Malliavin's calculus and applications in stochastic control and finance
Lecture Course at U Luxemburg
October-December 2020

Peter Imkeller
Institut für Mathematik
Humboldt-Universität zu Berlin

10. Dezember 2020

Malliavin's calculus: originally study of smoothness of measures on infinite dimensional spaces, typically laws of solutions of parabolic PDE; leads to measures on path spaces.

Finite dimensional setting: smoothness of measures clear. Start interpreting Gaussian (Wiener) space as sequence space, extend theory of differentiation and integration in Euclidean spaces; transfer to path spaces.

Applications to many areas of stochastics, e. g. stochastic finance. Here: interpretation of backward stochastic differential equations (BSDE), key stochastic tool for utility optimization in incomplete financial markets.

Here is the sequence space access.

1 The Wiener space as sequence space

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

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

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

Example 1:
Let $\Omega = C(\mathbf{R}_+, \mathbf{R})$, $\mathcal{F}$ Borel sets on Ω generated by topology of uniform convergence on compacts of $\mathbf{R}_+$, P canonical Wiener measure on $\mathcal{F}$. Let W be the canonical Wiener process.

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

0. Prove this for $\Omega = C([0,1],\mathbf{R})$ instead. Then extend by a sequence of copies of this space, arguing for $W^n(t) = W(t+n) - W(n), t \in [0,1], n \in \mathbf{N}$, use independence of increments.

1. **A particular CONS in $L^2([0,1])$:**
Haar functions $(h_{pm}, p \geq 0, 1 \leq m \leq 2^p)$. They form a *wavelet*: system of functions created by translations and scalings from one *mother* function.

Mother function:

$$h_{01}(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2} \leq t \leq 1, \\ 0, & \text{else.} \end{cases}$$

(define on $\mathbf{R}$, then restrict to $[0,1]$). Therewith

$$\begin{aligned} h_{00}(t) &= 1_{[0,1]}, \\ h_{11}(t) &= \sqrt{2}h_{01}(2t), \\ h_{12}(t) &= \sqrt{2}h_{01}(2t-1), \\ h_{21}(t) &= 2^{\frac{2}{2}}h_{01}(4t), \\ h_{22}(t) &= 2^{\frac{2}{2}}h_{01}(4t-1), \\ h_{23}(t) &= 2^{\frac{2}{2}}h_{01}(4t-2), \\ h_{24}(t) &= 2^{\frac{2}{2}}h_{01}(4t-3). \end{aligned}$$

In general for $p \geq 0$, $1 \leq m \leq 2^p$:

$$h_{pm}(t) := \begin{cases} 2^{\frac{p}{2}}, & t \in \left[\frac{m-1}{2^p}, \frac{2m-1}{2^{p+1}}\right), \\ -2^{\frac{p}{2}}, & t \in \left[\frac{2m-1}{2^{p+1}}, \frac{m}{2^p}\right), \\ 0, & \text{otherwise.} \end{cases}$$

It is clear that $(h_{00}, h_{pm})_{p\geq 0, 1\leq m\leq 2^p})$ is an ONS in $L^2([0,1])$. To see this, observe that

$$\langle h_{pn}, h_{pm}\rangle = 0,$$

if $m \neq n$, since the supports of the two functions are disjoint, and

$$\langle h_{pn}, h_{pn}\rangle = 2^p\lambda([\frac{m-1}{2^p}, \frac{m}{2^p}]) = 1.$$

And if $p, q \in \mathbb{N}$ are different, then for $1 \leq n \leq 2^q$, $1 \leq m \leq 2^p$, either the supports $[\frac{n-1}{2^q}, \frac{n}{2^q}]$ and $[\frac{m-1}{2^p}, \frac{m}{2^p}]$ are disjoint. Or, if $p > q$, $[\frac{m-1}{2^p}, \frac{m}{2^p}] \subset [\frac{n-1}{2^q}, \frac{n}{2^q}]$ and then $h_{qn}|_{[\frac{m-1}{2^p}, \frac{m}{2^p}]} = c$ is constant on the support of h_{pm}, and hence

$$\langle h_{pm}, h_{qn}\rangle = c\int_{\frac{m-1}{2^p}}^{\frac{m}{2^p}} h_{pm}(x)dx = 0.$$

