Prof. Dr. Peter Imkeller
Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin
imkeller@mathematik.hu-berlin.de

Mathematics Subject Classification (2000): Primary 60H07; Secondary 60H05, 60H30

ISBN-13: 978-83-86806-02-7
Printed in Poland

For additional information and updates on this book, visit the IMPAN Publications home page at http://journals.impan.pl/ln/

Contents

Introduction

Malliavin's calculus has been developed to provide a stochastic approach to Hörmander's theorem (see Bell [5], [6], Bismut [7], [8], Bouleau–Hirsch [9], Malliavin [26], [27], Stroock [37], [38], [39], Watanabe [40]). The latter states that parabolic partial differential equations on manifolds generated by vector fields whose Lie algebra creates action in all directions possess smooth solutions. Malliavin's original idea to construct an Ornstein–Uhlenbeck dynamics on the Wiener space for this purpose, in subsequent work by Bismut, Stroock, Watanabe and others, has led to a version of a stochastic calculus of variations on Wiener space. Its basic notion of Fréchet differentiability of random variables leads to criteria of smoothness of these variables' laws. In the context of stochastic differential equations corresponding to given parabolic partial differential equations the criteria are seen to be applicable provided Hörmander's Lie algebra conditions are satisfied by the underlying vector fields. For an account of this see Norris [31], Nualart [32], [33].

In the meantime the stochastic calculus of variations has proved its central importance to stochastic and applied analysis in a variety of different applications. Recent fields of application include the smoothness problem for solutions of stochastic partial differential equations such as the stochastic Navier–Stokes equation (see Mattingly–Pardoux [29]). Malliavin's calculus has proved to be very efficient in enhancing the speed of algorithms in the numerical treatment of stochastic equations, in particular in stochastic finance (see Fournié et al. [12], Kusuoka–Stroock [22], Malliavin–Thalmayer [28] for more references); it was also seen to play a significant role in problems of statistical analysis (see Privault–Réveillac [35]). It has been used in the investigation of fine structure properties of fractional stochastic processes and their functionals (see for example Mishura [30], Kuo [21], Imkeller et al. [18]), in anticipative calculus (see Nualart–Pardoux [34]) and in financial market models with asymmetric information structures (see Imkeller [16], [17], Imkeller et al. [19]). It has been known since Stroock, and underpinned in the formula by Clark, Haussmann and Ocone, to provide an explicit description of the stochastic integrand in the martingale representation theorem on Wiener spaces. This connection was deepened in work by El Karoui, Peng and Quenez to include explicit representations of the solutions of backward stochastic differential equations (BSDE) on a Wiener basis, and thus serve as a natural tool in stochastic optimization and control (see El Karoui et al. [10], Ma–Yong [23]). On this track, in this course of lectures I aim at explaining this connection, and apply it to some recently discovered results about the fine structure of option pricing and hedging in incomplete finance or insurance markets (see Ankirchner et al. [1], [2], [3], [4]).

Malliavin smoothness is nothing but an infinite-dimensional version of everybody's notion of smoothness from classical calculus on finite-dimensional Euclidean space. This is seen by analyzing the canonical Wiener space by means of any orthonormal basis of $L^2(\mathbb{R}_+)$ as an infinite-dimensional sequence space with the product topological, measurable and smoothness structures. The projection of these structures on finite-dimensional Euclidean space just reproduces the usual notion of smoothness in the Leibniz–Newton–Sobolev sense. I take this observation to define the starting position of this course. From there on I shall develop Malliavin's calculus by extending the classical differential calculus on Sobolev spaces stepwise from the finite-dimensional to infinite-dimensional sequence space, and then on to the canonical Wiener space. The natural dual of the derivative operator with respect to the Hilbert space structure created by Wiener measure will let emerge Skorokhod and Itô integrals as well as the Ornstein–Uhlenbeck operator. Once at this end of the calculus path, starting from the martingale optimality principle, we will first carefully develop a class of BSDE whose solutions provide a simple access to problems of utility optimization with not necessarily convex constraints. Then we shall develop the link between the calculus of variations and stochastic control theory by giving the representation of the control variable in the solution notion of BSDE by a Malliavin trace. The material of the course is organized in the following way.

After linking the canonical Wiener space to an infinite-dimensional Gaussian sequence space by a universal isomorphism, we develop in Chapter 1 the differential calculus on this Gaussian sequence space in several steps, partly following Malliavin [25]. We begin by explaining Malliavin's smoothness problem of laws in the one-dimensional Euclidean setting, then use the differential operator and its dual obtained in this setting to develop a Gaussian Sobolev calculus first for finite-dimensional Euclidean space. By exploiting a martingale-theoretic argument, it is then extended to the infinite-dimensional sequence space. In the background, we exhibit the calculus more explicitly in the case of integrability degree 2. Here Sobolev spaces are Hilbert, and the generalized Hermite polynomials provide an orthonormal basis.

Chapter 2 is devoted to developing the tools of Malliavin's calculus needed for later control-theoretic applications on canonical Wiener path space, along the lines of Nualart [32]. We first use the universal isomorphism to transfer the differential calculus of Chapter 1 to the canonical path space. This will provide the usual notion of differentiability in the sense of Malliavin's calculus and its Sobolev spaces. The dual operator, providing Skorokhod's integral, and, in the case of adapted integrands, the classical Itô integral, are investigated (for some background on stochastic calculus see Huang–Yan [13], Ikeda–Watanabe [14], Malliavin [24], and Revuz–Yor [36]). The orthonormal Hermite polynomial basis is translated into the somewhat more flexible language of multiple Wiener–Itô integrals on path space. Finally, to be able to treat smoothness questions for stochastic differential equations, we discuss differentiation for Itô integrals and integral processes.

Chapter 3 is entirely devoted to BSDE and their interpretation by means of Malliavin's calculus. We first discuss, following Hu et al. [15] the utility maximization problem on incomplete financial markets with exponential utility, formulate it in terms of the martingale optimality principle, and solve this problem by means of BSDE of quadratic type. We subsequently consider BSDE with Lipschitz generators, and carefully develop a theory of existence and uniqueness of solutions based on a priori inequalities, following El Karoui, Peng and Quenez [10]. In the last step, we interpret the solution pairs of those equations by means of the stochastic calculus of variations.

Acknowledgements. It is my special pleasure to thank an anonymous referee for valuable and constructive comments. I am sure they have been of great help in improving the readability of these notes. I have very good memories of the course period in Warsaw, especially the sessions at IMPAN. This is due to a very lively, knowledgeable and faithful audience, and to the wonderful hospitality of Feliks Przytycki, Łukasz Stettner and Jerzy Zabczyk.

1. Calculus on the sequence space

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 by 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. We follow Malliavin [25].

1.1. The Wiener space as a sequence space

DEFINITION 1.1.1. A probability space $(\Omega, \mathscr{F}, P)$ is called *Gaussian* if there is a family $(X_k)_{1\leq k\leq n}$ or a sequence $(X_k)_{k\in\mathbb{N}}$ of independent Gaussian unit random variables such that

$$\mathscr{F} = \sigma(X_k : 1 \leq k \leq n) \quad \text{resp.} \quad \sigma(X_k : k \in \mathbb{N})$$

(completed by sets of P-measure 0).

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

CLAIM. *$(\Omega, \mathscr{F}, P)$ is Gaussian.*

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

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

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

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

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

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

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

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

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

$$\mathscr{F} = \sigma(W^j(g_i) : i \in \mathbb{N}, 1 \le j \le m).$$

Moreover, since

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

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 ultimately interested in infinite-dimensional spaces, from now on we make the following

ASSUMPTION. The Gaussian space considered is generated by *infinitely many independent Gaussian unit variables.*

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

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

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

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

For $n \in \mathbb{N}$ let

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

This notation is consistent for $n = 1$.

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

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

is $\mathscr{F}$-measurable on Ω.

LEMMA 1.1.1. *Let F be $\mathscr{F}$-measurable on Ω. Then there exists a $\mathbf{B}^{\mathbb{N}}$-measurable function f on $\mathbb{R}^{\mathbb{N}}$ such that*

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

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

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

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

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

THEOREM 1.1.1. *Let $p \ge 1$. Then the mapping*

$$L^p(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu) \ni f \mapsto F = f \circ ((X_n)_{n\in\mathbb{N}}) \in L^p(\Omega, \mathscr{F}, P) \tag{1.1}$$

defines a linear isomorphism.

Proof. The mapping is well defined, since

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

and bijective by Lemma 1.1.1. Linearity is trivial. ∎

Theorem 1.1.1 allows us to develop a differential calculus on the sequence space $(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu)$, and then to transfer it to the canonical space $(\Omega, \mathscr{F}, P)$. For this purpose we are stimulated by the treatment of the one-dimensional situation.

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

1.2. Absolute continuity of measures on $\mathbb{R}$

Our aim is to study laws of random variables defined on $(\Omega, \mathscr{F}, P)$, i.e. the probability measures P_X for random variables X. By means of Theorem 1.1.1 these measures correspond to the measures $\nu \circ f^{-1}$ for $\mathbf{B}^{\mathbb{N}}$-measurable functions f on $\mathbb{R}^{\mathbb{N}}$. The one-dimensional version of these measures is given by $\nu_1 \circ f^{-1}$ for $\mathbf{B}^1$-measurable functions f defined on $\mathbb{R}$.

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

LEMMA 1.2.1. *Let μ be a finite measure on $\mathbf{B}^1$. Suppose there exists $c \in \mathbb{R}$ such that for all $\phi \in C^1(\mathbb{R})$ with bounded derivative we have*

$$\left|\int \phi'(x)\,\mu(dx)\right| \le c\|\phi\|_\infty.$$

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

Proof. Let $f \ge 0$ be continuous with compact support, and define

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

Then

$$\int f\,d\mu = \int \phi'(x)\,\mu(dx) \le c\|\phi\|_\infty = c\int f\,d\lambda.$$

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

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

which clearly implies $\mu \ll \lambda$. ∎

We aim at applying Lemma 1.2.1 to the probability measure $\nu_1 \circ f^{-1}$ with f $\mathbf{B}^1$-measurable. In this connection 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 $g, h \in C_0^\infty(\mathbb{R})$ (smooth functions with compact support) we have

$$\langle g', h\rangle = \int g'(x)h(x)\,dx = -\int g(x)h'(x)\,dx = -\langle g, h'\rangle. \tag{1.2}$$

This relationship can be extended to functions $g, h \in L^2(\mathbb{R})$ which vanish at $\pm\infty$ and which possess derivatives in the distributional sense. Let us for the moment assume this setting and denote by dg the *distributional derivative* of g, and by δh its *adjoint operator* in the sense of the duality (1.2). Then for $h \in C_0^\infty(\mathbb{R})$ we have

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

and we can interpret the duality relationship as

$$\langle dg, h\rangle = \langle g, \delta h\rangle. \tag{1.3}$$

Finally, the operator $\delta d = -d^2/dx^2$ 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 $g,h\in L^2(\mathbb{R},\nu_1)$ define

$$\langle g\,|\,h\rangle=\int g(x)h(x)\nu_1(dx). \tag{1.4}$$

To apply Lemma 1.1.1 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,

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

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

$$\begin{aligned}\langle g'|h\rangle&=\frac{1}{\sqrt{2\pi}}\int g'(x)h(x)\exp(-x^2/2)\,dx\\&=-\frac{1}{\sqrt{2\pi}}\int g(x)\frac{d}{dx}[h(x)\exp(-x^2/2)]\,dx\\&=-\int g(x)\exp(x^2/2)\frac{d}{dx}[h(x)\exp(-x^2/2)]\,\nu_1(dx)\\&=\langle g\,|-h'+xh\rangle.\end{aligned} \tag{1.5}$$

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

$$\delta h=-h'+xh. \tag{1.6}$$

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

$$\langle dg\,|\,h\rangle=\langle g\,|\,\delta h\rangle. \tag{1.7}$$

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

$$L=\delta d=-\frac{d^2}{dx^2}+x\frac{d}{dx}, \tag{1.8}$$

in a suitable distributional sense.

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

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

LEMMA 1.2.2. *Let $g, h \in L^2(\mathbb{R}, \nu_1)$ be such that $dg, \delta h \in L^2(\mathbb{R}, \nu_1)$. Then*

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

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

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

Proof. We continue the above calculation in the notation chosen. By duality and the Cauchy–Schwarz inequality we have

$$\begin{aligned} \left|\left\langle d(\phi \circ f) \,\middle|\, \frac{1}{df} \right\rangle\right| &= \left|\left\langle \phi \circ f \,\middle|\, \delta\left(\frac{1}{df}\right) \right\rangle\right| \\ &= \left|\int \phi \circ f(x) \delta\left(\frac{1}{df}\right)(x) \nu_1(dx)\right| \\ &\leq \|\phi\|_\infty \left\|\delta\left(\frac{1}{df}\right)\right\|_2. \end{aligned}$$

Hence Lemma 1.1.1 can be applied with $c = \left\|\delta\left(\frac{1}{df}\right)\right\|_2$, 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, for brevity, mostly concentrate on the operators d and δ. One natural orthonormal basis of $L^2(\mathbb{R}, \nu_1)$ proves to be very useful here.

1.3. Hermite polynomials and orthogonal developments

We continue denoting by d, δ and L the operators studied above. They are at least well defined on $C_0^\infty(\mathbb{R})$ (and this is the sense in which we use them), and even, by the integrability properties of the Gaussian density, on the space of polynomials in one real variable.

DEFINITION 1.3.1. For $n \geq 0$ let

$$H_n = \delta^n 1. \tag{1.9}$$

H_n is called the *Hermite polynomial* of degree n.

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

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

THEOREM 1.3.1. *H_n is a polynomial of degree n, with leading coefficient* 1. *Moreover, for $n \in \mathbb{N}$,*

$$\delta H_n = H_{n+1}, \tag{1.10}$$
$$dH_n = nH_{n-1}, \tag{1.11}$$
$$LH_n = nH_n. \tag{1.12}$$

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

Proof. (1.10) follows by definition.

To prove (1.11), we first compute the commutator of d and δ. For $h \in C_0^\infty(\mathbb{R})$ we have

$$\begin{aligned}(d\delta - \delta d)h &= d(-h' + xh) - (-h'' + xh') \\ &= -h'' + xh' + h - (-h'' + xh') = h.\end{aligned}$$

This means that

$$d\delta - \delta d = \mathrm{id}.$$

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

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

Finally, (1.12) follows from $LH_n = \delta dH_n = n\delta H_{n-1} = nH_n$. ∎

COROLLARY 1.3.1. *For $g \in L^2(\mathbb{R})$ define the* Fourier transform *by*

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

Then

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

Proof. Choose $u \in \mathbb{R}$. Then

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

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

THEOREM 1.3.2. $\left(\frac{1}{\sqrt{n!}}H_n\right)_{n\geq 0}$ *is an orthonormal basis of* $L^2(\mathbb{R},\nu_1)$.

Proof. 1. Let $n,k\in\mathbb{N}$, and suppose that $n<k$. Then by Theorem 1.3.1,

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

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 $L^2(\mathbb{R},\nu_1)$, i.e. the set of linear combinations of Hermite polynomials is dense in $L^2(\mathbb{R},\nu_1)$. For this purpose, it suffices to prove that if $\phi\in L^2(\mathbb{R},\nu_1)$ satisfies $\langle H_n\,|\,\phi\rangle=0$ for all $n\geq 0$, then $\phi=0$.

