligned}||\delta Y||_{2,\beta} &\leq \frac{T}{\beta}[E(\int_0^T e^{\beta s}|f(\cdot,s,y_s^1,z_s^1)-f(\cdot,s,y_s^2,z_s^2)|^2 ds)]^{\frac{1}{2}},\\ ||\delta Z||_{2,\beta} &\leq \frac{1}{\beta}[E(\int_0^T e^{\beta s}|f(\cdot,s,y_s^1,z_s^1)-f(\cdot,s,y_s^2,z_s^2)|^2 ds)]^{\frac{1}{2}}.\end{aligned}$$

Since f is Lipschitz continuous, we further obtain

$$\begin{aligned}||\delta Y||_{2,\beta} &\leq \frac{2TC}{\beta}[||\delta y||_{2,\beta}+||\delta z||_{2,\beta}],\\ ||\delta Z||_{2,\beta} &\leq \frac{2C}{\beta}[||\delta y||_{2,\beta}+||\delta z||_{2,\beta}].\end{aligned}$$

We summarize to obtain

$$(**)\quad ||\delta Y||_{2,\beta}+||\delta Z||_{2,\beta} \leq \frac{2C(T+1)}{\beta}[||\delta y||_{2,\beta}+||\delta z||_{2,\beta}].$$

By choice of β, Γ is a contraction.

3. Now let $(\overline{Y},\overline{Z})$ be the fixed point of Γ, which exists due to 2. Let

$$Y_t = E(\xi+\int_t^T f(\cdot,s,\overline{Y}_s,\overline{Z}_s)ds|\mathcal{F}_t),\quad 0\leq t\leq T.$$

Then Y is continuous and P-a.s. identical to $\overline{Y}$. Then $(Y,\overline{Z})$ is a solution of our BSDE.

4. Uniqueness follows from the contraction property of Γ and the uniqueness of the fixed point. •

The construction of solutions in the preceding proof rests upon a recursive algorithm. The algorithm converges, as we shall note in the following Corollary.

Corollary 10.1 *Let $\beta > 2(1+T)C$, $((Y^k,Z^k))_{k\geq 0}$ the sequence of processes, given by $Y^0 = Z^0 = 0$,*

$$Y_t^{k+1} = \xi - \int_t^T (Z_s^{k+1})^* dW_s + \int_t^T f(\cdot,s,Y_s^k,Z_s^k)ds$$

according to the proof of the preceding Theorem. Then $((Y^k,Z^k))_{k\geq 0}$ converges in $H^{2,\beta}(\mathbf{R}^m)\times H^{2,\beta}(\mathbf{R}^{n\times m})$ to the uniquely determined solution (Y,Z) of (BSDE).

Proof
The inequality (**) in the proof of Theorem 10.1 recursively yields

$$||Y^{k+1} - Y^k||_{2,\beta} + ||Z^{k+1} - Z^k||_{2,\beta} \le \varepsilon^k[||Y^1 - Y^0||_{2,\beta} + ||Z^1 - Z^0||_{2,\beta}],$$

with $\varepsilon = \frac{2C(T+1)}{\beta} < 1$. This implies

$$\sum_{k\in\mathbf{N}} [||Y^{k+1} - Y^k||_{2,\beta} + ||Z^{k+1} - Z^k||_{2,\beta}] < \infty.$$

Now a standard argument applies. •

11 Interpretation of backward stochastic differential equations in Malliavin's calculus

In this chapter we shall establish the vital connection between Malliavin's calculus and the structure of solutions of BSDE. We shall see that, provided the standard parameters are sufficiently smooth, the process Z can be interpreted as the *Malliavin trace* of the process Y. For doing this, we first have to introduce the version of the space $\mathbf{L}_1^2$ which corresponds to integrability in some arbitrary power $p \ge 2$. For simplicity, we let the dimension of our underlying Wiener process be one, i.e. for this chapter we set $n = 1$.

Definition 11.1 *Let* $p \ge 2$, *and*

$$\begin{aligned}\mathbf{L}_1^p(\mathbf{R}^m) \;=\; \{u|u \textit{ adapted}, \; & [\int_0^T |u_t|^2 dt]^{\frac{1}{2}} \in L^p(\Omega),\\ & u_t \in (D_1^p)^m \textit{ for } \lambda - \textit{ a.a. } t \ge 0,\\ & \textit{for some measurable version}\\ & \textit{of } (s,t) \mapsto D_s u_t \textit{ we have } E([\int_0^T \int_0^T |D_s u_t|^2 ds dt]^{\frac{p}{2}}) < \infty\}.\end{aligned}$$