The system is also complete: linear hull of set of indicator functions of dyadic intervals of the form $[\frac{m-1}{2^p}, \frac{m}{2^p}]$ dense in $L^2([0,1])$ (measure theory, definition of integral through integrals of step functions). Moreover $1_{[\frac{m-1}{2^p}, \frac{m}{2^p}]}$ is in the linear hull of

$(h_{pm})_{p\geq 0, 1\leq m\leq 2^p}$. To see this, we use induction, starting by

$$\begin{aligned} 1_{[0,1]} &= h_{00}, \\ 1_{[0,\frac{1}{2}]} &= \frac{1}{2}(h_{00}+h_{01}), \\ 1_{[\frac{1}{2},1]} &= \frac{1}{2}(h_{00}-h_{01}). \end{aligned}$$

2. **Construction of sequence of i. i. d. Gaussian unit variables:**
For $t\in[0,1]$

$$\begin{aligned} 1_{[0,t]} &= \langle 1_{[0,t]}, h_{00}\rangle\, h_{00} + \sum_{p\geq 0, 1\leq m\leq 2^p} \langle 1_{[0,t]}, h_{pm}\rangle\ h_{pm} \\ &= \int_0^t h_{00}(s)ds\, h_{00} + \sum_{p\geq 0, 1\leq m\leq 2^p} \int_0^t h_{pm}(s)ds h_{pm} \\ &= g_{00}(t)h_{00} + \sum_{p\geq 0, 1\leq m\leq 2^p} g_{pm}(t)\ h_{pm}, \end{aligned}$$

with the *Schauder functions* $(g_{00}, g_{pm})_{p\geq 0, 1\leq m\leq 2^p})$. Let

$$t^0_{pm} := \frac{m-1}{2^p}, \quad t^1_{pm} := \frac{2m-1}{2^{p+1}}, \quad t^2_{pm} := \frac{m}{2^p}, \quad p\geq 0, 1\leq m\leq 2^p,$$

and $t^0_{00}=0, \quad t^1_{00}:=1, \quad t^2_{00}:=1$. Define for f continuous on $[0,1]$, $0\leq s\leq t\leq 1$

$$\begin{aligned} \langle 1_{[s,t]}, df\rangle &:= f(t)-f(s), \quad \text{and so} \\ \langle h_{pm}, df\rangle &:= 2^{\frac{p}{2}}\left[\left(f\left(t^1_{pm}\right)-f\left(t^0_{pm}\right)\right)-\left(f\left(t^2_{pm}\right)-f\left(t^1_{pm}\right)\right)\right] \\ &= 2^{\frac{p}{2}}\left[2f\left(t^1_{pm}\right)-f\left(t^0_{pm}\right)-f\left(t^2_{pm}\right)\right], \quad p\geq 0, 1\leq m\leq 2^p, \\ \langle h_{00}, df\rangle &:= f(1)-f(0). \end{aligned}$$

Now by 1. we can write W sample by sample as a series with coefficients $\langle h_{00}, dW\rangle, \langle h_{pm}, dW\rangle, p\geq 0, 1\leq m\leq 2^p$:

$$W(t) = \langle 1_{[0,t]}, dW\rangle = g_{00}(t)\langle h_{00}, dW\rangle + \sum_{p\geq 0, 1\leq m\leq 2^p} g_{pm}(t)\langle h_{pm}, dW\rangle.$$

Show: $\langle h_{00}, dW\rangle$, $\langle h_{pm}, dW\rangle, p\geq 0, 1\leq m\leq 2^p$ is an i.i.d. sequence of standard normal variables.