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

$$F(z)=\int_{\mathbb{R}}\phi(v)e^{ivz-v^2/2}\,dv.$$

Then for $k\in\mathbb{N}$ and $t\in\mathbb{R}$ by Cauchy–Schwarz we have

$$\int_{\mathbb{R}}|v^k\phi(v)|e^{-vt-v^2/2}\,dv\leq\left[\int_{\mathbb{R}}\phi^2(v)e^{-v^2/2}\,dv\int_{\mathbb{R}}v^{2k}e^{-2vt-v^2/2}\,dv\right]^{1/2}<\infty.$$

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

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

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 $\mathbb{R}^{\mathbb{N}}$, the sequence space version of our infinite-dimensional Gaussian space. Our task will be to establish in this space suitable analogues 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 1.3.2. For $n\in\mathbb{N}$ let $E_n=\mathbb{Z}_+^n$, and let E be the set of sequences in $\mathbb{Z}_+$ with finitely many nonzero terms. For $p=(p_1,\ldots,p_k,0,\ldots)\in E$, let $|p|=\sum_{i=1}^k p_i$ and $p!=\prod_{i=1}^k p_i!$. For $x\in\mathbb{R}^k$ resp. $x\in\mathbb{R}^{\mathbb{N}}$, and $p\in E_k$

resp. $p \in E$ let

$$H_p(x) = \prod_{i=1}^{k} H_{p_i}(x_i) \quad \text{resp.} \tag{1.13}$$

$$= \prod_{i\in\mathbb{N}} H_{p_i}(x_i). \tag{1.14}$$

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

We can now extend Theorem 1.3.2 to the multidimensional setting.

THEOREM 1.3.3. *$\big(\frac{1}{\sqrt{p!}}H_p\big)_{p\in E_k}$ is an orthonormal basis of $L^2(\mathbb{R}^k, \mathbf{B}^k, \nu_k)$, and $\big(\frac{1}{\sqrt{p!}}H_p\big)_{p\in E}$ is an orthonormal basis of $L^2(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu)$.*

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

$$\langle g \,|\, h\rangle = \int_{\mathbb{R}^k} g(x)h(x)\,\nu_k(dx). \tag{1.15}$$

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

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

Hence $\big(\frac{1}{\sqrt{p!}}H_p\big)_{p\in E_k}$ is an orthonormal system. Moreover, the linear combinations of tensor products of functions of one of k variables are dense in $L^2(\mathbb{R}^k, \mathbf{B}^k, \nu_k)$. Hence the first claim follows from Theorem 1.3.2.

2. The set $\bigcup_{n\in\mathbb{N}} \pi_n^{-1}[L^2(\mathbb{R}^n, \mathbf{B}^n, \nu_n)]$ is dense in $L^2(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu)$. 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.

1.4. Finite-dimensional Gaussian Sobolev spaces

Let $k \in \mathbb{N}$. 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 $\mathbb{R}^k$.

DEFINITION 1.4.1. Let $p \geq 1$. For $f \in L^p(\mathbb{R}^k)$ and $a \in \mathbb{R}^k$, we say that f *has a directional (generalized) derivative in direction a* if there is a function $d_a f$ in $L^p(\mathbb{R}^k)$ such that

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

as $\varepsilon \to 0$. Let

$$W_1^p = \{f \in L^p(\mathbb{R}^k) : f \text{ has a directional derivative in direction } a \text{ for any } a \in \mathbb{R}^k\} \tag{1.16}$$

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

By linearity, it is clear that if for $1 \le i \le k$ we denote by $e_i \in \mathbb{R}^k$ the ith canonical basis vector, then for $f \in W_1^p(\mathbb{R}^k)$ and $a = (a_1, \ldots, a_k)$ we have $d_a f = \sum_{i=1}^k a_i d_{e_i} f$. We will write $d_i = d_{e_i}$.

DEFINITION 1.4.2. Let $p \ge 1$ and $s \in \mathbb{N}$. We define recursively

$$W_s^p = \{f \in W_1^p : d_a f \in W_{s-1}^p \text{ for any } a \in \mathbb{R}^k\} \tag{1.17}$$

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

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

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

$$\|f\|_{1,p} = \|f\|_p + \sum_{i=1}^k \|d_i f\|_p, \quad f \in W_1^p, \tag{1.18}$$

and analogous norms for higher order derivatives.

For any $p \ge 1$ and $s \in \mathbb{N}$ we have

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

and for $g \in C_0^\infty(\mathbb{R}^k)$ and $a = (a_1, \ldots, a_k) \in \mathbb{R}^k$ 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 the distributional sense" encountered above?

DEFINITION 1.4.3. Let $f \in L^1_{\text{loc}}(\mathbb{R}^k)$ and $a \in \mathbb{R}^k$. Then $u_a \in L^1_{\text{loc}}(\mathbb{R}^k)$ is called a *weak derivative* of f *in direction* a if for any $\phi \in C_0^\infty(\mathbb{R}^k)$ we have

$$\langle f, d_a \phi \rangle = -\langle u_a, \phi \rangle. \tag{1.19}$$

THEOREM 1.4.1. *Let $p \ge 1$ and $f \in L^p(\mathbb{R}^k)$. Then the following are equivalent:*

(i) $f \in W_1^p$,
(ii) *for any $a \in \mathbb{R}^k$, f has a weak derivative u_a, and $u_a \in L^p(\mathbb{R}^k)$.*

In this case, moreover, $d_a f = u_a$.

Proof. 1. Let us show that (i) implies (ii). For this purpose, assume $a \in \mathbb{R}^k$ and $\phi \in C_0^\infty(\mathbb{R}^k)$. Then for $\varepsilon > 0$ by translation 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) and 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 ϕ in $C_0^\infty(\mathbb{R}^k)$,

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

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

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

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

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

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

By means of this observation we can transfer the desired result from $\phi \in C_0^\infty(\mathbb{R}^k)$ to $\phi \in L^q(\mathbb{R}^k)$, since $C_0^\infty(\mathbb{R}^k)$ is dense in $L^q(\mathbb{R}^k)$, where q is the conjugate exponent of p. We then conclude by remarking that

$$||f||_p = \sup\{\langle f, g\rangle : g \in L^q(\mathbb{R}^k), ||g||_q \leq 1\}. \blacksquare$$

COROLLARY 1.4.1. *Let $e_1, \ldots, e_k$ denote the canonical basis of $\mathbb{R}^k$, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence in W_1^p such that*

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

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

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 $1 \leq i \leq k$, and $d_i f = \lim_{n\to\infty} d_i f_n \in L^p(\mathbb{R}^k)$. 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 $\phi \in C_0^\infty(\mathbb{R}^k)$ and $1 \leq i \leq k$,

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

This means that f has weak directional derivative in direction e_i and $d_i f = u_i \in L^p(\mathbb{R}^k)$. Now Theorem 1.4.1 is applicable and finishes the proof. $\blacksquare$

COROLLARY 1.4.2. *Let $p \geq 1$. Then W_1^p is a Banach space with respect to the norm $||\cdot||_{1,p}$, and for any $a \in \mathbb{R}^k$ the mapping $d_a : W_1^p \to L^p(\mathbb{R}^k)$ is continuous.*

Proof. We have to prove that W_1^p is complete with respect to $||\cdot||_{1,p}$. So let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in W_1^p. Then setting $f = \lim_{n\to\infty} f_n$ in $L^p(\mathbb{R}^k)$, we see that the hypotheses (i) and (ii) of Corollary 1.4.1 are satisfied, and it suffices to apply that corollary. $\blacksquare$

We finally need a local version of Sobolev spaces.

DEFINITION 1.4.4. For $p \geq 1$ and $s \in \mathbb{N}$ let

$$W_{s,\text{loc}}^p = \{f : \mathbb{R}^k \to \mathbb{R} \text{ measurable} : f\phi \in W_s^p \text{ for all } \phi \in C_0^\infty(\mathbb{R}^k)\} \tag{1.20}$$

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

THEOREM 1.4.2. *Let $p \geq 1$ and $s \in \mathbb{N}$. Then $f \in W_{s,\text{loc}}^p$ iff for any $x_0 \in \mathbb{R}^k$ there exists an open neighborhood V_{x_0} of x_0 such that for any $\phi \in C_0^\infty(\mathbb{R}^k)$ with support in V_{x_0} we have $\phi f \in W_s^p$.*

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

(i) $0 \le \phi_n \le 1$ for any $n \in \mathbb{N}$,
(ii) for any $n \in \mathbb{N}$ there exists $x_0(n)$ such that $\operatorname{supp}(\phi_n) \subset V_{x_0(n)}$,
(iii) $\sum_{n\in\mathbb{N}} \phi_n = 1$,
(iv) for any compact set $K \subset \mathbb{R}^k$ the intersection of K and $\operatorname{supp}(\phi_n)$ is non-empty for at most finitely many n.

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

$$\phi_k\phi f \in W_s^p, \quad k \in \mathbb{N}.$$

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

$$\phi f \in W_s^p. \ \blacksquare$$

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 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, first without reference to integrability. In fact, using Theorem 1.4.2, and for $s \in \mathbb{N}$, $p \ge 1$, $1 \le j_1, \dots, j_s \le k$, $f \in W^p_{s,\mathrm{loc}}$ 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 $x_0 \in \mathbb{R}^k$ by the corresponding generalized derivative of ϕf with $\phi \in C_0^\infty(\mathbb{R}^k)$ 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 1.4.5. Let $s \in \mathbb{N}$, $p \ge 1$, $1 \le j \le k$, $f \in W^p_{s,\mathrm{loc}}$, and denote by d_j the directional derivative in the direction of the jth unit vector in $\mathbb{R}^k$ according to the preceding remark. Let then

$$\nabla f = (d_1 f, \dots, d_k f), \tag{1.21}$$

$$\delta_j f = -d_j f + x_j f, \tag{1.22}$$

$$Lf = \sum_{j=1}^k \delta_j d_j f = \sum_{j=1}^k [-d_j d_j f + x_j d_j f]. \tag{1.23}$$

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

$$\nabla^r f = (d_{j_1} \cdots d_{j_r} f : 1 \leq j_1, \ldots, j_r \leq k). \tag{1.24}$$

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

DEFINITION 1.4.6. Let $p \geq 1$ and $s \in \mathbb{N}$. Then let

$$D_s^p(\mathbb{R}^k) = \Big\{ f \in W_{s,\mathrm{loc}}^p : \sum_{r=0}^{s} |||\nabla^r f|||_p < \infty \Big\}, \tag{1.25}$$

$$||f||_{s,p} = \sum_{r=0}^{s} |||\nabla^r f|||_p \tag{1.26}$$

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

REMARK. $D_s^p(\mathbb{R}^k)$ is a Banach space. This is seen by arguments as in the proof of Corollary 1.4.2.

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

$$f = \sum_{p \in E_k} \frac{c_p(f)}{p!} H_p \tag{1.27}$$

with coefficients $c_p(f) \in \mathbb{R}, p \in E_k$. By 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 $\mathscr{P}$ the linear hull of the k-dimensional Hermite polynomials. Plainly, $\mathscr{P} \subset W_{s,\mathrm{loc}}^p$ for any $s \in \mathbb{N}$ and $p \geq 1$. According to Chapter 3, $\mathscr{P}$ is dense in $L^2(\mathbb{R}^k, \nu_k)$. And for functions in $\mathscr{P}$, 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 any $p \in E_k$ and $1 \leq j \leq k$ 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 $f \sim (c_p(f)) \in \mathscr{P}$ and $1 \le j \le k$ we may write

$$d_j f = \sum_{p \in E_k} \frac{c_p(f)}{p!}\, p_j \prod_{i \neq j} H_{p_i} H_{p_j - 1}, \tag{1.28}$$

$$\delta_j f = \sum_{p \in E_k} \frac{c_p(f)}{p!} \prod_{i \neq j} H_{p_i} H_{p_j + 1}, \tag{1.29}$$

$$Lf = \sum_{p \in E_k} \frac{c_p(f)}{p!} |p| H_p. \tag{1.30}$$

According to Corollary 1.4.2 and the calculations just sketched, the natural domains of the operators extending ∇, δ_j and L beyond $\mathscr{P}$ must be those distributions in $\mathbb{R}^k$ 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 $D_1^2(\mathbb{R}^k)$. For $f \sim (c_p(f)) \in \mathscr{P}$ we have

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

If in addition $f \in L^2(\mathbb{R}^k, \nu_k)$, we may write $f \sim (c_p(f))$ and approximate it by $f_n = \sum_{p \in E_k, |p| \le n} c_p(f)/p! \in \mathscr{P}$, $n \in \mathbb{N}$. Hence, according to Corollary 1.4.2, f belongs to $D_1^2(\mathbb{R}^k)$ if the following series converges:

$$\begin{aligned}
|||\nabla f|||_2^2 &= \lim_{n \to \infty} |||\nabla f_n|||_2^2 \\
&= \lim_{n \to \infty} \sum_{p \in E_k, |p| \le n} |p| \frac{c_p(f)^2}{p!} \\
&= \sum_{p \in E_k} |p| \frac{c_p(f)^2}{p!} < \infty.
\end{aligned} \tag{1.31}$$

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 1.4.3. *Let* $r \in \mathbb{N}$ *and* $f \sim (c_p(f)) \in L^2(\mathbb{R}, \nu_1) \cap W^2_{r,\mathrm{loc}}$. *Define*

$$f_p = \frac{c_p(f)}{p!} H_p, \qquad p \geq 0.$$

Then the following are equivalent:

(i) $\nabla^r f \in L^2(\mathbb{R}, \nu_1)$,
(ii) $\sum_{p\geq 0} p^r \|f_p\|_2^2 < \infty$,
(iii) $f \in D^2_r(\mathbb{R})$,
(iv) $\delta^r f \in L^2(\mathbb{R}, \nu_1)$.

In particular, $D^2_r(\mathbb{R})$ *is the domain of* ∇^r *and* δ^r *in* $L^2(\mathbb{R}, \nu_1)$. *For* $f, g \in D^2_1(\mathbb{R})$ *we have*

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

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

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

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

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

whereas

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

This completes the proof. ■

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

THEOREM 1.4.4. *Let* $g \in D_1^4(\mathbb{R})$ *and* $\mu = \nu_1 \circ g^{-1}$. *If* $\phi \in L^4(\mathbb{R},\mu)$ *and* $\nabla\phi \in L^4(\mathbb{R},\mu)$, *then*

$$\nabla(\phi \circ g) = (\nabla \phi)\circ g \cdot \nabla g. \tag{1.32}$$

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

THEOREM 1.4.5. *Let* $f, g \in D_1^4(\mathbb{R})$. *Then* $f \cdot g \in D_1^2(\mathbb{R})$ *and*

$$\nabla(f\cdot g) = f \cdot \nabla g + \nabla f \cdot g. \tag{1.33}$$

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

Note that Theorem 1.4.5 provides a rigorous basis of justification for the calculations done in the context of the absolute continuity criterion in the one-dimensional setting. We now turn to arbitrary finite dimension k, and interpret Gaussian Sobolev spaces by convergence properties of Hermite expansions as above.

THEOREM 1.4.6. *Let* $f \sim (c_p(f)) \in L^2(\mathbb{R}^k, \nu_k) \cap W^2_{1,\mathrm{loc}}$. *Define* $f_p = (c_p(f)/p!)H_p$ *for* $p \in E_k$. *Then the following are equivalent:*

(i) $|\nabla f| = [\sum_{j=1}^k (d_j f)^2]^{1/2} \in L^2(\mathbb{R}^k, \nu_k)$,
(ii) $\sum_{p \in E_k} |p| \, \|f_p\|_2^2 < \infty$,
(iii) $f \in D_1^2(\mathbb{R}^k)$.

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

Proof. Analogous to the proof of Theorem 1.4.3. ∎

1.5. Infinite-dimensional Gaussian Sobolev spaces

To reduce the infinite-dimensional setting to a finite-dimensional one, we use the following observation.

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

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

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

LEMMA 1.5.1. *Let* $p \ge 1$ *and* $f \in L^p(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu)$. *Then*

$$\hat{f}_n = E(f \mid \mathbb{C}_n), \quad n \in \mathbb{N}, \tag{1.34}$$

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. ∎

In the following $f_n : \mathbb{R}^n \to \mathbb{R}$ is the n-dimensional factorization of $\hat{f}_n$, determikined by

$$f_n \circ \pi^n = \hat{f}_n, \quad n \in \mathbb{N}. \tag{1.35}$$

It is crucial for the definition of infinite-dimensional Sobolev spaces that the martingale property is essentially not destroyed by directional derivative operators.