For $u \in \mathbf{L}_1^p(\mathbf{R}^m)$ *define*

$$||u||_{1,p} = E([\int_0^T |u_t|^2 dt]^{\frac{p}{2}}) + E([\int_0^T \int_0^T |D_s u_t|^2 ds dt]^{\frac{p}{2}}).$$

To abbreviate, for $k \in \mathbf{N}, v \in L^2(\Omega \times [0,T]^k)$ denote

$$||v|| = [\int_{[0,T]^k} |v_t|^2 dt]^{\frac{1}{2}}.$$

In these terms, Jensen's inequality gives

$$E(||Du||^p) \leq T^{\frac{p}{2}-1} \int_0^T ||D_s u||_p^p ds.$$

We next prove that solutions of BSDE for regular standard parameters are Malliavin differentiable, and that Z allows an interpretation as a Malliavin trace of Y. We need some versions of the process spaces considered in the previous chapter that correspond to p-integrable random variables. For $p \geq 2$ denote by $S^p(\mathbf{R}^m)$ the linear space of all measurable $(\mathcal{F}_t)$-adapted continuous processes $X : \Omega \times [0,T] \to \mathbf{R}^m$, endowed with the norm $||X||_{S^p} = E(\sup_{0\leq t\leq T} |X_t|^p)^{\frac{1}{p}}$. Let further $H^p(\mathbf{R}^m)$ denote the linear space of measurable $(\mathcal{F}_t)$-adapted processes $X : \Omega \times [0,T] \to \mathbf{R}^m$ endowed with the norm $||X||_p = E(||X||^p)^{\frac{1}{p}}$. To abbreviate, let $B^p(\mathbf{R}^m) = S^p(\mathbf{R}^m) \times H^p(\mathbf{R}^m)$, with the norm $||(Y,Z)||_p = [||Y||_{S^p}^p + ||Z||_p^p]^{\frac{1}{p}}$.

We are ready to state our main result.

Theorem 11.1 *Let (f,ξ) be standard parameters such that $\xi \in D_1^2 \cap L^4(\mathbf{R}^m)$, $f : \Omega \times [0,T] \times \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}^m$ continuously differentiable in (y,z), with uniformly bounded and continuous partial derivatives, and such that for $(y,z) \in \mathbf{R}^m \times \mathbf{R}^m$ we have*

$$(H3) \quad f(\cdot,y,z) \in \mathbf{L}_1^2, \quad f(\cdot,0,0) \in H^4(\mathbf{R}^m),$$

for $t \in [0,T]$ and $(y^1,z^1,y^2,z^2) \in (\mathbf{R}^m \times \mathbf{R}^m)^2$ we have

$$(H4) \quad |D_s f(\cdot,t,y^1,z^1) - D_s f(\cdot,t,y^2,z^2)| \leq K_s(t)[|y^1-y^2| + |z^1-z^2|]$$

(in $L^2(\Omega \times [0,t]^2)$), with a real-valued measurable process $(K_s(t))_{0\leq s\leq t}$ which is $(\mathcal{F}_t)$-adapted in t, and satisfies

$$\int_0^T ||K_s||_4^4 ds < \infty.$$

For the unique solution (Y,Z) of the BSDE (*) we moreover suppose

$$\int_0^T ||D_s f(\cdot,Y,Z)||^2 ds < \infty$$

P-a.s..

Then we have:

$$(Y,Z) \in \mathbf{L}_1^2(\mathbf{R}^m) \times \mathbf{L}_1^2(\mathbf{R}^m),$$

and a (measurable) version of $(D_sY_t, D_sZ_t)_{0\le s,t\le T})$ *possesses the properties*

$$\begin{aligned}
D_sY_t &= D_sZ_t = 0, \quad 0 \le t < s \le T,\\
D_sY_t &= D_s\xi - \int_t^T D_sZ_u^* dW_u\\
&+ \int_t^T [\frac{\partial}{\partial y} f(\cdot, u, Y_u, Z_u)\, D_sY_u + \frac{\partial}{\partial z} f(\cdot, u, Y_u, Z_u)\, D_sZ_u\\
&\qquad + D_sf(\cdot, u, Y_u, Z_u)]du, \quad 0 \le s \le t \le T,\\
(D_sY_s)_{0\le s\le T} &\quad \text{is a version of } (Z_s)_{0\le s\le T}.
\end{aligned}$$

Proof

1. For further simplifying notation, we assume $m = 1$. As in chapter 10, our arguments are mainly based upon several *a priori estimates*. The first one is an analogue of Lemma 10.1 and investigates the properties of the contraction map on $B^2(\mathbf{R})$.