Proof:
a) First of all, note that since increments of Brownian motion are centered, we have

$$E(\langle h_{pm}, dW\rangle) = 0, \quad p\geq 0, 1\leq m\leq 2^p.$$

b) To calculate covariances, let $(p,m),(q,n)$ be given with $p,q\in\mathbf{N}, 1\leq m\leq 2^p, 1\leq n\leq 2^q$.

Case 1: $p = q, n = m$

In this case by independence of increments

$$\begin{aligned} E(\langle h_{pm}, dW\rangle^2) &= 2^p[E([(W(t^1_{pm}) - W(t^0_{pm}) - (W(t^2_{pm}) - W(t^1_{pm}))]^2)] \\ &= 2^p[E((W(t^1_{pm}) - W(t^0_{pm}))^2) + E((W(t^2_{pm}) - W(t^1_{pm})))^2) \\ &= 2^p[2^{-(p+1)} + 2^{-(p+1)}] = 1. \end{aligned}$$

Case 2: $p = q, m < n$

Here as a direct consequence of independence of increments

$$E(\langle h_{pm}, dW\rangle\langle h_{pn}, dW\rangle) = 0.$$

Case 3: $p < q, [\frac{m-1}{2^p}, \frac{m}{2^p}] \cap [\frac{n-1}{2^q}, \frac{n}{2^q}] = \emptyset$

In this case, as in the preceding one by independence of increments

$$E(\langle h_{pm}, dW\rangle\langle h_{qn}, dW\rangle) = 0.$$

Case 4: $p < q, [\frac{m-1}{2^p}, \frac{m}{2^p}] \supset [\frac{n-1}{2^q}, \frac{n}{2^q}]$

Here even w.l.o.g. $[\frac{m-1}{2^p}, \frac{2m-1}{2^{p+1}}] \supset [\frac{n-1}{2^q}, \frac{n}{2^q}]$. Hence

$$\begin{aligned} & E(\langle h_{pm}, dW\rangle\langle h_{qn}, dW\rangle) \\ & \quad = 2^{\frac{p+q}{2}}[E((W(t^1_{pm}) - W(t^0_{pm}))(W(t^1_{qn}) - W(t^0_{qn}))) \\ & \quad -E((W(t^1_{pm}) - W(t^0_{pm}))(W(t^2_{qn}) - W(t^1_{qn}))) \\ & \quad -E((W(t^2_{pm}) - W(t^1_{pm}))(W(t^1_{qn}) - W(t^0_{qn}))) \\ & \quad +E((W(t^2_{pm}) - W(t^1_{pm}))(W(t^2_{qn}) - W(t^1_{qn}))) \\ & \quad = E((W(t^1_{qn}) - W(t^0_{qn}))^2) \\ & \qquad -E((W(t^2_{qn}) - W(t^1_{qn}))^2) \quad \text{(independence of increments)} \\ = \ & 0 \quad \text{(equal length of intervals).} \end{aligned}$$

This proves the claim.□

3. Let If $W_p(t) = \langle h_{00}, dW\rangle g_{00}(t) + \sum_{k=0}^{p}\sum_{1\le m\le 2^k} g_{km}(t)\langle h_{pm}, dW\rangle,\ p \in \mathbf{N}$.
Show: For $t \in [0,1]$ $E((W(t) - W_p(t))^2) \to 0$ as $p \to \infty$.

(Absolute and uniform convergence pathwise: see course manuscript of 2019.)

In fact, for $t \in [0,1]$, $p \ge 1$ by the law properties of $W(t)$ and $\langle h_{pm}, dW\rangle, q \ge 1, 1 \le m \le 2^q$, proved before, and since $W_p(t)$ is the projection of $W(t)$ on span$(\langle h_{qm}, dW\rangle : q \le p, 1 \le m \le 2^q)$, we have

$$\begin{aligned} E((W(t) - W_p(t))^2) &= t + E(W_p(t)^2) - 2E(W(t)W_p(t)) \\ &= t - E(W_p(t)^2) = t - g_{00}^2(t) - \sum_{n=0}^{p}\sum_{1\le m\le 2^n} g_{nm}(t)^2, \end{aligned}$$

which converges to 0 as $p \to \infty$ by Parseval's identity.