LEMMA 1.5.2. *Let* $p > 1$, $f \in L^p(\mathbb{R}^{\mathbb{N}})$ *and* $(f_n)_{n \in \mathbb{N}}$ *the sequence defined above. Suppose that* $\sup_{n \in \mathbb{N}} \|f_n\|_{1,p} < \infty$. *Then for any* $j \in \mathbb{N}$ *the sequence* $(d_j f_n \circ \pi^n)_{n \in \mathbb{N}}$ *converges in* $L^p(\mathbb{R}^{\mathbb{N}})$, *to a limit that we denote by* $d_j f$. *Corresponding statements hold true for higher order derivatives.*

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

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

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

$$\sup_{n\geq j} \|d_j f_n \circ \pi^n\|_p \leq \sup_{n\in\mathbb{N}} \|f_n\|_{1,p} < \infty,$$

is bounded in $L^p(\mathbb{R}^{\mathbb{N}})$ and hence converges in $L^p(\mathbb{R}^{\mathbb{N}})$, as $p > 1$. ∎

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

DEFINITION 1.5.1. Let $p \geq 1$ and $s \in \mathbb{N}$. Then we define

$$D_s^p(\mathbb{R}^{\mathbb{N}}) = \{f \in L^p(\mathbb{R}^{\mathbb{N}}, \nu) : f_n \in D_s^p(\mathbb{R}^n),\, n \in \mathbb{N},\, \sup_{n\in\mathbb{N}} \|f_n\|_{s,p} < \infty\} \tag{1.36}$$

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

$$\|f\|_{s,p} = \sup_{n\in\mathbb{N}} \|f_n\|_{s,p}, \quad f \in D_s^p(\mathbb{R}^{\mathbb{N}}). \tag{1.37}$$

This definition makes sense, for the following reasons.

THEOREM 1.5.1. *Let* $p > 1$ *and* $s \in \mathbb{N}$. *Then* $D_s^p(\mathbb{R}^{\mathbb{N}})$ *is a Banach space with the norm* $\|\cdot\|_{s,p}$.

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

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

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

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

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

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

Moreover,

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

Hence by definition $f \in D_1^p(\mathbb{R}^{\mathbb{N}})$, 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 1.5.2, the gradient on infinite-dimensional Gaussian Sobolev spaces is defined as follows.

DEFINITION 1.5.2. Let $p > 1$ and $f \in D_1^p(\mathbb{R}^{\mathbb{N}})$. Then let

$$\nabla f = (d_j f)_{j\in\mathbb{N}} \tag{1.38}$$

(*Malliavin gradient* or *Malliavin derivative*), where for any $j \in \mathbb{N}$ according to Lemma 1.5.2,

$$d_j f = \lim_{n\to\infty} d_j f_n \circ \pi^n. \tag{1.39}$$

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

REMARK. The gradient ∇ being a continuous mapping from $D_1^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n, \nu_n)$ for any finite dimension n, Lemma 1.5.2 and the definition of the Malliavin gradient imply that ∇ is a continuous mapping from $D_1^p(\mathbb{R}^{\mathbb{N}})$ to $L^p(\mathbb{R}^{\mathbb{N}}, \nu)$.

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 $f = \sum_{p\in E}(c_p(f)/p!)H_p \in L^2(\mathbb{R}^{\mathbb{N}}, \nu)$. We shall continue to use the notation $f \sim (c_p(f))$. Then for $n \in \mathbb{N}$, we have $f_n = \sum_{p\in E_n}(c_{(p,0)}(f)/p!)H_p$, where we put $(p,0) = (p_1,\dots,p_n,0,0,\dots)$ for $p = (p_1,\dots,p_n) \in E_n$. Therefore, we also have $\hat{f}_n = \sum_{p\in E_n}(c_{(p,0)}(f)/p!)H_{(p,0)}$. Let again $\mathscr{P}$ be the linear hull of all generalized Hermite polynomials.

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

$$\begin{aligned} d_j f = \lim_{n\to\infty} d_j f_n \circ \pi^n &= \lim_{n\to\infty} \sum_{p\in E_n} \frac{c_{(p,0)}(f)}{p!}\, p_j \prod_{i\ne j} H_{p_i} H_{p_j-1} \\ &= \sum_{p\in E} \frac{c_p(f)}{p!}\, p_j \prod_{i\ne j} H_{p_i} H_{p_j-1}. \end{aligned} \tag{1.40}$$

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

$$\begin{aligned} \infty > \sup_{n\in\mathbb{N}} \||\nabla f_n \circ \pi^n|\|_2^2 &= \||\nabla f|\|_2^2 \\ &= \sup_{n\in\mathbb{N}} \sum_{p\in E_n} |p| \frac{c_{(p,0)}(f)^2}{p!} = \sum_{p\in E} |p| \frac{c_p(f)^2}{p!}. \end{aligned} \tag{1.41}$$

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

THEOREM 1.5.2. *For $f \in L^2(\mathbb{R}^{\mathbb{N}}, \nu)$ the following are equivalent:*

(i) $f \in D_1^2(\mathbb{R}^{\mathbb{N}})$,
(ii) $\sum_{p \in E} |p| c_p(f)^2 / p! < \infty$,
(iii) $|\nabla f_n| \circ \pi^n = [\sum_{j \in \mathbb{N}} (d_j f_n)^2 \circ \pi^n]^{1/2}$ *converges in* $L^2(\mathbb{R}^{\mathbb{N}}, \nu)$ *to* $|\nabla f|$.

Moreover, $D_1^2(\mathbb{R}^{\mathbb{N}})$ *is a Hilbert space with respect to the scalar product*

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

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

1.6. 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 concerns the smoothness of laws of random variables defined on the Gaussian space. We start with a generalization of Lemma 1.1.1 to finite measures on $\mathbf{B}^d$ for $d \in \mathbb{N}$.

LEMMA 1.6.1. *Let $\mu | \mathbf{B}^d$ be a finite measure. Assume there exists $c \in \mathbb{R}$ such that for all $\phi \in C^1(\mathbb{R}^d)$ with bounded partial derivatives, and any $1 \leq j \leq d$, we have*

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

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

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

1. Assume that $\phi \in C^1(\mathbb{R}^2)$ has compact support. We show that

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

In fact, we have

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

2. Let $u \geq 0$ be continuous with compact support and such that $\int u \, d\lambda = 1$. For $\varepsilon > 0$ define

$$u_\varepsilon = \frac{1}{\varepsilon^2} u\left(\frac{1}{\varepsilon}\cdot\right).$$

Moreover, let

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

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

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

3. We show that

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

is a continuous linear functional.

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

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

Generalizing this inequality to bounded measurable ϕ, and then taking

$\phi = \operatorname{sgn}(\psi_\varepsilon)$ yields the inequality

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

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

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

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

$$\left| \iint g(x)\, \mu(dx) \right| \leq c \|g\|_2.$$

Finally, extend this inequality to $g \in L^2(\mathbb{R}^2)$ 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 $f = (f^1, \ldots, f^d)$ with components in $L^2(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu)$. 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 1.6.1. Let $\phi \in C^1(\mathbb{R}^d)$ have 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

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

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

THEOREM 1.6.1. *Let $p \geq 2$, $f \in D_1^p(\mathbb{R}^{\mathbb{N}})^d$, and $\phi \in C^1(\mathbb{R}^d)$ with bounded partial derivatives. Then*

$$\phi \circ f \in D_1^p(\mathbb{R}^{\mathbb{N}}) \quad \textit{and} \quad \nabla[\phi \circ f] = \sum_{i=1}^d \frac{\partial}{\partial x_i} \phi(f) \cdot \nabla f^i. \tag{1.42}$$

Proof. Use Theorem 1.5.2 to choose a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathscr{P}^d$ such that for any $1 \le i \le d$,

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

For each $n \in \mathbb{N}$ 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 $D_1^p(\mathbb{R}^{\mathbb{N}})$, and since the partial derivatives of ϕ are bounded, we furthermore obtain

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

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

We next present a calculation leading to the verification of the absolute continuity criterion of Lemma 1.6.1. 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 $\phi \in C^1(\mathbb{R}^d)$ with bounded partial derivatives, the expression $\frac{\partial}{\partial x_i}\phi(f)$, $1 \le i \le d$. Recall the notation

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

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

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

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

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

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

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

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

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

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

$$c = \Big\|\sum_{1\le k\le d}\sum_{j=1}^{\infty}\delta_j(d_j f^k \sigma_{ki}^{-1})\Big\|_2$$

in $L^2(\mathbb{R}^{\mathbb{N}},\nu)$. It can be seen (in analogy to Theorem 1.4.3) that this series makes sense under the hypotheses of the following main theorem.

THEOREM 1.6.2. *Suppose that* $f = (f^1,\dots,f^d) \in L^2(\mathbb{R}^{\mathbb{N}},\nu)$ *satisfies*

(i) $f^i \in D_2^4(\mathbb{R}^{\mathbb{N}})$ *for* $1\le i\le d$,
(ii) $\sigma_{ik} = (\nabla f^i, \nabla f^k)$, $1\le i,k\le d$, *is* ν*-a.s. invertible and* $\sigma_{ki}^{-1}\in D_1^4(\mathbb{R}^{\mathbb{N}})$ *for* $1\le i,k\le d$.

Then $\nu\circ f^{-1} \ll \lambda^d$.

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

2. Calculus on the path space

2.1. The canonical Wiener space: multiple integrals

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

$$T : L^p(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu) \to L^p(\Omega, \mathscr{F}, P), \quad f \mapsto f \circ ((W(g_i))_{i\in\mathbb{N}}), \tag{2.1}$$

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

In the approach to differential calculus on Gaussian sequence spaces via the Hilbert space setting, the most important tool proved to be the Hermite expansions of functions in $L^2(\mathbb{R}^{\mathbb{N}}, \nu)$. In the setting on the canonical space, they can be given a different interpretation which we shall now develop. We follow Nualart [32].

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

$$\prod_{i=1}^{\infty} H_{p_i}(W(g_i)), \quad p \in E.$$

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

DEFINITION 2.1.1. For $m \in \mathbb{N}$ let

$$\mathscr{E}_m = \Big\{ f : \mathbb{R}^m_+ \to \mathbb{R} : f = \sum_{i_1,\dots,i_m=1}^{n} a_{i_1\dots i_m} 1_{A_{i_1}\times\cdots\times A_{i_m}},$$
$$(A_i)_{1\le i\le n} \subset \mathbf{B}^1_+ \text{ p.d.}, a_{i_1\dots i_m} = 0 \text{ in case } i_k = i_l \text{ for some } k \neq l \Big\}. \tag{2.2}$$

REMARK. For $f \in L^2(\mathbb{R}_+)$ of the form

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

let

$$W(f) = \sum_{i=1}^{n} a_i W(J_i) = \sum_{i=1}^{n} a_i (W_{t_i} - W_{s_i}), \quad \text{if } J_i =]s_i, t_i],\ 1 \le i \le n.$$

Then by Itô's isometry we have

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

Since $\mathscr{E}_1$ is dense in $L^2(\mathbb{R}_+)$, we can extend the linear mapping $f \mapsto W(f)$ to $L^2(\mathbb{R}_+)$. 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 $\mathbb{R}_+^m$ with unspecified integer m.

DEFINITION 2.1.2. For $f = \sum_{i_1,\dots,i_m=1}^{n} a_{i_1 \dots i_m} 1_{A_{i_1} \times \dots \times A_{i_m}} \in \mathscr{E}_m$ let

$$I_m(f) = \sum_{i_1,\dots,i_m=1}^{n} a_{i_1 \dots i_m} W(A_{i_1}) \cdots W(A_{i_m}). \tag{2.3}$$

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

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

LEMMA 2.1.1. *Let $m, q \in \mathbb{N}$ and $f \in \mathscr{E}_m$, $g \in \mathscr{E}_q$.*

(i) $I_m : \mathscr{E}_m \to \mathbb{R}$ *is linear.*
(ii) *If* $\tilde{f}(t_1, \dots, t_m) = \frac{1}{m!} \sum_{\sigma \in \mathscr{S}_m} f(t_{\sigma(1)}, \dots, t_{\sigma(m)})$ (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. (i) follows from the additivity of the map $A \mapsto W(A)$.

(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.

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

$$f = \sum_{i_1,\ldots,i_m=1}^{n} a_{i_1\ldots i_m} 1_{A_{i_1}\times\cdots\times A_{i_m}},$$

$$g = \sum_{i_1,\ldots,i_q=1}^{n} b_{i_1\ldots i_q} 1_{A_{i_1}\times\cdots\times A_{i_q}}.$$

Now if $m \neq q$, by the assumptions that $(A_i)_{1\le i\le n}$ 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,

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

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

LEMMA 2.1.2. *$\mathscr{E}_m$ is dense in $L^2(\mathbb{R}_+^m)$ for any $m \in \mathbb{N}$.*

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

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

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

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

where $b_{i_1\ldots i_m} = 1$ if $B_{i_1} \times \cdots \times B_{i_m} \subset A_1 \times \cdots \times A_m$, and 0 otherwise. Let $I = \{(i_1,\ldots,i_m) : i_k \neq i_l \text{ for } k \neq l\}$, and $J = \{1,\ldots,n\}^m \setminus I$. Then by definition

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

and we have

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

Finally, we have to choose δ small enough. ■

Using Lemma 2.1.2, we may now extend I_m to $L^2(\mathbb{R}_+^m)$.

DEFINITION 2.1.3. The linear and continuous extension of I_m to $L^2(\mathbb{R}_+^m)$ which exists according to Lemma 2.1.2 is called the *multiple Wiener–Itô integral* of degree m and is also denoted by I_m.

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

THEOREM 2.1.1. *Let $m,q\in\mathbb{N}$ and $f\in L^2(\mathbb{R}_+^m)$, $g\in L^2(\mathbb{R}_+^q)$.*

(i) *$I_m : L^2(\mathbb{R}_+^m)\to\mathbb{R}$ 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) *$I_1(f)=W(f)$ for $f\in L^2(\mathbb{R}_+)$.*

NOTATION. We write

$$\begin{aligned}I_m(f) &= \int_{\mathbb{R}_+^m} f(t_1,\dots,t_m)\,dW_{t_1}\cdots dW_{t_m} \qquad (2.4)\\ &= \int_{\mathbb{R}_+^m} f(t_1,\dots,t_m)\,W(dt_1)\cdots W(dt_m).\end{aligned}$$

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:

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

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

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

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

DEFINITION 2.1.4. Let $m \in \mathbb{N}$, $f \in L^2(\mathbb{R}_+^m)$, $g \in L^2(\mathbb{R}_+)$. Then for $t_1, \dots, t_m, t \in \mathbb{R}_+$,

$$f \otimes g(t_1, \dots, t_m, t) = f(t_1, \dots, t_m) \cdot g(t) \qquad \text{(tensor product)}, \tag{2.5}$$

$$f \otimes_1 g(t_1, \dots, t_{m-1}) = \int_{\mathbb{R}_+} \tilde{f}(t_1, \dots, t_m) g(t_m)\, dt_m \qquad \text{(contraction)}. \tag{2.6}$$

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

LEMMA 2.1.3. *Let $m \in \mathbb{N}$ and $f \in L^2(\mathbb{R}_+^m)$, $g \in L^2(\mathbb{R}_+)$. Then*

$$I_m(f) I_1(g) = I_{m+1}(f \otimes g) + m I_{m-1}(f \otimes_1 g). \tag{2.7}$$

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

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

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 $B_1, \dots, B_n \in \mathbf{B}_+^1$ such that $A_1 = \bigcup_{i=1}^n B_i$, and $\lambda(B_i) < \varepsilon$ for any $1 \le i \le n$. Then

$$\begin{aligned} I_m(f) I_1(g) &= W(A_1)^2 W(A_2) \cdots W(A_m) \\ &= \sum_{i \neq j} W(B_i) W(B_j) W(A_2) \cdots W(A_m) \\ &\quad + \sum_{1 \le i \le n} [W(B_i)^2 - \lambda(B_i)] W(A_2) \cdots W(A_m) \\ &\quad + \lambda(A_1) W(A_2) \cdots W(A_m). \end{aligned} \tag{2.8}$$

(a) We now prove that the first term on the right hand side of (2.8) is close to $I_{m+1}(f \otimes g)$. In fact, let

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

Then

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

(b) Let us next prove that the second term on the right hand side of (2.8) is negligible in the limit $\varepsilon \to 0$. In fact, denote it by R_ε. Then, since for $1 \le i \le n$ the variance of $W^2(B_i) - \lambda(B_i)$ is given by $c\lambda(B_i)^2$ with some constant c, we obtain

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

(c) To evaluate the last term, note that

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

Therefore

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

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 2.1.2. *Let $m \in \mathbb{N}$ and $h \in L^2(\mathbb{R}_+)$ be such that $||h||_2 = 1$. Denote by $h^{\otimes_m}$ the m-fold tensor product of h with itself. Then*

$$H_m(W(h)) = I_m(h^{\otimes_m}). \tag{2.9}$$

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

$$L^2(\Omega, \mathscr{F}, P) = \bigoplus_{m=0}^{\infty} \mathscr{H}_m. \tag{2.10}$$

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

$$F = \sum_{m=0}^{\infty} I_m(f_m). \tag{2.11}$$

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

Proof. 1. We first have to prove that

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

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 holds for m. Then Lemma 2.1.3 and the recursion formula for Hermite polynomials given above combine to yield, remembering that $||h||_2 = 1$,

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

2. Let $L^2_s(\mathbb{R}^m_+)$ be the linear space of symmetric functions in $L^2(\mathbb{R}^m_+)$. Then by Theorem 2.1.1,

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

hence $\mathscr{H}_m = I_m(L^2_s(\mathbb{R}^m_+))$ is closed. Orthogonality is also a consequence of Theorem 2.1.1.