Lemma 11.1 *Let* $p \ge 2$*, assume* $f(\cdot, 0, 0) \in H^p(\mathbf{R})$*, and define*

$$\Gamma : B^p(\mathbf{R}) \to B^p(\mathbf{R}), (y, z) \mapsto (Y, Z),$$

where (Y, Z) *is the solution of the BSDE*

$$(+) \quad Y_t = \xi - \int_t^T Z_s dW_s + \int_t^T f(\cdot, s, y_s, z_s)ds.$$

Let further for $i = 1, 2$ (Y^i, Z^i) *be the solutions corresponding to* (y^i, z^i) *in (+), and let*

$$\delta Y = Y^1 - Y^2, \quad \delta Z = Z^1 - Z^2, \quad \delta y = y^1 - y^2, \quad \delta z = z^1 - z^2.$$

Then there exists a constant C_p *not depending on* (y, z, Y, Z) *such that*

$$\begin{aligned}
\text{(i)} \quad ||Y||_{S^p}^p &\le C_pE([|\xi|^p) + T^{\frac{p}{2}}(\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds)^{\frac{p}{2}}]),\\
\text{(ii)} \quad ||Z||_p^p &\le C_pE([|\xi|^p) + T^{\frac{p}{2}}(\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds)^{\frac{p}{2}}]),\\
\text{(iii)} \quad ||(\delta Y, \delta Z)||_p^p &\le C_pT^{\frac{p}{2}}||(\delta y, \delta z)||_p^p.
\end{aligned}$$

Proof

a) Taking up the notation of the proof of Lemma 10.1, **we show:** Γ is well defined.

To do this, recall for $0 \le t \le T$

$$Y_t = E(\xi + \int_t^T f(\cdot, s, y_s, z_s) ds | \mathcal{F}_t).$$

We have

$$|Y_t| \le E(|\xi| + \int_t^T |f(\cdot, s, y_s, z_s)| ds | \mathcal{F}_t),$$

hence Doob's inequality provides a universal constant C_p^1 such that

$$||Y||_{S^p}^p \le C_p^1 E([|\xi| + \int_0^T |f(\cdot, s, y_s, z_s)| ds]^p).$$

Moreover, by Cauchy-Schwarz' inequality

$$\int_0^T |f(\cdot, s, y_s, z_s)| ds \le T^{\frac{1}{2}} [\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds]^{\frac{1}{2}},$$

hence with another universal constant C_p^2 we have

$$(*) \quad ||Y||_{S^p}^p \le C_p^2 E([|\xi|^p + [\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds]^{\frac{p}{2}}]).$$

Invoke $f(\cdot, 0, 0) \in H^p(\mathbf{R})$, that f is Lipschitz continuous, and that $(y, z) \in B^p(\mathbf{R})$, to see that the right hand side of the preceding inequality is finite.

We next prove that $Z \in H^p(\mathbf{R})$. For this purpose we shall use the inequality of Burkholder. It yields further universal constants $C_p^3, \cdot, C_p^5$ such that

$$\begin{aligned} (**) \quad E(||Z||^p) &\le C_p^3 E(|\int_0^T Z_s dW_s|^p) \\ &\le C_p^4 E(|\xi + \int_0^T f(\cdot, s, y_s, z_s) ds - Y_0|^p) \\ &\le C_p^5 E(|\xi|^p + T^{\frac{p}{2}} [\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds]^{\frac{p}{2}}). \end{aligned}$$

Hence we obtain $Z \in H^p(\mathbf{R})$, and therefore $(Y, Z) \in B^p(\mathbf{R})$. The inequalities (i) and (ii) have also been established.

b) **We prove:** (iii). The solution $(\delta Y, \delta Z)$ belongs to the generator $f(\cdot, t, y_t^1, z_t^1) - f(\cdot, t, y_t^2, z_t^2)$, and $\xi = 0$. Therefore (i) and (ii), as well as an appeal to the Lipschitz condition, give, with universal constants C_p^6, C_p^7

$$\begin{aligned} ||(\delta Y, \delta Z)||_p^p &\le C_p^6 T^{\frac{p}{2}} E([\int_0^T |f(\cdot, t, y_t^1, z_t^1) - f(\cdot, t, y_t^2, z_t^2)|^2 dt]^{\frac{p}{2}}) \\ &\le C_p^7 T^{\frac{p}{2}} [||\delta y||_{S^p}^p + ||\delta z||_p^p] \\ &= C_p^7 T^{\frac{p}{2}} ||(\delta y, \delta z)||_p^p. \end{aligned}$$

This completes the proof. •

Let us return to the proof of Theorem 11.1. We define approximations of the solution of the BSDE recursively. Let for $k \geq 0, 0 \leq t \leq T$

$$\begin{aligned} Y^0 &= Z^0 = 0, \\ Y_t^{k+1} &= \xi - \int_t^T Z_s^{k+1} dW_s + \int_t^T f(\cdot, s, Y_s^k, Z_s^k) ds. \end{aligned}$$