4. (Modulo completion) we have

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

Let $t \in [0,1]$. Then by the representation of 2. $W(t)$ (modulo completion) measurable w. r. t. $(\langle h_{00}, dW\rangle, \langle h_{pm}, dW\rangle)_{p\geq 0, 1\leq m\leq 2^p}$. Therefore (modulo completion)

$$\mathcal{F} = \sigma((\langle h_{00}, dW\rangle, \langle h_{pm}, dW\rangle)_{p\geq 0, 1\leq m\leq 2^p}).$$

•

Goal: construct abstract isomorphism between *canonical Wiener space* and *sequence space*. Main interest: infinite dimensional spaces. Therefore

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

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

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

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

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

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

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

Notation consistent for $n = 1$.

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

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

is $\mathcal{F}$-measurable on Ω.

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

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

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

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

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

2. By part 1. and linearity: claimed equation verified for linear combinations of indicator functions of sets in $\mathcal{F}$. Equation stable for monotone limits. Hence: valid for $\mathcal{F}$-measurable functions (monotone class theorem). •

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

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

defines a linear isomorphism.

Proof
The mapping is well defined due to

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

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

Now: develop differential calculus on sequence space $(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu)$. Then use Theorem 1.1 to transfer it to canonical space $(\Omega, \mathcal{F}, P)$. Start with one-dimensional situation.

Focus: questions of smoothness of probability measures. Start with measures on $\mathbf{R}$.

2 Dimension 1: absolute continuity and duality

Goal: study laws of random variables on $(\Omega, \mathcal{F}, P)$, i.e. measures $P_F = P \circ F^{-1}$ for random variables F. By Theorem 1.1 $F = f \circ ((X_n)_{n\in\mathbf{N}})$, and for $A \in \mathbf{B}^1$

$$P_F(A) = P(F \in A) = P(f((X_n)_{n\in\mathbf{N}}) \in A) = \nu(f \in A),$$

so study $\nu \circ f^{-1}$ for $\mathbf{B}^{\mathbf{N}}$-measurable f on $\mathbf{R}^{\mathbf{N}}$.

One-dimensional version: $\nu_1 \circ f^{-1}$ for $\mathbf{B}^1$-measurable $f : \mathbf{R} \to \mathbf{R}$.

The following is a simple analytic criterion for absolute continuity.

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

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

Then $\mu \ll \lambda$, the Lebesgue measure on $\mathbf{R}$.

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

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

Then

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

This extends to bounded measurable f (measure theoretic standard argument). Therefore for $A \in \mathbf{B}^1$

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

which implies $\mu \ll \lambda$. •

Aim: apply Lemma 2.1 to $\nu_1 \circ f^{-1}$ with f $\mathbf{B}^1$-measurable. Here *integration by parts on Gaussian spaces* appears, which is central for Malliavin's calculus.

2.1 Classical integration by parts

Indeed, for $g, h \in C_0^\infty(\mathbf{R})$

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

This can be extended to $g, h \in L^2(\mathbf{R})$ vanishing at $\pm\infty$ with derivatives in the distributional sense. Assume this setting, and denote dg the *derivative in distributional sense* of g, (*) with δh its *adjoint operator* in the sense of (*). For $h \in C_0^\infty(\mathbf{R})$

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

and (*) reads

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

Finally, $\delta d = -\frac{d^2}{dx^2}$, here the *negative Laplace operator*, plays an important role. In classical calculus: no need to distinguish d and δ (modulo sign).

2.2 Gaussian integration by parts

For $g, h \in L^2(\mathbf{R}, \nu_1)$ let

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

To apply Lemma 1.1 to $\mu = \nu_1 \circ f^{-1}$ for f $\mathbf{B}^1$-measurable, we write (assuming operations justified)

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

Aim: transfer derivation to other argument. In fact, for $g, h \in C_0^\infty(\mathbf{R})$

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

Now define basic operators: *dg derivative in distributional sense*, its dual operator for (**) (domain to be described later)

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

Then duality relationship reads

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

Combination of derivative and dual operator:

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

in suitable sense.

d: *Malliavin derivative*, δ: *Skorokhod integral*, L: *Ornstein-Uhlenbeck operator.* Domains: precisely defined in higher dimensional setting.