3. Let $(g_i)_{i\in\mathbb{N}}$ be an orthonormal basis of $L^2(\mathbb{R}_+)$, and let $F \in L^2(\Omega, \mathscr{F}, P)$. Let $f \in L^2(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu)$ be such that $T(f) = F$. Assume $f \sim (c_p(f))$, according to the notation of Chapter 1. For $m \geq 0$, define

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

considered as a function of m variables. Then $f_m \in L^2(\mathbb{R}^m_+)$ and with the help of a slight generalization of Lemma 2.1.3 we see that

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

Summing this expression over m yields the desired

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

The remaining claims are obvious. ∎

2.2. 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 $(g_i)_{i\in\mathbb{N}}$ of $L^2(\mathbb{R}_+)$, and recall the isomorphism

$$T : L^2(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu) \to L^2(\Omega, \mathscr{F}, P), \quad f \mapsto f((W(g_i))_{i\in\mathbb{N}}). \tag{2.12}$$

Of course, every permutation of the basis functions gives another orthonormal basis. So here we encounter the 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 1.5.2,

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

takes values in l^2. The corresponding object on the side of the canonical space is $L^2(\mathbb{R}_+)$. It is therefore reasonable to make

DEFINITION 2.2.1. For $n \in \mathbb{N}$ let $C_p^\infty(\mathbb{R}^n)$ denote the set of smooth functions with all partial derivatives of polynomial growth. Let

$$\mathscr{S} = \{F : F = f(W(h_1),\dots,W(h_n)),\ h_1,\dots,h_n \in L^2(\mathbb{R}_+),\\ f \in C_p^\infty(\mathbb{R}^n),\ n \in \mathbb{N}\}. \tag{2.13}$$

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

$$D_t F = \sum_{i=1}^n \frac{\partial}{\partial x_i} f(W(h_1),\dots,W(h_n)) h_i(t). \tag{2.14}$$

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

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

For $1 \leq i \leq n$ write

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

Then, setting

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

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

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

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

If $h \in L^2(\mathbb{R}_+)$ is another function and $F = f(W(g_1), \dots, W(g_m))$, we have by definition

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

in particular for $i \in \mathbb{N}$,

$$\langle D_\bullet F, g_i\rangle = \frac{\partial}{\partial x_i} f(W(g_1), \dots, W(g_m)) = d_i f(W(g_1), \dots, W(g_m)). \tag{2.15}$$

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

$$\begin{aligned}\langle DF, DF\rangle = \sum_{i=1}^{m}\langle DF, g_i\rangle^2 &= \sum_{i=1}^{m}(d_i f)^2(W(g_1), \dots, W(g_m))\\ &= |\nabla f|^2(W(g_1), \dots, W(g_m)).\end{aligned} \tag{2.16}$$

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 simply transfer the definitions of Gaussian Sobolev spaces to the setting of the canonical Wiener space.

DEFINITION 2.2.2. Let $p \geq 2$ and $s \in \mathbb{N}$. For $F \in L^2(\Omega, \mathscr{F}, P)$ denote by $f \in L^2(\mathbb{R}^{\mathbb{N}}, \mathbf{B}^{\mathbb{N}}, \nu)$ the function for which we have $F = f((W(g_i))_{i\in\mathbb{N}})$. Then let

$$D_s^p = \{F : f \in D_s^p(\mathbb{R}^{\mathbb{N}})\} \tag{2.17}$$

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

$$||F||_{s,p} = ||f||_{s,p}. \tag{2.18}$$

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

$$D^r F = \sum_{j_1,\dots,j_r=1}^{\infty} d_{j_1}\cdots d_{j_r} f(W(g_i)_{i\in\mathbb{N}}) g_{j_1} \otimes \cdots \otimes g_{j_r} \tag{2.19}$$

(the *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 $p \geq 2$ and $s \in \mathbb{N}$, and that for $F \in D_s^p$ we have

$$||F||_{s,p} = \sum_{r=0}^{s} ||\langle D^r F, D^r F\rangle^{1/2}||_p.$$

We know that D_s^p is the closure of $\mathscr{S}$ 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

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

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 $L^2(\Omega,\mathscr{F},P)$.

Let us now investigate how D acts on the decomposition into Wiener–Itô integrals.

THEOREM 2.2.1. *Let $F=\sum_{m=0}^{\infty} I_m(f_m)\in L^2(\Omega,\mathscr{F},P)$ be given, with f_m symmetric for any $m\geq 0$. Then*

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

In this case

$$D_tF=\sum_{m=1}^{\infty} mI_{m-1}(f_m(\cdot,t)) \tag{2.20}$$

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

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

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

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

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

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

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

$$D_tI_m(f_m)=mI_{m-1}(f_m(\cdot,t))$$

follows.

3. For $n \in \mathbb{N}$ 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

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

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 $n, m \in \mathbb{N}$ with $n \geq m$ be given. Then

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

Hence $(DF_n)_{n\in\mathbb{N}}$ 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 2.2.2. *Let $p \geq 2$ and $d \in \mathbb{N}$, $\phi \in C^1(\mathbb{R}^d)$ with bounded partial derivatives, and let $F = (F^1, \ldots, F^d) \in (D_1^p)^d$. Then $\phi \circ F \in D_1^p$ and*

$$D\phi \circ F = \sum_{i=1}^{d} \frac{\partial}{\partial x_i} \phi(F) DF^i. \tag{2.21}$$

Proof. The proof of Theorem 1.6.1 translates. ∎

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

THEOREM 2.2.3. *Let $F \in \mathcal{S}$ and $h \in L^2(\mathbb{R}_+)$. Then*

$$E(\langle DF, h\rangle) = E(F\, W(h)). \tag{2.22}$$

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

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

This completes the proof. ∎

THEOREM 2.2.4. *Let $F, G \in \mathscr{S}$ and $h \in L^2(\mathbb{R}_+)$. Then*

$$E(G\langle DF, h\rangle) = E(FGW(h) - F\langle DG, h\rangle). \tag{2.23}$$

Proof. Apply Theorem 2.2.3 to the function FG. ∎

2.3. The canonical Wiener space: Skorokhod's integral

In this chapter we focus on the dual operator (in the sense of Hilbert space theory) of the Malliavin gradient rather than on 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

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

is densely defined and linear.

DEFINITION 2.3.1. Let

$$\operatorname{dom}(\delta) = \{u \in L^2(\Omega \times \mathbb{R}_+) : \text{there is } c \in \mathbb{R} \text{ such that for any } F \in D_1^2 \text{ we have } E(\langle DF, u\rangle) \le c\|F\|_2\}. \tag{2.24}$$

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

$$E(\langle DF, u\rangle) = E(F \cdot \delta(u)), \quad F \in D_1^2. \tag{2.25}$$

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

DEFINITION 2.3.2. For $u \in \operatorname{dom}(\delta)$ the uniquely determined random variable $\delta(u) \in L^2(\Omega)$ is called the *Skorokhod integral* of u.

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

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

DEFINITION 2.3.3. Let

$$\mathscr{S}_{L^2(\mathbb{R}_+)} = \left\{u : u = \sum_{i=1}^n F_i h_i,\ F_i \in \mathscr{S},\ h_i \in L^2(\mathbb{R}_+),\ n \in \mathbb{N}\right\}. \tag{2.26}$$

LEMMA 2.3.1. *Let $u = \sum_{i=1}^n F_i h_i \in \mathscr{S}_{L^2(\mathbb{R}_+)}$. Then*

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

Proof. By linearity of δ we may assume that $u = Fh$ with $F \in \mathscr{S}$ and $h \in L^2(\mathbb{R}_+)$. Then for $G \in \mathscr{S}$, by Theorem 2.2.4,

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

Hence we have

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

This completes the proof. ■

Recall now the standard Wiener filtration $(\mathscr{F}_t)_{t\geq 0}$, which for $t \geq 0$ is given by the P-completion $\mathscr{F}_t$ of $\sigma(W_s : s \leq t)$. Lemma 2.3.1 yields the elementary Itô integral if F_i is $\mathscr{F}_{t_i}$-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 2.3.2. *Let $F \in D_1^2$, $A \in \mathbf{B}_+^1$ and $\mathscr{F}_A = \sigma(W(1_B) : B \subset A, \lambda(B) < \infty)$. Then*

$$E(F \mid \mathscr{F}_A) \in D_1^2$$

and

$$D_t E(F \mid \mathscr{F}_A) = E(D_t F \mid \mathscr{F}_A) 1_A(t) \tag{2.27}$$

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

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

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

Let

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

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

$$E(F \mid \mathscr{F}_A) = \int g(W(h_1 1_A), \ldots, W(h_n 1_A), y_1, \ldots, y_n)\, dQ(y_1, \ldots, y_n).$$

Hence $E(F \mid \mathscr{F}_A) \in \mathscr{S}$ and

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

2. It remains to approximate $F \in D_1^2$ using standard arguments. ■

THEOREM 2.3.1. *Let $u \in L^2(\Omega \times \mathbb{R}_+)$ be $(\mathscr{F}_t)$-adapted. Then*

$$u \in \operatorname{dom}(\delta) \quad \textit{and} \quad \delta(u) = \int_{\mathbb{R}_+} u_t\, dW_t \quad \textit{(Itô integral).} \tag{2.28}$$

Proof. 1. Let $0 \le s < t$ and $F \in L^2(\Omega, \mathscr{F}_s, P)$. We prove

$$u = F 1_{]s,t]} \in \mathrm{dom}(\delta) \quad \text{and} \quad \delta(u) = F(W_t - W_s).$$

(a) Let first $F \in \mathscr{S}$. Then by Lemmas 2.3.1 and 2.3.2 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 $F \in L^2(\Omega, \mathscr{F}_s, P)$ let $(F^n)_{n\in\mathbb{N}} \subset \mathscr{S}$ be such that $F^n \to F$ in $L^2(\Omega, \mathscr{F}, P)$. Then also $\mathscr{S} \ni G^n = E(F^n \mid \mathscr{F}_s) \to F$ in $L^2(\Omega, \mathscr{F}, P)$. Hence by (a) for any $n \in \mathbb{N}$,

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

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

$$F 1_{]s,t]} \in \mathrm{dom}(\delta) \quad \text{and} \quad \delta(F 1_{]s,t]}) = F(W_t - W_s).$$

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

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

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

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

This completes the proof. ■

We next ask how the Skorokhod integral acts on the decomposition into multiple Wiener–Itô integrals.

Lemma 2.3.3. *Let $u \in L^2(\Omega \times \mathbb{R}_+)$. Then for any $m \ge 0$ there exist functions $f_m \in L^2(\mathbb{R}_+^{m+1})$ such that f_m is symmetric in its first m variables, and*

$$u_t = \sum_{m=0}^\infty I_m(f_m(\cdot, t)) \quad \text{in } L^2(\Omega \times \mathbb{R}_+). \tag{2.29}$$

Moreover,

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

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

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

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

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

For $n\in\mathbb{N}$ and $1\le k\le m_n$ let

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

Then

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

Define

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

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

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

Hence for any $m\ge 0$, $(f_m^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathbb{R}_+^{m+1})$ which converges to a function f_m which is also symmetric in the first m variables. For $u^{n,M} = \sum_{m=0}^{M} I_m(f_m^n)$, $n,M\in\mathbb{N}$, we obtain

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

By a similar argument and by definition we must have

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

This completes the proof. ■

THEOREM 2.3.2. *Let $u\in L^2(\Omega\times\mathbb{R}_+)$, $u=\sum_{m=0}^{\infty} I_m(f_m(\cdot,\cdot))$ according to Lemma 2.3.3. Then*

$$u\in \mathrm{dom}(\delta) \quad \textit{if and only if} \quad \sum_{m=0}^{\infty} (m+1)!\,||\tilde{f}_m||_2^2 < \infty.$$

In this case

$$\delta(u) = \sum_{m=0}^{\infty} I_{m+1}(f_m). \tag{2.30}$$

Proof. 1. Let $n \in \mathbb{N}$, $g \in L^2(\mathbb{R}_+^n)$ and $G = I_n(g)$. We show that

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

In fact, by Theorem 2.2.1,

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

2. Let us now prove the "if" part of the claim. For this purpose, let $u \in \mathrm{dom}(\delta)$ and $G = I_n(g) \in \mathscr{H}_n$, $n \in \mathbb{N}$. 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

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

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 $g_n \in L^2(\mathbb{R}_+^n)$, $n \in \mathbb{N}$, symmetric, and finally let $G^n = \sum_{k=0}^{n} I_k(g_k)$, $n \in \mathbb{N}$. Then we may again appeal to the first part of the proof to get

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

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

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

DEFINITION 2.3.4. Let

$$\begin{aligned} \mathbf{L}_1^2 = \Big\{ & u \in L^2(\Omega \times \mathbb{R}_+) : u_t \in D_1^2 \text{ for } \lambda\text{-a.a. } t \geq 0, \text{ and} \\ & \text{for some measurable version of } (s,t) \mapsto D_s u_t \text{ we have} \\ & E\left(\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} |D_s u_t|^2\,ds\,dt\right) < \infty \Big\}. \end{aligned} \tag{2.31}$$

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 $u \in L^2(\Omega \times \mathbb{R}_+)$, $u = \sum_{m=0}^{\infty} I_m(f_m(\cdot,\cdot))$, according to Lemma 2.3.3. 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 2.2.1, $u_t \in D_1^2$ means that

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

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

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

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 2.3.2. Since for $m \geq 0$ we have $||\tilde{f}_m||_2 \leq ||f_m||_2$, we obviously have

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

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

THEOREM 2.3.3. *Let $u, v \in \mathbf{L}_1^2$. Then*

$$E(\delta(u)\delta(v)) = E\left(\int_{\mathbb{R}_+} u_t v_t \, dt\right) + E\left(\int_{\mathbb{R}_+^2} D_t u_s D_s v_t \, ds \, dt\right). \tag{2.32}$$

Proof. 1. Let first $u \in \mathscr{S}_{L^2(\mathbb{R}_+)}$. We show that

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

By linearity, we may assume that $u = Fh$, where $F \in \mathscr{S}$ and $h \in L^2(\mathbb{R}_+)$. Then according to Lemma 2.3.1 we may write $\delta(u) = F\,W(h) - \langle DF, h\rangle$, and therefore

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

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

$$\begin{aligned} E(\delta(u)\delta(v)) &= E(\langle D\delta(u), v\rangle) \\ &= E(\langle u, v\rangle + \langle \delta(D_\bullet u), v\rangle) \\ &= E(\langle u, v\rangle) + E\left(\int_{\mathbb{R}_+^2} D_t u_s D_s v_t \, ds\, dt\right). \end{aligned}$$

3. It remains to approximate u by functions in $\mathscr{S}_{L^2(\mathbb{R}_+)}$, and to use the fact that convergence in $\mathbf{L}_1^2$ 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 applications of Malliavin's calculus to stochastic analysis.