We show:

$$||(Y^k, Z^k) - (Y, Z)||_4 \to 0 \quad (k \to \infty).$$

Recall the universal constant C_4 from Lemma 11.1, (iii). We may (modulo repeating the argument finitely often on successive subintervals of $[0,T]$) assume that $[C_4 T^2]^{\frac{1}{4}} < 1$. With this condition, Lemma 11.1 implies that Γ is a contraction, and the solution (Y,Z) of the BSDE its unique fixed point in $B^4(\mathbf{R})$. From this observation, we obtain our assertion via the Cauchy sequence property of the approximate solutions which follows from

$$||(Y^k, Z^k) - (Y^l, Z^l)||_4 \leq \sum_{r=k+1}^{l} ||(Y^r, Z^r) - (Y^{r-1}, Z^{r-1})||_4 \to 0 \quad (k, l \to \infty).$$

2. **We prove by recursion on k:**

$$(Y^k, Z^k) \in \mathbf{L}_1^2(\mathbf{R}) \times \mathbf{L}_1^2(\mathbf{R}).$$

This is trivial for $k = 0$. Let it be guaranteed for k. According to the chain rule for the Malliavin derivative and our hypotheses concerning the standard parameters we know for $0 \leq t \leq T$

$$\xi + \int_t^T f(\cdot, s, Y_s^k, Z_s^k) ds \in D_1^2,$$

with Malliavin derivative

$$\begin{aligned} & D_s \xi + \int_t^T [\frac{\partial}{\partial y} f(\cdot, u, Y_u^k, Z_u^k)\, D_s Y_u^k + \frac{\partial}{\partial z} f(\cdot, u, Y_u^k, Z_u^k)\, D_s Z_u^k \\ + & \; D_s f(\cdot, u, Y_u^k, Z_u^k)] du. \end{aligned}$$

This is seen by discretizing the Lebesgue integral, using the chain rule, and then approximating by means of the boundedness properties of the partial derivatives, the Lipschitz continuity properties of f and closedness of the operator D. Consequently, Lemma 9.2 yields for fixed $0 \le t \le T$

$$Y_t^{k+1} = E(\xi + \int_t^T f(\cdot, s, Y_s^k, Z_s^k) ds | \mathcal{F}_t) \in D_1^2$$

as well. Consequently, also

$$\int_t^T Z_s^{k+1} dW_s = \xi + \int_t^T f(\cdot, s, Y_s^k, Z_s^k) ds - Y_t^{k+1} \in D_1^2.$$

Now an appeal to Theorem 9.4 implies that $Z^{k+1} \in \mathbf{L}_1^2(\mathbf{R})$ and in $L^2(\Omega \times [0,T]^2)$ we have the equation

$$\begin{aligned} D_s \int_t^T Z_u^{k+1} dW_u &= \int_t^T D_s Z_u^{k+1} dW_u, \quad s \le t, \\ D_s \int_t^T Z_u^{k+1} dW_u &= Z_s^{k+1} + \int_s^T D_s Z_u^{k+1} dW_u, \quad s > t. \end{aligned}$$

All stated differentiabilities go along with square integrability of the Malliavin derivatives in all variables. This means that

$$(Y^{k+1}, Z^{k+1}) \in \mathbf{L}_1^2(\mathbf{R}) \times \mathbf{L}_1^2(\mathbf{R}),$$

and the recursion step is completed. We also can identify the Malliavin derivative by the formula valid for $0 \le s \le t \le T$ in the usual sense

$$\begin{aligned} (***) \quad D_s X_t^{k+1} &= D_s \xi - \int_t^T D_s Z_u^{k+1} dW_u \\ &+ \int_t^T [\frac{\partial}{\partial y} f(\cdot, u, Y_u^k, Z_u^k)\, D_s Y_u^k + \frac{\partial}{\partial z} f(\cdot, u, Y_u^k, Z_u^k)\, D_s Z_u^k \\ &\qquad + D_s f(\cdot, u, Y_u^k, Z_u^k)] du. \end{aligned}$$

3. In this step **we show:**

$$(DY^k, DZ^k) \to (Y^\cdot, Z^\cdot) \text{ in } L^2(\Omega \times [0,T]^2),$$

where for $0 \le s \le T$ (Y^s, Z^s) is the solution of the BSDE

$$\begin{aligned} (\natural) Y_t^s &= D_s \xi - \int_t^T Z_u^s dW_u + \int_t^T [\frac{\partial}{\partial y} f(\cdot, u, Y_u, Z_u)\, Y_u^s + \frac{\partial}{\partial z} f(\cdot, u, Y_u, Z_u)\, Z_u^s \\ &\qquad + D_s f(\cdot, u, Y_u, Z_u)] du, \quad 0 \le s \le t \le T, \\ Y_t^s &= Z_t^s = 0, \quad 0 \le t < s \le T. \end{aligned}$$