THEOREM 2.3.4. *Let $u \in L^2(\Omega \times [0,1])$ be adapted, and $X_t = \int_0^t u_s \, dW_s$, $0 \leq t \leq 1$, its Itô integral process. Then*

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

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

$$D_t X_T = u_t 1_{[0,T]}(t) + \int_t^T D_t u_r \, dW_r, \tag{2.33}$$

and

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

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 2.3.3 to adapted u. Let $u = \sum_{m=0}^\infty I_m(f_m(\cdot,\cdot))$ be the representation according to Lemma 2.3.3. Here for any $m \geq 0$ the function $f_m \in L^2(\mathbb{R}_+^{m+1})$ is symmetric in its first m variables. We show that

$$f_m(\cdot, t) = f_m(\cdot, t) 1_{[0,t]^m}(\cdot) \quad (\text{in } L^2(\mathbb{R}_+^{m+1})), \tag{2.34}$$

and thus

$$||\tilde{f}_m||_2^2 = \frac{1}{m+1} ||f_m||_2^2. \tag{2.35}$$

In fact, resume the notation of the proof of Lemma 2.3.3, specialized to the adapted case. We approximate u in $L^2(\Omega \times [0,1])$ by functions

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

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

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

of F_k^n, by its measurability properties and Lemma 2.3.2, for $t_1, \dots, t_m \in \mathbb{R}_+$ we have

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

Hence the functions

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

satisfy

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

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

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

To prove the second assertion, note that by 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 supports, and their norms in $L^2(\mathbb{R}_+^{m+1})$ are identical to those 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 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 2.2.1, $X_1 \in D_1^2$ 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 by Lemma 2.3.2,

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

Moreover, for $t \leq s \leq 1$ we have, according to Theorem 2.2.1,

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

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 $(D_t u_s)_{t \leq s \leq 1}$ is adapted, we know from Theorem 2.3.1 that $(D_t u_s)_{t \leq s \leq 1}$ is Itô integrable and that its Itô and Skorokhod integrals are identical. More formally,

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

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 $L^2(\Omega \times [0,1])$)

$$\begin{aligned} D_t \delta(u) &= D_t \sum_{m=0}^{\infty} I_{m+1}(f_m) \\ &= D_t \sum_{m=0}^{\infty} I_{m+1}(\tilde{f}_m) \end{aligned}$$

$$= \sum_{m=1}^{\infty}(m+1)I_m(\tilde{f}_m(\cdot,t))$$
$$= \sum_{m=1}^{\infty} I_m(f_m(\cdot,t)) + \sum_{m=1}^{\infty} m I_m(f_m(\ldots,t,\cdot)).$$

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. ■

3. Backward stochastic differential equations

3.1. Martingale optimality and 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. Let us validate this claim in this introductory section by looking at a typical utility optimization problem arising in financial stochastics in the context of incomplete markets. In fact, we shall derive an important class of BSDE by applying martingale optimality to solve an optimal investment problem of a small financial agent whose income is partly affected by market external risk. We follow Hu et al. [15].

Let $T > 0$ be a finite time horizon which determines an interval during which financial investment decisions are taken. For the probabilistic setting, we fix a canonical probability space $(\Omega, \mathscr{F}, P)$ carrying an n-dimensional Brownian motion $(W_t)_{t\in[0,T]}$. The filtration $(\mathscr{F}_t)_{t\geq 0}$ is the natural filtration of W completed by sets of P-measure 0. We imagine a small investor active on the market who receives his income from two sources. As a *first source*, he is able to invest in the financial market. The market is supposed to be modelled by one bond with interest rate zero and $d < n$ stocks. In this case we face an incomplete market. The price process of stock i evolves according to the equation

$$\frac{dS_t^i}{S_t^i} = b_t^i\, dt + \sigma_t^i\, dW_t, \quad i = 1, \dots, d, \tag{3.1}$$

where b^i (resp. σ^i) is an $\mathbb{R}$-valued (resp. $\mathbb{R}^{1\times n}$-valued) predictable uniformly bounded stochastic process. The rows of the $d \times m$-matrix σ are the vectors σ_t^i, $i = 1, \dots, d$. The volatility matrix $\sigma = (\sigma^i)_{i=1,\dots,d}$ has full rank and we assume that $\sigma\sigma^*$ is uniformly elliptic, i.e. $KI_d \geq \sigma\sigma^* \geq \varepsilon I_d$, P-a.s. with constants $K > \varepsilon > 0$. The predictable $\mathbb{R}^n$-valued process, called the *market price of risk* and defined by

$$\theta_t = \sigma_t^*(\sigma_t\sigma_t^*)^{-1} b_t, \quad t \in [0, T],$$

is then also uniformly bounded.

A d-dimensional $(\mathcal{F}_t)_{t\in[0,T]}$-predictable process $\pi = (\pi_t)_{0\leq t\leq T}$ is called a *trading strategy* if $\int \pi \frac{dS}{S}$ is well defined, e.g. $\int_0^T \|\pi_t\sigma_t\|^2\,dt < \infty$ P-a.s. For $1 \leq i \leq d$, the process π_t^i describes the amount of money invested in stock i at time t. The number of shares is π_t^i/S_t^i. The wealth process X^π based on the trading strategy π with initial capital x satisfies the stochastic differential equation

$$X_t^\pi = x + \sum_{i=1}^d \int_0^t \frac{\pi_{i,u}}{S_{i,u}}\,dS_{i,u} = x + \int_0^t \pi_u\sigma_u(dW_u + \theta_u\,du), \qquad t\in[0,T]. \quad (3.2)$$

In this notation π has to be taken as a vector in $\mathbb{R}^{1\times d}$. We allow further constraints on the trading strategies the agent chooses. Our approach to the trader's utility optimization problem by means of martingale optimality instead of duality theory of convex analysis will allow us to work with arbitrary closed, not necessarily convex constraint sets. Hence we shall require that the trading strategies take their values in a *closed set*, i.e. $\pi_t(\omega)\in\tilde{C}$, with $\tilde{C}\subseteq\mathbb{R}^{1\times d}$ closed. For technical reasons we impose some further integrability assumptions which at the same time guarantee that there is no arbitrage.

In our concept of admissible trading strategies we use martingales of *bounded mean oscillation*, briefly BMO-martingales. We recall a few well known facts from this theory following [20]. Let $(\mathbf{G}_t)_{0\leq t\leq T}$ be a complete, right-continuous filtration, P a probability measure and M a continuous local $(P,(\mathbf{G}_t))$-martingale satisfying $M_0 = 0$. Let $1\leq p<\infty$. Then M belongs to the normed linear space BMO_p if

$$\|M\|_{\mathrm{BMO}_p} := \sup_{\tau\ (\mathbf{G}_t)\text{-stopping time}} E[|M_T - M_\tau|^p\,|\,\mathbf{G}_\tau]^{1/p} < \infty.$$

By Corollary 2.1 in [20], M is a BMO_p-martingale if and only if it is a BMO_q-martingale for every $q\geq 1$. Therefore it is simply called a BMO-martingale. In particular, M is a BMO-martingale if and only if

$$\|M\|_{\mathrm{BMO}_2} = \sup_{\tau\ (\mathbf{G}_t)\text{-stopping time}} E[\langle M\rangle_T - \langle M\rangle_\tau\,|\,\mathbf{G}_\tau]^{1/2} < \infty.$$

This means that local martingales of the form $M_t = \int_0^t \xi_s\,dW_s$ are BMO-martingales if and only if

$$\|M\|_{\mathrm{BMO}_2} = \sup_{\tau\ (\mathbf{G}_t)\text{-stopping time}} E\left[\int_\tau^T \|\xi_s\|^2\,ds \,\middle|\, \mathbf{G}_\tau\right]^{1/2} < \infty. \quad (3.3)$$

By finiteness of time horizon, this condition is satisfied for bounded integrands. According to Theorem 2.3 of [20], the stochastic exponential $\exp(M - \frac{1}{2}\langle M\rangle)$ of a BMO-martingale M is a uniformly integrable martingale. If Q is a probability measure defined by $dQ = \exp(M - \frac{1}{2}\langle M\rangle)_T dP$ for a P-BMO martingale M, then a P-BMO martingale is a BMO-martingale under Q (Theorem 3.6 in [20]).

DEFINITION 3.1.1. Let $\tilde{C}$ be a closed set in $\mathbb{R}^{1\times d}$. The set $\tilde{\mathscr{A}}$ of *admissible trading strategies* consists of all d-dimensional predictable processes $\pi = (\pi_t)_{0\le t\le T}$ which satisfy $\pi_t \in \tilde{C}$ $\lambda \otimes P$-a.s., $\int_0^\cdot \pi_s \sigma_s \, dW_s$ is a P-BMO-martingale, and $E(U(X_T^\pi)) > -\infty$.

The boundedness of θ following from our hypotheses, and Theorem 3.6 in [20], imply that the wealth process X^π is a BMO-martingale under the equivalent probability measure (*risk neutral measure*)

$$Q^0 = \exp\left(-\int_0^T \theta_s \, dW_s - \frac{1}{2}\int_0^T |\theta_s|^2 \, ds\right) \cdot P. \tag{3.4}$$

Therefore the set $\tilde{\mathscr{A}}$ is free of arbitrage, i.e. in this set there is no trading strategy π with initial capital $X_0^\pi = 0$, terminal wealth $X_T^\pi \ge 0$ P-a.s. and $P[X_T^\pi > 0] > 0$. For technical reasons, we define for $t \in [0,T]$ and $\omega \in \Omega$ the effective constraint set by

$$C_t(\omega) = \tilde{C}\sigma_t(\omega). \tag{3.5}$$

If we write

$$p_t = \pi_t \sigma_t, \quad X_t^p = x + \int_0^t p_u (dW_u + \theta_u \, du), \quad t \in [0,T], \tag{3.6}$$

the set of admissible trading strategies $\tilde{\mathscr{A}}$ for π is transformed into a constraint set $\mathscr{A}$ for the $\mathbb{R}^{1\times n}$-valued predictable stochastic process p which will also be called a *strategy*, characterized by $p_t(\omega) \in C_t(\omega)$ for $P\otimes\lambda$-a.e. $(\omega, t) \in \Omega \times [0,T]$, and $\int_0^\cdot p_s \, dW_s$ is a Q^0-BMO-martingale, hence a P-BMO-martingale (see Kazamaki [20, Theorem 3.6]).

Our investor's *second source* of income is given by some bounded random liability F which is assumed to be $\mathscr{F}_T$-measurable, but not necessarily positive. The uncertainty inherent in this random *loss* may be thought of as being independent of or only correlated to the financial market: remember that $d < n$, i.e. there are components of the Wiener process of our model independent of the stocks, which may cause part of the uncertainty in F. One may think of a weather or climate dependent liability, such as the random loss of a re-insurance company due to hurricane risk.

So the investor's total income for an investment strategy p at terminal time T of the trading interval is then given by $X_T^p - F$. Now suppose he draws utility from this wealth according to an exponential utility function

$$U(x) = -\exp(-\alpha x), \quad x \in \mathbb{R}, \tag{3.7}$$

where $\alpha > 0$ describes his risk aversion. In these terms the investor finally wants to find the maximum in the utility optimization problem

$$\begin{aligned} V(x) &= \sup_{p\in\mathscr{A}} E[-\exp(-\alpha(X_t^p - F))] \\ &= \sup_{p\in\mathscr{A}} E\left[-\exp\left(-\alpha\left(x+\int_0^T p_t(dW_t+\theta_t\,dt)-F\right)\right)\right], \end{aligned} \tag{3.8}$$

where x is the initial wealth. Here V is called the *value function*.

In order to find the value function and an optimal strategy we use the *martingale optimality principle.* This is done by constructing a family $(R^{(p)})_{p\in\mathscr{A}}$ of stochastic processes with the following properties:

(i) $R_T^{(p)} = -\exp(-\alpha(X_T^p - F))$ for all $p \in \mathscr{A}$,
(ii) $R_0^{(p)} = R_0$ is a constant independent of $p \in \mathscr{A}$,
(iii) $R^{(p)}$ is a supermartingale for all $p \in \mathscr{A}$,
(iv) there exists a $p^* \in \mathscr{A}$ such that $R^{(p*)}$ is a martingale.

The processes $R^{(p)}$ and the initial value R_0 depend of course on the initial capital x.

Let us first show that once such a family of processes is constructed, our utility optimization problem is solved. In fact, we can compare the expected utilities of general strategies $p \in \mathscr{A}$ corresponding to supermartingales $R^{(p)}$ and $p^* \in \mathscr{A}$ corresponding to martingales $R^{(p^*)}$ by using properties (i)–(iv) and writing

$$E[-\exp(-\alpha(X_T^p - F))] \le R_0(x) = E[-\exp(-\alpha(X_T^{p*} - F))] = V(x), \tag{3.9}$$

whence any p^* (there may be more than one) gives an optimal strategy.

To construct such a family, we try to find a function $f : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, adapted in (ω, t), and a pair of adapted control processes (Y, Z) indexed by $t \in [0, T]$ which serve to effectively create a *dynamical version* of the random liability F in the following way. (Y, Z) is a solution of the equation

$$Y_t = Y_0 + \int_0^t Z_s\,dW_s + \int_0^t f(s, Z_s)\,ds, \qquad Y_T = F,$$

which means that f creates a dynamics which controls the process Y into the terminal condition F. Note that if $f = 0$, (Y, Z) will be given by the martingale representation theorem. Written in a backward version, the conditions determining (Y, Z) may be expressed by the simpler equation

$$Y_t = F - \int_t^T Z_s\,dW_s - \int_t^T f(s, Z_s)\,ds, \qquad t \in [0, T]. \tag{3.10}$$

We shall call this a *backward stochastic differential equation (BSDE)* with terminal condition F and generator f. In fact, in the subsequent section, we shall study the problem of existence and uniqueness of solutions for BSDE with generators which, besides ω, time and state z of Z, may also depend on the state y of the process Y, and which are Lipschitz continuous in (y,z). Here, the dynamics introduced by this BSDE is used in order to define the above stipulated family of processes by setting

$$R_t^{(p)} := -\exp(-\alpha(X_t^{(p)} - Y_t)), \quad t \in [0,T],\ p \in \mathscr{A}, \tag{3.11}$$

where (Y,Z) is a solution of (BSDE). Recalling the properties (i)–(iv) we require the family $(R^{(p)})_{p\in\mathscr{A}}$ to fulfill, and noting that by choice of (BSDE) conditions (i) and (ii) are already guaranteed, we are therefore bound to find a function f for which $R^{(p)}$ is a supermartingale for all $p \in \mathscr{A}$ and there exists a $p^* \in \mathscr{A}$ such that $R^{(p^*)}$ is a martingale. This function f, via $\mathscr{A}$, of course also depends on the constraint set C where p takes its values (see (3.5)). We get

$$V(x) = R_0^{(p,x)} = -\exp(-\alpha(x - Y_0)) \quad \text{for all } p \in \mathscr{A}.$$

Let us now calculate f. For $p \in \mathscr{A}$ we write $R^{(p)}$ as the product of a (local) martingale $M^{(p)}$ and a process of bounded variation $\tilde{A}^{(p)}$. For $t \in [0,T]$ define

$$M_t^{(p)} = \exp(-\alpha(x - Y_0)\exp\left(-\int_0^t \alpha(p_s - Z_s)\,dW_s - \frac{1}{2}\int_0^t \alpha^2(p_s - Z_s)^2\,ds\right). \tag{3.12}$$

Comparing $R^{(p)}$ and $M^{(p)}\tilde{A}^{(p)}$ yields

$$\tilde{A}_t^{(p)} = -\exp\left(\int_0^t v(s,p_s,Z_s)\,ds\right), \quad t \in [0,T], \tag{3.13}$$

with

$$v(t,p,z) = -\alpha p\theta_t + \alpha f(t,z) + \frac{1}{2}\alpha^2|p-z|^2. \tag{3.14}$$

Now we turn the requirements that $R^{(p)}$ be a supermartingale for any strategy p and a martingale for a certain p^* into conditions on v and thus f. As $M^{(p)}$ is a (local) martingale, we see that (iii) and (iv) are equivalent to the conditions that $A^{(p)}$ be decreasing for any $p \in \mathscr{A}$ and constant for a certain particular p^*. In other words,

$$\begin{aligned} v(t,p_t,Z_t) &\geq 0 \quad \text{for all } p \in \mathscr{A}, \text{ and} \\ v(t,p_t^*,Z_t) &= 0 \quad \text{for some } p^* \in \mathscr{A}. \end{aligned}$$

To see from these two conditions how f has to be chosen, note that for $t \in [0,T]$ we obtain by completing squares

$$\begin{aligned}\frac{1}{\alpha}v(t,p_t,Z_t) &= \frac{\alpha}{2}|p_t|^2 - \alpha p_t\left(Z_t + \frac{1}{\alpha}\theta_t\right) + \frac{\alpha}{2}|Z_t|^2 + f(t,Z_t)\\ &= \frac{\alpha}{2}\left|p_t - \left(Z_t + \frac{1}{\alpha}\theta_t\right)\right|^2 - \frac{\alpha}{2}\left|Z_t + \frac{1}{\alpha}\theta_t\right|^2 + \frac{\alpha}{2}Z_t^2 + f(t,Z_t)\\ &= \frac{\alpha}{2}\left|p_t - \left(Z_t + \frac{1}{\alpha}\theta_t\right)\right|^2 - Z_t\theta_t - \frac{1}{2\alpha}|\theta_t|^2 + f(t,Z_t).\end{aligned}$$

Recall that p is supposed to be constrained to the closed random set C. So if we set

$$f(\cdot,t,z) = -\frac{\alpha}{2}\operatorname{dist}^2\left(z + \frac{1}{\alpha}\theta_t, C_t(\cdot)\right) + z\theta_t(\cdot) + \frac{1}{2\alpha}|\theta_t(\cdot)|^2,$$

the function is well defined, since so is the distance between a point and a closed set. Note that f is not globally Lipschitz continuous in z. For this choice we obviously get $v \geq 0$. If for a closed set $C \subset \mathbb{R}^n$ and $a \in \mathbb{R}^n$ we set

$$\Pi_C(a) = \{b \in C : |a-b| = \operatorname{dist}(C,a)\},$$

then for

$$p_t^* \in \Pi_{C_t(\omega)}\left(Z_t + \frac{1}{\alpha}\theta_t\right), \quad t \in [0,T], \tag{3.15}$$

we obtain $v(\cdot,p^*,Z) = 0$. Hence the family $(R^{(p)})_{p\in\mathscr{A}}$ has the properties (i)–(iv) of the martingale optimality principle with this choice. We summarize our findings in the following main theorem.

THEOREM 3.1.1. *Suppose that F is bounded and $\mathscr{F}_T$-measurable. Choose*

$$f(\cdot,t,z) = -\frac{\alpha}{2}\operatorname{dist}^2\left(z + \frac{1}{\alpha}\theta_t(\cdot), C_t(\cdot)\right) + z\theta_t(\cdot) + \frac{1}{2\alpha}|\theta(\cdot)|^2. \tag{3.16}$$

Suppose that the BSDE corresponding to the pair (F,f), given by

$$Y_t = F - \int_t^T Z_s\, dW_s - \int_t^T f(s,Z_s)\,ds, \quad t \in [0,T], \tag{3.17}$$

has a solution (Y,Z). Then the value function of the optimization problem (3.9) is given by

$$V(x) = -\exp(-\alpha(x - Y_0)). \tag{3.18}$$

There exists an optimal trading strategy $p^* \in \mathscr{A}$ with

$$p_t^* \in \Pi_{C_t(\cdot)}\left(Z_t + \frac{1}{\alpha}\theta_t\right), \quad t \in [0,T],\ P\text{-a.s.} \tag{3.19}$$

Theorem 3.1.1 states that provided there is a suitable pair (Y,Z) of processes solving (3.17), the problem of optimal investment for our small investor is solved. In the following section, we shall exhibit a theory of existence and uniqueness for BSDE with Lipschitz continuous generators, which can be extended to also cover the generator of 3.1.1. The only thing to verify in addition is that this generator has the required measurability properties, and that p^* is indeed admissible. This will be shown in the following proof.

Proof. 1. In the first part we establish the *measurable selection* result which induces the measurability properties of p^*, given by (3.19).

LEMMA 3.1.1 (measurable selection). *Let $(a_t)_{t\in[0,T]}$, $(\sigma_t)_{t\in[0,T]}$ be $\mathbb{R}^{1\times n}$-valued resp. $\mathbb{R}^{d\times n}$-valued predictable stochastic processes, $\tilde{C} \subset \mathbb{R}^d$ a closed set and $C_t = \tilde{C}\sigma_t$, $t \in [0,T]$.*

(a) *The process*

$$D = (\operatorname{dist}(a_t, C_t))_{t\in[0,T]}$$

is predictable.

(b) *There exists a predictable process a^* with*

$$a_t^* \in \Pi_{C_t}(a_t) \quad \textit{for all } t \in [0,T]. \tag{3.20}$$

Proof. (a) We show the first claim. Observe that D is the composition of continuous mappings with predictable processes. For $k \in \mathbb{N}$ let H^k denote the space of compact subsets of $\mathbb{R}^k$ equipped with the Hausdorff metric and $\mathscr{B}(H^k)$ the Borel sigma-algebra with respect to this metric. The mapping $\operatorname{dist} : \mathbb{R}^n \times H^n \to \mathbb{R}$ is jointly continuous, hence $\mathscr{B}(\mathbb{R}^n) \otimes \mathscr{B}(H^n)$-$\mathscr{B}(\mathbb{R})$-measurable. Now consider $j : \mathbb{R}^{d\times n} \times H^d \to H^n$ that maps a compact subset $\tilde{C}$ in $\mathbb{R}^d$ by applying an $d \times n$-matrix $\tilde{\sigma}$ to a compact subset $\tilde{K}$ of $\mathbb{R}^n$. More formally, j maps $\tilde{C}$ to the following set:

$$\tilde{K} = \{b \in \mathbb{R}^n : \exists \tilde{c} \in \tilde{C},\, b = \tilde{c}\tilde{\sigma}\}.$$

The mapping j is also jointly continuous and therefore $\mathscr{B}(\mathbb{R}^{n\times d}) \otimes \mathscr{B}(H^d)$-$\mathscr{B}(H^n)$-measurable. Hence (a) follows for compact $\tilde{C}$.

If more generally $\tilde{C}$ is closed but not bounded, take $\tilde{C}_m = \tilde{C} \cap B_m$ where B_m is the closed ball with radius m centered at the origin. According to what has already been shown, for $m \in \mathbb{N}$, $\operatorname{dist}(a_t, \tilde{C}_m\sigma_t)$ defines a predictable process and $\operatorname{dist}(a_t, \tilde{C}_m\sigma_t)$ converges to $\operatorname{dist}(a_t, \tilde{C}\sigma_t)$ as $m \to \infty$. This proves the first claim.

(b) In order to prove the second claim, we first concentrate on the case of compact $\tilde{C}$. We have to show that for $z \in \mathbb{R}^n$ and a compact set $\tilde{K} \subset \mathbb{R}^n$ there exists a $\mathscr{B}(\mathbb{R}^n) \otimes \mathscr{B}(H^n)$-$\mathscr{B}(\mathbb{R}^n)$-measurable mapping $\xi(z,\tilde{K})$ with $\xi(z,\tilde{K}) \in \Pi_{\tilde{K}}(z)$. This is achieved by defining of a sequence of mappings $\xi_n(z,\tilde{K})$ with randomly chosen indices that converges to an element of $\Pi_{\tilde{K}}(z)$. The choice of the converging subsequence will depend in a measurable way on z and $\tilde{K}$.

For $m \in \mathbb{N}$, let $G_m = (x_i^m)_{i\in\mathbb{N}}$ be a dyadic grid with $\min_{x\in G_m} \operatorname{dist}(\bar{z}, x) \leq 1/m$ for all $\bar{z} \in \mathbb{R}^n$. Let the elements of the grid G_m be enumerated as $\{g_i^m : i \in \mathbb{N}\}$. Let $\tilde{K}_m$ be the elements of the grid with distance at most $1/m$ from G_m. Since we can describe the sets $\tilde{K}_m$ as the intersections of the discrete set G_m with the closed set of all points in $\mathbb{R}^n$ at distance at most $1/n$ from $\tilde{K}$, and this closed set depends continuously on $\tilde{K}$, it follows that $\tilde{K}_m$ is measurable in $\tilde{K}$. For any $z \in \mathbb{R}^n$, let $\Pi_m(z,\tilde{K})$ be the set of all points in $\tilde{K}_m$ with minimal distance from z. Since $\tilde{K}_m$ is measurable in $\tilde{K}$, $\Pi_m(z,\tilde{K})$ is obviously measurable in $(z,\tilde{K})$. To define $\xi_m(z,\tilde{K})$, we have to choose one point in $\Pi_m(z,\tilde{K})$. Let it be the one with minimal index in the enumeration of G_m. This choice preserves the measurability in $(z,\tilde{K})$. Hence $\xi_m(z,\tilde{K})$ is $\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(H^n)$-$\mathscr{B}(\mathbb{R}^n)$-measurable.

We next have to choose the subsequence with random indices. This is done as in Lemma 1.55 of [11]. Let us argue with respect to an arbitrary probability measure Q on the Borel sets of $\mathbb{R}^n \times H^n$. First note that $\eta = \liminf_{m\to\infty} |\xi_m| < \infty$. Let $\tau_1 = 1$, and

$$\tau_m = \inf\{l > \tau_{m-1} : |\xi_l - \eta| \le 1/l\} \quad \text{for } m > 1.$$

Then it is plain that τ_m is $\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(H^n)$-$\mathscr{B}(\mathbb{R}^n)$-measurable. Now take

$$\xi = \liminf_{m\to\infty} \xi_{\tau_m},$$

which obviously inherits the measurability properties of the subsequence with random indices. But ξ is a selection. Indeed, for every $m \in \mathbb{N}$, $\operatorname{dist}(\xi_{\tau_m}(z,\tilde{K}),\Pi_{\tilde{K}}(z)) \le 1/\tau_m$. Thus by construction, $\xi(z,\tilde{K}) \in \Pi_{\tilde{K}}(z)$ for all $(z,\tilde{K}) \in \mathbb{R}^n \times H^n$.

We may then choose

$$a^* = \xi(a, C\sigma)$$

to satisfy the requirements of the second part of the assertion in the compact case.

Finally, if C is only closed, we may proceed as in the proof for (a).

2. In the second part we show that p^* also satisfies the integrability properties required for admissible strategies.

LEMMA 3.1.2. *Let Z be the second component of a solution of the BSDE* (3.17), *and let p^* be given by Lemma* 3.1.1 *for $a = Z + \frac{1}{\alpha}\theta$. Then the processes*

$$\int_0^\cdot Z_s\,dW_s, \quad \int_0^\cdot p_s^*\,dW_s \tag{3.21}$$

are P-BMO martingales.

Proof. As we shall see in the following section, as a consequence of the boundedness of F, the process Y corresponding to Z in a solution of the BSDE (3.17) is uniformly bounded P-a.s. Let k denote the upper bound. Applying Itô's formula to the nonpositive process $Y-k$, we obtain, for stopping times $\tau \leq T$,

$$E\left[\int_\tau^T Z_s^2\,ds \,\middle|\, \mathscr{F}_\tau\right] = E[(H-k)^2 \,|\, \mathscr{F}_\tau] - |Y_\tau - k|^2$$
$$-2E\left[\int_\tau^T (Y_s-k) f(s,Z_s)\,ds \,\middle|\, \mathscr{F}_\tau\right]$$

The definition of f yields, for all $(t,z) \in [0,T] \times \mathbb{R}^n$,

$$f(t,z) \leq z\theta_t + \frac{1}{2\alpha}|\theta_t|^2.$$

Therefore there exist positive constants c_1, c_2 and $\tilde{c}_1$ such that

$$E\left[\int_\tau^T |Z_s|^2\,ds \,\middle|\, \mathscr{F}_\tau\right] \leq c_1 + c_2 E\left[\int_\tau^T |Z_s+1|\,ds \,\middle|\, \mathscr{F}_\tau\right]$$
$$\leq \tilde{c}_1 + \frac{1}{2}E\left[\int_\tau^T |Z_s|^2\,ds \,\middle|\, \mathscr{F}_\tau\right].$$

Hence, $\int_0^\cdot Z_s\,dW_s$ is a BMO-martingale.

We next deal with the stochastic integral process of p^*. The triangle inequality implies

$$|p^*| \leq \left|Z + \frac{1}{\alpha}\theta\right| + \left|p^* - \left(Z + \frac{1}{\alpha}\theta\right)\right|.$$

The definition of p^* yields, for some constants k_1, k_2,

$$|p_t^*| \leq 2|Z_t| + \frac{2}{\alpha}|\theta_t| + k_1 \leq 2|Z_t| + k_2, \qquad t \in [0,T],$$

and thus for every stopping time $\tau \leq T$,

$$E\left[\int_\tau^T |p_t^*|^2\,dt \,\middle|\, \mathscr{F}_\tau\right] \leq E\left[\int_\tau^T 8|Z_t|^2\,dt + 2Tk_2^2 \,\middle|\, \mathscr{F}_\tau\right].$$

This implies the P-BMO property of $\int_0^\cdot p_s^* dW_s$, and finishes the proof of the theorem. ∎

3.2. Existence theory under Lipschitz conditions

In this chapter we shall establish the basic existence and uniqueness theory for backward stochastic differential equations in case the coefficients are Lipschitz continuous.

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

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

$$\|X\|_{2,\beta}^2 = E\left(\int_0^T e^{\beta t}|X_t|^2\,dt\right), \tag{3.22}$$

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

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

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

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

$$f(\cdot,\cdot,0,0)\in H^2(\mathbb{R}^m). \tag{3.23}$$

Moreover, f is *uniformly Lipschitz*, i.e. there exists $C>0$ such that for any $(y_1,z_1),(y_2,z_2)\in\mathbb{R}^m\times\mathbb{R}^{n\times m}$, $P\otimes\lambda$-a.e. $(\omega,t)\in\Omega\times\mathbb{R}_+$,

$$|f(\omega,t,y_1,z_1)-f(\omega,t,y_2,z_2)|\leq C[|y_1-y_2|+|z_1-z_2|]. \tag{3.24}$$

Here for $z\in\mathbb{R}^{n\times m}$ we write $|z|=(\mathrm{tr}(zz^*))^{1/2}$.

DEFINITION 3.2.1. A pair of functions (f,ξ) satisfying, besides the above mentioned measurability requirements, hypotheses (3.23), (3.24), is said to be a *standard parameter*.

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

$$dY_t = Z_t^* dW_t - f(\cdot,t,Y_t,Z_t)dt,\quad Y_T=\xi, \tag{3.25}$$

is satisfied.