We first consult our hypotheses to verify that, at least for λ-a.e. $0 \le s \le T$ the parameters $(F^s, D_s\xi)$ with

$$F^s(\cdot,t,y,z) = \frac{\partial}{\partial y} f(\cdot,t,Y_t,Z_t)\, y + \frac{\partial}{\partial z} f(\cdot,t,Y_t,Z_t)\, z + D_s f(\cdot,t,Y_t,Z_t),$$

$0 \le t \le T, y, z \in \mathbf{R}$, are standard. Hence (Y^s, Z^s) is well defined (and set trivial on the set of s where the parameters eventually fail to be standard). Also in this case our arguments will be based on *a priori inequalities.*

Lemma 11.2 *Let* $(f^i, \xi^i), i = 1,2,$ *be standard parameters of a BSDE,* $p \ge 2$. *Suppose*

$$\xi^i \in L^p(\Omega), \quad f^i(\cdot,0,0) \in H^p(\mathbf{R}), \quad i = 1,2.$$

Let $(Y^i, Z^i) \in B^p(\mathbf{R})$ *be the corresponding solutions,* C *a Lipschitz constant for* f^1. *Put*

$$\delta Y = Y^1 - Y^2, \quad \delta Z = Z^1 - Z^2, \quad \delta_2 f_t = f^1(\cdot,t,Y_t^2,Z_t^2) - f^2(\cdot,t,Y_t^2,Z_t^2),$$

$0 \le t \le T$. *Then for* T *small enough there exists a constant* $C_{p,T}$ *such that*

$$\begin{aligned} ||(\delta Y, \delta Z)||_p^p &\le C_{p,T}[E(|\delta Y_T|^p) + E((\int_0^T |\delta_2 f_s| ds)^p)] \\ &\le C_{p,T}[E(|\delta Y_T|^p) + T^{\frac{p}{2}}||\delta_2 f_s||^p]. \end{aligned}$$

Proof
With a calculation analogous to the one used to prove (i) and (ii) in Lemma 11.1 we arrive at the following inequality which is valid with universal constants $C_p^1, \cdot, C_p^3$, and for which we also use Doob's and Cauchy-Schwarz' inequalities,

$$\begin{aligned} ||\delta Y||_{S^p}^p + ||\delta Z||_p^p \le\ & C_p^1 E(|\delta Y_T|^p \\ & +(\int_0^T |f^1(\cdot,t,Y_t^1,Z_t^1) - f^2(\cdot,t,Y_t^2,Z_t^2)| dt)^p) \\ & \le C_p^2 E(|\delta Y_T|^p + (\int_0^T [|\delta Y_s| + |\delta Z_s| + |\delta_2 f_s|] ds)^p) \\ & \le C_p^3 [E(|\delta Y_T|^p \\ & +(\int_0^T |\delta_2 f_s| ds)^p) + (T^p ||\delta Y||_{S^p}^p + T^{\frac{p}{2}} ||\delta Z||_p^p)]. \end{aligned}$$

Now choose T small enough to ensure $C_p^3(T^p+T^{\frac{p}{2}}) < 1$. This being done, we may take the last two expressions in the previous inequality from the right to the left hand side, to obtain the desired estimate. •

Let us now apply Lemma 11.2 to prove that for λ-a.a. $0 \leq s \leq T$ we have $(Y^s, Z^s) \in B^2(\mathbf{R})$. For this purpose, we apply the Lemma with

$$\begin{aligned} Y^1 &= Y^s, Y^2 = 0, \\ \xi^1 &= D_s\xi, \xi^2 = 0, \\ f^1(\cdot, t, y, z) &= (\frac{\partial}{\partial y} f(\cdot, t, Y_t, Z_t)\, y + \frac{\partial}{\partial z} f(\cdot, t, Y_t, Z_t)\, z + D_s f(\cdot, t, Y_t, Z_t), \\ f^2 &= 0, \end{aligned}$$

$0 \leq t \leq T, y, z \in \mathbf{R}$. Then we have

$$\delta_2 f_t = D_s f(\cdot, t, Y_t, Z_t), \quad 0 \leq t \leq T.$$

We obtain with some universal constant C the inequality

$$||(Y^s, Z^s)||_2^2 \leq CE(|D_s\xi|^2 + ||D_s f(\cdot, Y, Z)||^2),$$

and therefore

$$\int_0^T ||(Y^s, Z^s)||_2^2 ds < \infty.$$

This implies the desired integrability.