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

LEMMA 3.2.1. *For $i=1,2$, let (f^i,ξ^i) be standard parameters and $(Y^i,Z^i)\in H^2(\mathbb{R}^m)\times H^2(\mathbb{R}^{n\times m})$ solutions of* (3.25) *with the corresponding standard parameters. Let C be a Lipschitz constant for f^1. For $0\leq t\leq T$ define*

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

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

$$||\delta Y||^2_{2,\beta} \leq T\left[e^{\beta T}E(|\delta Y_T|^2)+\frac{1}{\mu^2}||\delta_2 f||^2_{2,\beta}\right], \tag{3.26}$$

$$||\delta Z||^2_{2,\beta} \leq \frac{\lambda^2}{\lambda^2 - C}\left[e^{\beta T}E(|\delta Y_T|^2)+\frac{1}{\mu^2}||\delta_2 f||^2_{2,\beta}\right]. \tag{3.27}$$

Proof. 1. Let $(Y,Z)\in H^2(\mathbb{R}^m)\times H^2(\mathbb{R}^{n\times m})$ be a solution of (3.25) with standard parameters (f,ξ). This means that for $0\leq t\leq T$ we may write

$$Y_t = \xi - \int_t^T Z_s^* \, dW_s + \int_t^T f(\cdot, s, Y_s, Z_s)\, ds. \tag{3.28}$$

We show that

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

In fact, by (3.28) we have

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

and, with the help of Doob's inequality,

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

Since in addition (3.23) and (3.24) guarantee that $|\xi|+\int_0^T|f(\cdot,s,Y_s,Z_s)|\,ds\in L^2(\mathbb{R})$, we obtain the desired

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

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

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

By reordering the terms in the equation we obtain

$$
\begin{aligned}
e^{\beta t}|\delta Y_t|^2 + \beta \int_t^T e^{\beta s}|\delta Y_s|^2\,ds + \int_t^T e^{\beta s}|\delta Z_s|^2\,ds \\
= e^{\beta T}|\delta Y_T|^2 + 2\int_t^T e^{\beta s}\langle \delta Y_s, \delta Z_s^* dW_s\rangle \\
- 2\int_t^T e^{\beta s}\langle \delta Y_s, f^1(\cdot,s,Y_s^1,Z_s^1) - f^2(\cdot,s,Y_s^2,Z_s^2)\rangle\,ds. \qquad (3.29)
\end{aligned}
$$

3. We prove that for $0 \le t \le T$,

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

To prove this, first take expectations on both sides of the inequality (3.29) to get

$$
\begin{aligned}
E(e^{\beta t}|\delta Y_t|^2) + \beta E\left(\int_t^T e^{\beta s}|\delta Y_s|^2\,ds\right) + E\left(\int_t^T e^{\beta s}|\delta Z_s|^2\,ds\right) \\
\le E(e^{\beta T}|\delta Y_T|^2) \\
+ 2E\left(\int_t^T e^{\beta s}\langle \delta Y_s, f^1(\cdot,s,Y_s^1,Z_s^1) - f^2(\cdot,s,Y_s^2,Z_s^2)\rangle\,ds\right). \qquad (3.30)
\end{aligned}
$$

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

$$
\begin{aligned}
|f^1(\cdot,s,Y_s^1,Z_s^1) - f^2(\cdot,s,Y_s^2,Z_s^2)| &\le |f^1(\cdot,s,Y_s^1,Z_s^1) - f^1(\cdot,s,Y_s^2,Z_s^2)| + |\delta_2 f_s| \\
&\le C[|\delta Y_s| + |\delta Z_s|] + |\delta_2 f_s|. \qquad (3.31)
\end{aligned}
$$

(3.31) implies

$$
\begin{aligned}
\int_t^T E(2e^{\beta s}|\langle \delta Y_s, f^1(\cdot,s,Y_s^1,Z_s^1) - f^2(\cdot,s,Y_s^2,Z_s^2)\rangle|\,ds \\
\le \int_t^T 2e^{\beta s} E(|\delta Y_s|[C(|\delta Y_s| + |\delta Z_s|) + |\delta_2 f_s|])\,ds \\
= \int_t^T 2e^{\beta s}[CE(|\delta Y_s|^2) + E(|\delta Y_s|(C|\delta Z_s|) + |\delta_2 f_s|)]\,ds. \qquad (3.32)
\end{aligned}
$$

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

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

Using this we can estimate the last term in (3.32) further:

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

Summing our estimates in (3.30), we obtain, using our assumptions on the parameters,

$$\begin{aligned}
E(e^{\beta t}|\delta Y_t|^2)&\le E\left(\int_t^T e^{\beta s}|\delta Y_s|^2\,ds\right)[-\beta+C(2+\lambda^2)+\mu^2]\\
&\quad+E\left(\int_t^T e^{\beta s}|\delta Z_s|^2\,ds\right)\left[\frac{C}{\lambda^2}-1\right]+E(e^{\beta T}|\delta Y_T|^2)\\
&\quad+\frac{1}{\mu^2}E\left(\int_t^T e^{\beta s}|\delta_2 f_s|^2\,ds\right)+E(e^{\beta T}|\delta Y_T|^2)\\
&\le E(e^{\beta T}|\delta Y_T|^2)+\frac{1}{\mu^2}E\left(\int_t^T e^{\beta s}|\delta_2 f_s|^2\,ds\right). \qquad (3.33)
\end{aligned}$$

This is the claimed inequality.

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

5. The second inequality in the assertion follows from (3.33) by moving 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 (3.28).

THEOREM 3.2.1. *Let (f,ξ) be standard parameters. Then there exists a unique pair $(Y,Z)\in H^2(\mathbb{R}^m)\times H^2(\mathbb{R}^{n\times m})$ with the property*

$$Y_t=\xi-\int_t^T Z_s^*\,dW_s+\int_t^T f(\cdot,s,Y_s,Z_s)\,ds,\qquad 0\le t\le T. \qquad (3.34)$$

Proof. Consider

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

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

$$Y_t = \xi - \int_t^T Z_s^* \, dW_s + \int_t^T f(\cdot, s, y_s, z_s) \, ds, \quad 0 \le t \le T. \tag{3.35}$$

1. We prove that (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

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

is a well defined martingale. M has a continuous version, as we are working in a Wiener filtration. Moreover, M is square integrable. Hence we may apply the martingale representation theorem, which provides (a unique) $Z \in H^2(\mathbb{R}^{n\times m})$ 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

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

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 $0 \le t \le T$,

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

2. We prove that for $\beta > 2(1+T)C$ the mapping Γ is a contraction. For this purpose, let $(y^1,z^1),(y^2,z^2) \in H^{2,\beta}(\mathbb{R}^m) \times H^{2,\beta}(\mathbb{R}^{n\times m})$, and let (Y^1,Z^1), (Y^2,Z^2) be the corresponding solutions of (3.35) according to part 1. We apply

Lemma 3.2.1 with $C = 0$, $\beta = \mu^2$, and $f^i = f(\cdot, y^i, z^i)$. With this choice we obtain

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

Since f is Lipschitz continuous, we further obtain

$$||\delta Y||_{2,\beta} \leq \frac{2TC}{\beta}\,[||\delta y||_{2,\beta}+||\delta z||_{2,\beta}], \tag{3.36}$$
$$||\delta Z||_{2,\beta} \leq \frac{2C}{\beta}\,[||\delta y||_{2,\beta}+||\delta z||_{2,\beta}]. \tag{3.37}$$

We sum these to obtain

$$||\delta Y||_{2,\beta}+||\delta Z||_{2,\beta} \leq \frac{2C(T+1)}{\beta}\,[||\delta y||_{2,\beta}+||\delta z||_{2,\beta}]. \tag{3.38}$$

By the choice of β, Γ is a contraction.

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

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

Then Y is continuous and P-a.s. identical to $\overline{Y}$. Then $(Y,\overline{Z})$ is a solution of (3.34).

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 see in the following corollary.

COROLLARY 3.2.1. *Let $\beta > 2(1+T)C$, and let $((Y^k,Z^k))_{k\geq 0}$ be the sequence of processes given by $Y^0 = Z^0 = 0$,*

$$Y_t^{k+1} = \xi - \int_t^T (Z_s^{k+1})^*\,dW_s + \int_t^T f(\cdot,s,Y_s^k,Z_s^k)\,ds \tag{3.39}$$

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

Proof. The inequality (3.33) recursively yields

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

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

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

Now a standard argument applies. ■

3.3. Interpretation of solutions in Malliavin's calculus

In this chapter we shall establish the crucial 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. To do this, we first have to introduce the version of the space $\mathbf{L}_1^2$ which corresponds to integrability to some arbitrary power $p \geq 2$. For simplicity, we let the dimension of our underlying Wiener process be one, i.e. for this chapter we set $n = 1$.

DEFINITION 3.3.1. Let $p \geq 2$, and

$$\begin{aligned}\mathbf{L}_1^p(\mathbb{R}^m) = \Big\{ u : u \text{ adapted}, &\left[\int_0^T |u_t|^2\, dt\right]^{1/2} \in L^p(\Omega),\\ &u_t \in (D_1^p)^m \text{ for } \lambda\text{-a.a. } t \geq 0, \text{ and}\\ &\text{for some measurable version of } (s,t) \mapsto D_s u_t\\ &\text{we have } E\left(\left[\int_0^T \int_0^T |D_s u_t|^2\, ds\, dt\right]^{p/2}\right) < \infty \Big\}.\end{aligned} \tag{3.40}$$

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

$$||u||_{1,p} = E\left(\left[\int_0^T |u_t|^2\, dt\right]^{p/2}\right) + E\left(\left[\int_0^T \int_0^T |D_s u_t|^2\, ds\, dt\right]^{p/2}\right). \tag{3.41}$$

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

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

In this notation, Jensen's inequality gives

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

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 $S^p(\mathbb{R}^m)$ the linear space of all measurable $(\mathscr{F}_t)$-adapted continuous processes $X : \Omega \times [0,T] \to \mathbb{R}^m$, endowed with the norm $||X||_{S^p} = E(\sup_{0\leq t\leq T} |X_t|^p)^{1/p}$. Let further $H^p(\mathbb{R}^m)$ denote the linear space of measurable $(\mathscr{F}_t)$-adapted processes $X : \Omega\times[0,T] \to \mathbb{R}^m$ endowed with the norm $||X||_p = E(||X||^p)^{1/p}$. To abbreviate, let $B^p(\mathbb{R}^m) = S^p(\mathbb{R}^m) \times H^p(\mathbb{R}^m)$, with the norm $||(Y,Z)||_p = [||Y||^p_{S^p} + ||Z||^p_p]^{1/p}$.

We are ready to state our main result.

THEOREM 3.3.1. *Let (f,ξ) be standard parameters such that $\xi \in D^2_1 \cap L^4(\mathbb{R}^m)$, $f : \Omega\times[0,T]\times\mathbb{R}^m\times\mathbb{R}^m \to \mathbb{R}^m$ is continuously differentiable in (y,z), with uniformly bounded and continuous partial derivatives, and such that for $(y,z)\in\mathbb{R}^m\times\mathbb{R}^m$ we have*

$$f(\cdot,y,z)\in \mathbf{L}^2_1, \quad f(\cdot,0,0)\in H^4(\mathbb{R}^m), \tag{3.42}$$

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

$$|D_s f(\cdot,t,y^1,z^1) - D_s f(\cdot,t,y^2,z^2)| \le K_s(t)[|y^1-y^2|+|z^1-z^2|] \tag{3.43}$$

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

$$\int_0^T ||K_s||_4^4\, ds < \infty. \tag{3.44}$$

For the unique solution (Y,Z) of the BSDE (3.34) *we moreover suppose*

$$\int_0^T ||D_s f(\cdot,Y,Z)||^2\, ds < \infty \quad P\text{-a.s.} \tag{3.45}$$

Then

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

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

$$\begin{aligned}
&D_sY_t = D_sZ_t = 0, \quad 0\le t<s\le T,\\
&D_sY_t = D_s\xi - \int_t^T D_sZ^*_u\, dW_u\\
&\qquad + \int_t^T \Big[\frac{\partial}{\partial y} f(\cdot,u,Y_u,Z_u)D_sY_u + \frac{\partial}{\partial z} f(\cdot,u,Y_u,Z_u)D_sZ_u\\
&\qquad + D_s f(\cdot,u,Y_u,Z_u)\Big]\, du, \quad 0\le s\le t\le T, \qquad (3.46)\\
&(D_sY_s)_{0\le s\le T} \text{ is a version of } (Z_s)_{0\le s\le T}. \qquad (3.47)
\end{aligned}$$

Proof. 1. To further simplify notation, we assume $m = 1$. As in Chapter 3.2, our arguments are mainly based upon several *a priori* estimates. The first one is an analogue of Lemma 3.2.1 and concerns the properties of the contraction map on $B^2(\mathbb{R})$.

LEMMA 3.3.1. *Let $p \geq 2$, assume $f(\cdot,0,0) \in H^p(\mathbb{R})$, and define*

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

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

$$Y_t = \xi - \int_t^T Z_s \, dW_s + \int_t^T f(\cdot,s,y_s,z_s)\,ds. \tag{3.48}$$

Let further, for $i=1,2$, (Y^i,Z^i) be the solutions corresponding to (y^i,z^i) in (3.48), *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

$$||Y||_{S^p}^p \leq C_p E\left(|\xi|^p + T^{p/2}\left(\int_0^T |f(\cdot,s,y_s,z_s)|^2\,ds\right)^{p/2}\right), \tag{3.49}$$

$$||Z||_p^p \leq C_p E\left(|\xi|^p + T^{p/2}\left(\int_0^T |f(\cdot,s,y_s,z_s)|^2\,ds\right)^{p/2}\right), \tag{3.50}$$

$$||(\delta Y,\delta Z)||_p^p \leq C_p T^{p/2} ||(\delta y,\delta z)||_p^p. \tag{3.51}$$

Proof. (a) Taking up the notation of the proof of Lemma 3.2.1, we show that Γ is well defined. To do this, recall that for $0 \leq t \leq T$,

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

We have

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

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

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

Moreover, by Cauchy–Schwarz' inequality

$$\int_0^T |f(\cdot,s,y_s,z_s)|\,ds \leq T^{1/2}\left[\int_0^T |f(\cdot,s,y_s,z_s)|^2\,ds\right]^{1/2},$$

hence with another universal constant C_p^2 we have

$$||Y||_{S^p}^p \leq C_p^2 E\left(|\xi|^p + \left[\int_0^T |f(\cdot,s,y_s,z_s)|^2\,ds\right]^{p/2}\right). \tag{3.52}$$

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

We next prove that $Z \in H^p(\mathbb{R})$. For this purpose we shall use the inequality of Burkholder. It yields further universal constants C_p^3, C_p^4, C_p^5 such that

$$\begin{aligned}
E(\|Z\|^p) &\le C_p^3 E\left(\left|\int_0^T Z_s\, dW_s\right|^p\right) \\
&\le C_p^4 E\left(\left|\xi + \int_0^T f(\cdot, s, y_s, z_s)\, ds - Y_0\right|^p\right) \\
&\le C_p^5 E\left(|\xi|^p + T^{p/2}\left[\int_0^T |f(\cdot, s, y_s, z_s)|^2\, ds\right]^{p/2}\right). \qquad (3.53)
\end{aligned}$$

Hence $Z \in H^p(\mathbb{R})$, and therefore $(Y, Z) \in B^p(\mathbb{R})$. The inequalities (3.49) and (3.50) have also been established.

(b) We prove (3.51). 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 (3.49) and (3.50), as well as an appeal to the Lipschitz condition, give, with universal constants C_p^6, C_p^7,

$$\begin{aligned}
\|(\delta Y, \delta Z)\|_p^p &\le C_p^6 T^{p/2} E\left(\left[\int_0^T |f(\cdot, t, y_t^1, z_t^1) - f(\cdot, t, y_t^2, z_t^2)|^2\, dt\right]^{p/2}\right) \\
&\le C_p^7 T^{p/2}[\|\delta y\|_{S^p}^p + \|\delta z\|_p^p] \\
&= C_p^7 T^{p/2}\|(\delta y, \delta z)\|_p^p. \qquad (3.54)
\end{aligned}$$

This completes the proof. ∎

Let us return to the proof of Theorem 3.3.1. We define approximations of the solution of the BSDE recursively. For $k \ge 0$ and $0 \le t \le T$ let

$$\begin{aligned}
Y^0 &= Z^0 = 0, \\
Y_t^{k+1} &= \xi - \int_t^T Z_s^{k+1}\, dW_s + \int_t^T f(\cdot, s, Y_s^k, Z_s^k)\, ds. \qquad (3.55)
\end{aligned}$$

We show that

$$\|(Y^k, Z^k) - (Y, Z)\|_4 \to 0 \quad (k \to \infty). \qquad (3.56)$$

Recall the universal constant C_4 from (3.51) of Lemma 3.3.1. We may (modulo repeating the argument finitely many times on successive subintervals of $[0, T]$) assume that $[C_4 T^2]^{1/4} < 1$. With this condition, Lemma 3.3.1 implies that Γ is a contraction, and the solution (Y, Z) of the BSDE is its unique fixed point in $B^4(\mathbb{R})$. From this observation, we obtain our assertion since the approximate solutions form a Cauchy sequence, as follows from

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

2. We prove by recursion on k that

$$(Y^k, Z^k) \in \mathbf{L}_1^2(\mathbb{R}) \times \mathbf{L}_1^2(\mathbb{R}). \tag{3.57}$$

This is trivial for $k = 0$. Assume that it holds for k. According to the chain rule for the Malliavin derivative and our hypotheses concerning the standard parameters we know that for $0 \le t \le T$,

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

with Malliavin derivative

$$D_s \xi + \int_t^T \left[\frac{\partial}{\partial y} f(\cdot, u, Y_u^k, Z_u^k) D_s Y_u^k + \frac{\partial}{\partial z} f(\cdot, u, Y_u^k, Z_u^k) D_s Z_u^k \right.$$
$$\left. + D_s f(\cdot, u, Y_u^k, Z_u^k) \right] du. \tag{3.58}$$

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 2.3.2 yields, for fixed $0 \le t \le T$,

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

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 2.3.4 implies that $Z^{k+1} \in \mathbf{L}_1^2(\mathbb{R})$, and in $L^2(\Omega \times [0, T]^2)$ we have the equation

$$D_s \int_t^T Z_u^{k+1}\, dW_u = \int_t^T D_s Z_u^{k+1}\, dW_u, \qquad s \le t,$$
$$D_s \int_t^T Z_u^{k+1}\, dW_u = Z_s^{k+1} + \int_s^T D_s Z_u^{k+1}\, dW_u, \qquad s > t.$$

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

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

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

$$D_s X_t^{k+1} = D_s \xi - \int_t^T D_s Z_u^{k+1}\, dW_u \\ + \int_t^T \Big[\frac{\partial}{\partial y} f(\cdot, u, Y_u^k, Z_u^k) D_s Y_u^k + \frac{\partial}{\partial z} f(\cdot, u, Y_u^k, Z_u^k) D_s Z_u^k \\ + D_s f(\cdot, u, Y_u^k, Z_u^k) \Big]\, du. \tag{3.59}$$

3. In this step we show that

$$(DY^k, DZ^k) \to (Y^\cdot, Z^\cdot) \quad \text{in } L^2(\Omega \times [0,T]^2), \tag{3.60}$$

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

$$Y_t^s = D_s \xi - \int_t^T Z_u^s\, dW_u \\ + \int_t^T \Big[\frac{\partial}{\partial y} f(\cdot, u, Y_u, Z_u) Y_u^s + \frac{\partial}{\partial z} f(\cdot, u, Y_u, Z_u) Z_u^s \\ + D_s f(\cdot, u, Y_u, Z_u) \Big]\, du, \quad 0 \le s \le t \le T, \tag{3.61}$$

$$Y_t^s = Z_t^s = 0, \quad 0 \le t < s \le T.$$

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

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

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

LEMMA 3.3.2. *Let (f^i, ξ^i), $i = 1, 2$, be standard parameters of a BSDE, $p \ge 2$. Suppose*

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

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

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

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

$$\|(\delta Y, \delta Z)\|_p^p \le C_{p,T} \left[E(|\delta Y_T|^p) + E\left(\left(\int_0^T |\delta_2 f_s|\, ds \right)^p \right) \right] \\ \le C_{p,T} [E(|\delta Y_T|^p) + T^{p/2} \|\delta_2 f_s\|^p]. \tag{3.62}$$

Proof. With a calculation analogous to the one used to prove (3.49) and (3.50) in Lemma 3.3.1 we arrive at the following inequality which is valid with universal constants C_p^1, C_p^2, C_p^3, and for which we also use Doob's and Cauchy–Schwarz' inequalities

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

Now choose T small enough to ensure $C_p^3(T^p + T^{p/2}) < 1$. Having done this, we may shift 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 3.3.2 to prove that for λ-a.a. $0 \le s \le T$ we have $(Y^s, Z^s) \in B^2(\mathbb{R})$. For this purpose, we apply the lemma with

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

for $0 \le t \le T$ and $y, z \in \mathbb{R}$. Then we have

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

We obtain with some universal constant C the inequality

$$||(Y^s, Z^s)||_2^2 \le C E(|D_s \xi|^2 + ||D_s f(\cdot,Y,Z)||^2),$$

and therefore

$$\int_0^T ||(Y^s, Z^s)||_2^2\,ds < \infty. \tag{3.63}$$

This implies the desired integrability.

To obtain estimates for differences of (DY^k, DZ^k) and $(Y^\cdot, Z^\cdot)$, let us next, fixing $k \in \mathbb{N}$, apply Lemma 3.3.2 to the following parameters:

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

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

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

The lemma yields the inequality

$$||(D_s Y^{k+1} - Y^s, D_s Z^{k+1} - Z^s)||_2^2 \le C_1 E\left(\left(\int_s^T |\delta_t^k| \, dt\right)^2\right) \tag{3.64}$$

with a universal constant C_1. Let us now further estimate the right hand side of this inequality. We have, fixing $0 \le s \le T$,

$$E\left(\left(\int_s^T |\delta_t^k| \, dt\right)^2\right) \le C_2 [A_k^s(T) + B_k^s(T) + C_k^s(T)], \tag{3.65}$$

where

$$A_k^s(T) = E\left(\left[\int_s^T |D_s f(\cdot, t, Y_t, Z_t) - D_s f(\cdot, t, Y_t^k, Z_t^k)| \, dt\right]^2\right), \tag{3.66}$$

$$\begin{aligned} B_k^s(T) = {} & E\left(\left[\int_s^T \left|\frac{\partial}{\partial y} f(\cdot, t, Y_t^k, Z_t^k)(Y_t^s - D_s Y_t^k)\right| dt\right]^2\right) \\ & + E\left(\left[\int_s^T \left|\frac{\partial}{\partial z} f(\cdot, t, Y_t^k, Z_t^k)(Z_t^s - D_s Z_t^k)\right| dt\right]^2\right), \end{aligned} \tag{3.67}$$

$$\begin{aligned} C_k^s(T) = {} & E\left(\left[\int_s^T \left|\frac{\partial}{\partial y} f(\cdot, t, Y_t, Z_t) - \frac{\partial}{\partial y} f(\cdot, t, Y_t^k, Z_t^k)\right| |Y_t^s| \, dt\right]^2\right) \\ & + E\left(\left[\int_s^T \left|\frac{\partial}{\partial z} f(\cdot, t, Y_t, Z_t) - \frac{\partial}{\partial z} f(\cdot, t, Y_t^k, Z_t^k)\right| |Z_t^s| \, dt\right]^2\right). \end{aligned} \tag{3.68}$$

With further universal constants C_3, C_4 we deduce, using (3.43), that

$$\begin{aligned} A_k^s(T) & \le E\left(\left[\int_s^T K_s(t)[|Y_t - Y_t^k| + |Z_t - Z_t^k|] \, dt\right]^2\right) \\ & \le C_3 E\left(\int_s^T K_s(t)^2 \, dt \left[\int_s^T |Y_t - Y_t^k|^2 \, dt + \int_s^T |Z_t - Z_t^k|^2 \, dt\right]\right) \\ & \le C_4 \left[E\left(\int_s^T K_s(t)^4 \, dt\right)\right]^{1/2} \\ & \quad \times \left[E\left(\int_s^T |Y_t - Y_t^k|^4 \, dt\right)^{1/2} + E\left(\int_s^T |Z_t - Z_t^k|^4 \, dt\right)^{1/2}\right]. \end{aligned}$$

Hence by part 1 of the proof,

$$\lim_{k\to\infty}\int_0^T A_k^s(T)\,ds = 0. \tag{3.69}$$

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 us to conclude that

$$\lim_{k\to\infty}\int_0^T C_k^s(T)\,ds = 0. \tag{3.70}$$

Let us finally discuss the convergence of $B_k^s(T)$ as $k\to\infty$. Again by boundedness of the partial derivatives of f we obtain, with a universal constant C_5,

$$B_k^s(T) \le C_5 T^2 \|(D_s Y^k - Y^s, D_s Z^k - Z^s)\|_2^2.$$

Now choose T small enough to ensure $\alpha = C_5 T^2 < 1$. Let $\varepsilon > 0$. Then by what has been shown there exists $N\in\mathbb{N}$ large enough so that for $k\ge N$ we have

$$\begin{aligned} &E\left(\int_0^T \|(D_s Y^{k+1} - Y^s, D_s Z^{k+1} - Z^s)\|^2\,ds\right) \\ &\qquad\qquad \le \varepsilon + \alpha E\left(\int_0^T \|(D_s Y^k - Y^s, D_s Z^k - Z^s)\|^2\,ds\right). \end{aligned}$$

By induction we obtain, for $k\ge N$,

$$\begin{aligned} &E\left(\int_0^T \|(D_s Y^k - Y^s, D_s Z^k - Z^s)\|^2\,ds\right) \\ &\qquad \le \varepsilon(1+\alpha+\alpha^2+\cdots+\alpha^{k-N-1}) \\ &\qquad\quad + \varepsilon\alpha^{k-N} E\left(\int_0^T \|(D_s Y^N - Y^s, D_s Z^N - Z^s)\|^2\,ds\right) \\ &\qquad \le \frac{\varepsilon}{1-\alpha} + \alpha^{k-N} E\left(\int_0^T \|(D_s Y^N - Y^s, D_s Z^N - Z^s)\|^2\,ds\right). \end{aligned}$$

Now let $k\to\infty$. Since ε is arbitrary, we conclude that

$$\lim_{k\to\infty}\int_0^T B_k^s(T)\,ds = 0. \tag{3.71}$$

3. Since $\mathbf{L}_1^2(\mathbb{R})$ is a Hilbert space, and D is a closed operator, we infer that $(Y,Z)\in \mathbf{L}_1^2(\mathbb{R})\times\mathbf{L}_1^2(\mathbb{R})$, and $(Y^s,Z^s)_{0\le s\le T}$ is a version of $(D_sY, D_sZ)_{0\le s\le T}$ in the usual sense.

4. We show that

$$(D_t Y_t)_{0\le t\le T} \text{ is a version of } (Z_t)_{0\le t\le T}. \tag{3.72}$$

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 2.3.4, for $t < u \leq s$,

$$\begin{aligned} D_u Y_s = Z_u &+ \int_u^s D_u Z_r \, dW_r \\ &- \int_u^s \left[\frac{\partial}{\partial y} f(\cdot, r, Y_r, Z_r) D_u Y_r + \frac{\partial}{\partial z} f(\cdot, r, Y_r, Z_r) D_u Z_r \right. \\ &\left. + D_u f(\cdot, r, Y_r, Z_r) \right] dr. \end{aligned}$$

By continuity in t of (Y^s, Z^s) we may choose $u = s$ to obtain the desired identity. ■

References

[1] S. Ankirchner and P. Imkeller, *Quadratic hedging of weather and catastrophe risk by using short term climate predictions*, preprint, HU Berlin, 2008.

[2] S. Ankirchner, P. Imkeller and A. Popier, *On measure solutions of backward stochastic differential equations*, preprint, HU Berlin, 2008.

[3] S. Ankirchner, P. Imkeller and G. Reis, *Classical and variational differentiability of BSDEs with quadratic growth*, Electron. J. Probab. 12 (2007), 1418–1453.

[4] —, —, —, *Pricing and hedging of derivatives based on non-tradable underlyings*, preprint, HU Berlin, 2007.

[5] D. Bell, *The Malliavin Calculus*, Pitman Monogr. Surveys Pure Appl. Math. 34, Longman and Wiley, 1987.

[6] —, *Degenerate Stochastic Differential Equations and Hypoellipticity*, Pitman Monogr. Surveys Pure Appl. Math. 79, Longman, Harlow, 1995.

[7] J. M. Bismut, *Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions*, Z. Wahrsch. Verw. Gebiete 56 (1981), 469–505.

[8] —, *Large Deviations and the Malliavin Calculus*, Progr. Math. 45, Birkhäuser, Basel, 1984.

[9] N. Bouleau and F. Hirsch, *Dirichlet Forms and Analysis on Wiener Space*, de Gruyter, Berlin, 1991.

[10] N. El Karoui, S. Peng and M. C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance 7 (1997), 1–71.

[11] H. Föllmer and A. Schied, *Stochastic Finance. An Introduction in Discrete Time*, Stud. Math. 27, de Gruyter, Berlin, 2002.

[12] E. Fournié, J.-M. Lasry, J. Lebouchoux, P.-L. Lions and N. Touzi, *Applications of Malliavin's calculus to Monte Carlo methods in finance*, Finance Stochastics 3 (1999), 391–412.

[13] Z-Y. Huang and J-A. Yan, *Introduction to Infinite Dimensional Stochastic Analysis*, Math. Appl. 502, Kluwer, Dordrecht, 2000.

[14] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, 1989.

[15] Y. Hu, P. Imkeller and M. Müller, *Utility maximization in incomplete markets*, Ann. Appl. Probab. 15 (2005), 1691–1712.

[16] P. Imkeller, *Malliavin's calculus in insider models: additional utility and free lunches*, Math. Finance 13 (2003), 153–169.

[17] P. Imkeller, *Enlargement of the Wiener filtration by an absolutely continuous random variable via Malliavin's calculus*, Probab. Theory Related Fields 106 (1996), 105–135.

[18] P. Imkeller, V. Perez-Abreu and J. Vives, *Chaos expansions of double intersection local times of Brownian motion in $\mathbb{R}^d$ and renormalization*, Stochastic Process. Appl. 56 (1995), 1–34.

[19] P. Imkeller, M. Pontier and F. Weisz, *Free lunch and arbitrage possibilities in a financial market model with an insider*, Stochastic Process. Appl. 92 (2001), 103–130.

[20] N. Kazamaki, *Continuous Exponential Martingales and BMO*, Lecture Notes in Math. 1579, Springer, Berlin, 1994.

[21] H. H. Kuo, *Donsker's delta function as a generalized Brownian functional and its application*, in: Lecture Notes in Control Inform. Sci. 49, Springer, Berlin, 1983, 167–178.

[22] S. Kusuoka and D. W. Stroock, *The partial Malliavin calculus and its applications to nonlinear filtering*, Stochastics 12 (1984), 83–142.

[23] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Math. 1702, Springer, Berlin, 1999.

[24] P. Malliavin, *Stochastic Analysis*, Grundlehren Math. Wiss. 313, Springer, Berlin, 1997.

[25] —, *Integration and Probability*, Grad. Texts in Math. 157, Springer, New York, 1995.

[26] —, *Stochastic calculus of variations and hypoelliptic operators*, in: Proc. Intern. Sympos. SDE (Kyoto, 1976), Kinokuniya, Tokyo, 1978, 195-263.

[27] —, *C^k-hypoellipticity with degeneracy*, in: Stochastic Analysis, A. Friedman and M. Pinsky (eds.), Academic Press, New York, 1978, 199–214, 327–340.

[28] P. Malliavin and A. Thalmaier, *Stochastic Calculus of Variations in Mathematical Finance*, Springer, Berlin, 2006.

[29] J. C. Mattingly and E. Pardoux, *Malliavin calculus for the stochastic 2D Navier Stokes equation*, Comm. Pure Appl. Math. 59 (2006), 1742–1790.

[30] Y. S. Mishura, *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, Lecture Notes in Math. 1969, Springer, Berlin, 2008.

[31] J. Norris, *Simplified Malliavin calculus*, in: Séminaire de Probabilités XX, Lecture Notes in Math. 1204, Springer, Berlin, 1986, 101–130.

[32] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer, Berlin, 1995.

[33] —, *Malliavin's Calculus and Anticipative Calculus*, St Flour Lecture Notes, 1996.

[34] D. Nualart and E. Pardoux, *Stochastic calculus with anticipating integrands*, Probab. Theory Related Fields 78 (1988), 535–581.

[35] N. Privault and A. Réveillac, *Stein estimation of the drift of Gaussian processes using Malliavin calculus*, Ann. Statist., to appear.

[36] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer, Berlin, 1991.

[37] D. W. Stroock, *The Malliavin calculus. Functional analytic approach*, J. Funct. Anal. 44 (1981), 212–257.

[38] —, *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.

[39] —, *Some applications of stochastic calculus to partial differential equations*, in: École d'Été de Probabilité de St Flour, Lecture Notes in Math. 976, Springer, 1983, 267–382.

[40] S. Watanabe, *Stochastic Differential Equations and Malliavin Calculus*, Tata Inst. Fund. Res., Springer, Berlin, 1984.