To obtain estimates for differences of (DY^k, DZ^k) and $(Y^\cdot, Z^\cdot)$, let us next, fixing $k \in \mathbf{N}$, apply Lemma 11.2 to the following parameters

$$\begin{aligned} \xi^1 &= \xi^2 = D_s\xi, \\ f^1(t) &= \frac{\partial}{\partial y} f(\cdot, t, Y_t^k, Z_t^k)\, D_s Y_t^k + \frac{\partial}{\partial z} f(\cdot, t, Y_t^k, Z_t^k)\, D_s Z_t^k + D_s f(\cdot, t, Y_t^k, Z_t^k), \\ f^2(t) &= \frac{\partial}{\partial y} f(\cdot, t, Y_t, Z_t)\, Y_t^s + \frac{\partial}{\partial z} f(\cdot, t, Y_t, Z_t)\, Z_t^s + D_s f(\cdot, t, Y_t, Z_t), \end{aligned}$$

$s \leq t \leq T$. Set for abbreviation

$$\delta_t^k = f^1(t) - f^2(t), \quad s \leq t \leq T.$$

The Lemma yields the inequality

$$||(D_s Y^{k+1} - Y^s, D_s Z^{k+1} - Z^s)||_2^2 \leq C_1 E((\int_s^T |\delta_t^k| dt)^2)$$

with a universal constant C_1. Let us now further estimate the right hand side of this inequality. We have, fixing $0 \leq s \leq T$

$$E((\int_s^T |\delta_t^k| dt)^2) \leq C_2[A_k^s(T) + B_k^s(T) + C_k^s(T)],$$

where

$$\begin{aligned}
A_k^s(T) &= E([\int_s^T |D_s f(\cdot, t, Y_t, Z_t) - D_s f(\cdot, t, Y_t^k, Z_t^k)| dt]^2), \\
B_k^s(T) &= E([\int_s^T |\frac{\partial}{\partial y} f(\cdot, t, Y_t^k, Z_t^k)(Y_t^s - D_s Y_t^k)| dt]^2) \\
&+ E([\int_s^T |\frac{\partial}{\partial z} f(\cdot, t, Y_t^k, Z_t^k)(Z_t^s - D_s Z_t^k)| dt]^2), \\
C_k^s(T) &= E([\int_s^T |\frac{\partial}{\partial y} f(\cdot, t, Y_t, Z_t) - \frac{\partial}{\partial y} f(\cdot, t, Y_t^k, Z_t^k)| \, |Y_t^s| dt]^2) \\
&+ E([\int_s^T |\frac{\partial}{\partial z} f(\cdot, t, Y_t, Z_t) - \frac{\partial}{\partial z} f(\cdot, t, Y_t^k, Z_t^k)| \, |Z_t^s| dt]^2).
\end{aligned}$$

With further universal constants C_3, C_4 we deduce, using (H4)

$$\begin{aligned}
A_k^s(T) &\leq E([\int_s^T K_s(t)[|Y_t - Y_t^k| + |Z_t - Z_t^k|] dt]^2) \\
&\leq C_3 E(\int_s^T K_s(t)^2 dt [\int_s^T |Y_t - Y_t^k|^2 dt + \int_s^T |Z_t - Z_t^k|^2 dt]) \\
&\leq C_4 [E(\int_s^T K_s(t)^4 dt]^{\frac{1}{2}} [E(\int_s^T |Y_t - Y_t^k|^4 dt)^{\frac{1}{2}} + E(\int_s^T |Z_t - Z_t^k|^4 dt)^{\frac{1}{2}}].
\end{aligned}$$

Hence by part 1. of the proof

$$\lim_{k \to \infty} \int_0^T A_k^s(T) ds = 0.$$

Furthermore, since the partial derivatives of f with respect to y, z are bounded and continuous, and since $E(\int_0^T ||(Y^s, Z^s)||^2 ds < \infty$, dominated convergence allows to conclude

$$\lim_{k \to \infty} \int_0^T C_k^s(T) ds = 0.$$

Let us finally discuss the convergence of the $B_k^s(T)$ as $k \to \infty$. Again by boundedness of the partial derivatives of f we obtain with a universal constant C_5

$$B_k^s(T) \leq C_5 T^2 ||D_s Y^k - Y^s, D_s Z^k - Z^s||_2^2.$$

Now choose T small enough to ensure $\alpha = C_5 T^2 < 1$. Let $\varepsilon > 0$. Then by what has been shown there exists $N \in \mathbf{N}$ large enough so that for $k \geq N$ we have

$$E(\int_0^T ||(D_s Y^{k+1} - Y^s, D_s Z^{k+1} - Z^s)||^2 ds)$$
$$\leq \varepsilon + \alpha E(\int_0^T ||(D_s Y^k - Y^s, D_s Z^k - Z^s)||^2 ds).$$

By recursion we obtain for $k \geq N$

$$E(\int_0^T ||(D_s Y^k - Y^s, D_s Z^k - Z^s)||^2 ds) \leq \varepsilon(1 + \alpha + \alpha^2 + \cdots + \alpha^{k-N-1})$$
$$+\varepsilon\alpha^{k-N} E(\int_0^T ||(D_s Y^N - Y^s, D_s Z^N - Z^s)||^2 ds)$$
$$\leq \frac{\varepsilon}{1-\alpha} + \alpha^{k-N} E(\int_0^T ||(D_s Y^N - Y^s, D_s Z^N - Z^s)||^2 ds).$$

Now let $k \to \infty$. Since ε is arbitrary, we conclude

$$\lim_{k\to\infty} \int_0^T B_k^s(T) ds = 0.$$

3. Since $\mathbf{L}_1^2(\mathbf{R})$ is a Hilbert space, and D is a closed operator, we obtain that $(Y, Z) \in \mathbf{L}_1^2(\mathbf{R}) \times \mathbf{L}_1^2(\mathbf{R})$, and that $(Y^s, Z^s)_{0\leq s\leq T}$ is a version of $(D_s Y, D_s Z)_{0\leq s\leq T}$ in the usual sense.

4. **We show:**

$$(D_t Y_t)_{0\leq t\leq T} \text{ is a version of } (Z_t)_{0\leq t\leq T}.$$

For $t \leq s$ we have

$$Y_s = Y_t + \int_t^s Z_r dW_r - \int_t^s f(\cdot, r, Y_r, Z_r) dr.$$

Hence by Theorem 9.4 for $t < u \leq s$

$$\begin{aligned} D_u Y_s &= Z_u + \int_u^s D_u Z_r dW_r \\ &\quad - \int_u^s [\frac{\partial}{\partial y} f(\cdot, r, Y_r, Z_r)\, D_u Y_r + \frac{\partial}{\partial z} f(\cdot, r, Y_r, Z_r)\, D_u Z_r \\ &\quad + D_u f(\cdot, r, Y_r, Z_r)] dr. \end{aligned}$$

By continuity in t of (Y^s, Z^s) we may choose $u = s$, to obtain the desired identity. •

Literatur

[1] Amendinger, J., Imkeller, P., Schweizer, M. *Additional logarithmic utility of an insider.* Stoch. Proc. Appl. 75 (1998), 263-286.

[2] S. Ankirchner, P. Imkeller, A. Popier *On measure solutions of backward stochastic differential equations.* Preprint, HU Berlin (2008).

[3] S. Ankirchner, P. Imkeller, G. Reis *Classical and Variational Differentiability of BSDEs with Quadratic Growth.* Electron. J. Probab. 12 (2007), 1418–1453 (electronic).

[4] S. Ankirchner, P. Imkeller, G. Reis *Pricing and hedging of derivatives based on non-tradable underlyings.* Preprint, HU Berlin (2007).

[5] S. Ankirchner, P. Imkeller *Quadratic hedging of weather and catastrophe risk by using short term climate predictions.* Preprint, HU Berlin (2008).

[6] Bell, D. *The Malliavin calculus.* Pitman Monographs and Surveys in Pure and Applied Math. 34. Longman and Wiley: 1987.

[7] Bell, D. *Degenerate stochastic differential equations and hypoellipticity.* Pitman Monographs and Surveys in Pure and Applied Mathematics. 79. Harlow, Essex: Longman (1995).

[8] Bismut, J. M. *Martingales, the Malliavin Calculus and Hypoellipticity under General Hörmander's Conditions.* Z. Wahrscheinlichkeitstheorie verw. Geb. 56 (1981), 469-505.

[9] Bismut, J. M. *Large deviations and the Malliavin calculus.* Progress in Math. 45, Birkhäuser, Basel 1984.

[10] Bouleau, N., Hirsch, F. *Dirichlet forms and analysis on Wiener space.* W. de Gruyter, Berlin 1991.

[11] El Karoui, N.; Peng, S.; Quenez, M.C. Backward stochastic differential equations in finance. Math. Finance 7, No.1, 1-71 (1997).

[12] Fournié, E., Lasry, J.-M., Lebouchoux, J., Lions, P.-L., Touzi, N. *Applications of Malliavin's calculus to Monte Carlo methods in finance.* Finance Stochast. 3 (1999), 391-412.

[13] Huang, Z-Y.; Yan, J-A. *Introduction to infinite dimensional stochastic analysis.* Mathematics and its Applications (Dordrecht). 502. Dordrecht: Kluwer Academic Publishers (2000).

[14] Ikeda, N., Watanabe, S. *Stochastic differential equations and diffusion processes.* North Holland (2 nd edition): 1989.

[15] Imkeller, P. *Malliavin's calculus in insider models: additional utility and free lunches.* Preprint, HU Berlin (2002).

[16] Imkeller, P., Pontier, M., Weisz, F. *Free lunch and arbitrage possibilities in a financial market model with an insider.* Stochastic Proc. Appl. 92 (2001), 103-130.

[17] Imkeller, P. *Enlargement of the Wiener filtration by an absolutely continuous random variable via Malliavin's calculus.* PTRF 106(1996), 105-135.

[18] Imkeller, P., Perez-Abreu, V., Vives, J. *Chaos expansions of double intersection local time of Brownian motion in* $\mathbf{R}^d$ *and renormalization.* Stoch. Proc. Appl. 56(1994), 1-34.

[19] Jacod, J. *Grossissement initial, hypothèse (H'), et théorème de Girsanov.* in: Grossissements de filtrations: exemples et applications. T. Jeulin, M.Yor (eds.). LNM 1118. Springer: Berlin 1985.

[20] Janson, Svante *Gaussian Hilbert spaces.* Cambridge Tracts in Mathematics. 129. Cambridge: Cambridge University Press (1997).

[21] Kuo, H. H. *Donsker's delta function as a generalized Brownian functional and its application.* Lecture Notes in Control and Information Sciences, vol. 49 (1983), 167-178. Springer, Berlin.

[22] Kusuoka, S. , Stroock, D.W. *The partial Malliavin calculus and its applications to nonlinear filtering.* Stochastics 12 (1984), 83-142.

[23] Ma, J., and Yong, J. *Forward-Backward Stochastic Differential Equations and Their Applications.* LNM 1702. Springer: Berlin (1999).

[24] Malliavin, P. Stochastic analysis. Grundlehren der Mathematischen Wissenschaften. 313. Berlin: Springer (1997).

[25] Malliavin, P. Integration and probability. Graduate Texts in Mathematics. 157. New York, NY: Springer (1995).

[26] Malliavin, P. *Stochastic calculus of variation and hypoelliptic operators.* Proc. Intern. Symp. SDE Kyoto, 1976, 195-263. Kinokynia: Tokyo 1978.

[27] Malliavin, P. C^k-*hypoellipticity with degeneracy.* Stochastic Analysis. ed: A. Friedman and M. Pinsky, 199-214, 327-340. Acad. Press: New York 1978.

[28] **Malliavin, P., Thalmaier, A. Stochastic calculus of variations in mathematical finance. Springer: Berlin 2006.**

[29] Norris, J. *Simplified Malliavin calculus.* Séminaire de Probabilités XX. LNM 1204, 101-130. Springer: Berlin 1986.

[30] Nualart, D., Pardoux, E. *Stochastic calculus with anticipating integrands.* Probab. Th. Rel. Fields 78 (1988), 535-581.

[31] Nualart, D. The Malliavin calculus and related topics. Springer: Berlin 1995.

[32] Nualart, D. Malliavin's calculus and anticipative calculus. St Flour Lecture Notes (1996).

[33] Nualart, D., Vives, J. *Chaos expansion and local times.* Publicacions Matematiques 36(2), 827-836.

[34] Revuz, D., Yor, M. *Continuous martingales and Brownian motion.* Springer, Berlin 1991.

[35] Stroock, D. W. *The Malliavin Calculus. Functional Analytic Approach.* J. Funct. Anal. 44 (1981), 212-257.

[36] Stroock, D. W. *The Malliavin calculus and its applications to second order parabolic differential equations.* Math. Systems Theory, Part I, 14 (1981), 25-65, Part II, 14 (1981), 141-171.

[37] Stroock, D. W. *Some applications of stochastic calculus to partial differential equations.* In: Ecole d'Eté de Probabilité de St Flour. LNM 976 (1983), 267-382.

[38] Watanabe, H. *The local time of self-intersections of Brownian motions as generalized Brownian functionals.* Letters in Math. Physics 23 (1991), 1-9.

[39] Watanabe, S. *Stochastic differential equations and Malliavin calculus.* Tata Institut of Fundamental Research. Springer: Berlin 1984.

[40] Yor, M. *Grossissement de filtrations et absolue continuité de noyaux.* in: Grossissements de filtrations: exemples et applications. T. Jeulin, M.Yor (eds.). LNM 1118. Springer: Berlin 1985.

[41] Yor, M. *Entropie d'une partition, et grossissement initial d'une filtration.* in: Grossissements de filtrations: exemples et applications. T. Jeulin, M.Yor (eds.). LNM 1118. Springer: Berlin 1985.