Back to smoothness of $\nu_1 \circ f^{-1}$:

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

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

Moreover for $f \in L^2(\mathbf{R}, \nu_1)$ with $\delta(\frac{1}{df}) \in L^2(\mathbf{R}, \nu_1)$:

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

Proof
By duality and Cauchy-Schwarz

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

Apply Lemma 1.1 with $c = ||\delta(\frac{1}{df})||_2$. •

Goal: develop calculus of $d, \delta, (L)$ in finite dimension. Useful: eigenbasis of (1-dim.) L in $L^2(\mathbf{R}, \nu_1)$.

3 Hermite polynomials; orthogonal developments

d, δ and L well defined on $C_0^\infty(\mathbf{R})$, extend to space of polynomials in one real variable (integrability properties of ν_1).

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

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

H_n*:* Hermite polynomial *of degree n.*

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

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

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

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

H_n is eigenvector of L with eigenvalue n.

Proof

(i): follows by definition.

(ii): Study commutator of d and δ : for $h \in C_0^\infty(\mathbf{R})$

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

Hence

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

Induction on n: claim clear for $n = 1$; assume it holds for $n - 1$. Then

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

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

Hermite polynomials provide CONS of Gaussian space in one dimension.

Theorem 3.2. $(\frac{1}{\sqrt{n!}} H_n)_{n\geq 0}$ *CONS in $L^2(\mathbf{R}, \nu_1)$.*

Proof

1. Let $n, k \in \mathbf{N}$, s. th. $n < k$. By Theorem 3.1

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

while

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

2. **show:** $(H_n)_{n\geq 0}$ complete in $L^2(\mathbf{R}, \nu_1)$, i.e. the set of linear combinations of Hermite polynomials is dense in $L^2(\mathbf{R}, \nu_1)$.

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

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

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

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

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

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

Therefore $F = 0$, and by uniqueness of Fourier transforms $\phi = 0$. •

Goal: suitable notions of operators d and δ on sequence space $\mathbf{R}^{\mathbf{N}}$. Start with describing CONS of $L^2(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu)$. First: extend Hermite polynomials.

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

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

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

Next extend Theorem 3.2 to multidimensional setting.

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

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

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

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

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

Hence $(\frac{1}{\sqrt{p!}} H_p)_{p\in E_k}$ is orthonormal system. Moreover, linear combinations of tensor products of functions of one of k variables are dense in $L^2(\mathbf{R}^k, \mathbf{B}^k, \nu_k)$ (approximation theorem by Stone-Weierstrass).

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

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

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

Lemma 3.1. *Let $p \ge 1$, $f \in L^p(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu)$. Then*

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

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

Proof
Standard result of discrete martingale theory. •

Now $\hat{f}_n = f_n \circ \pi_n$ with $f_n \in L^2(\mathbf{R}^n, \mathbf{B}^n, \nu_n)$. Hence, first assertion implies second one. •

Next: study Sobolev spaces in finite and infinite dimension; use knowledge of CONS.

4 Finite dimensional Gaussian Sobolev spaces

Let $k \in \mathbf{N}$. Consider k-dimensional spaces first.

4.1 Classical Sobolev spaces on $\mathbf{R}^k$

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

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

as $\varepsilon \to 0$. Let

$$\begin{aligned} W_1^p &= \{f \in L^p(\mathbf{R}^k) : f \textit{ possesses directional derivative in direction} \\ &\quad a \textit{ for any } a \in \mathbf{R}^k\} \end{aligned}$$

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

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

Definition 4.2. *Let $p \ge 1, s \in \mathbf{N}$. Define recursively*

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

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

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

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

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

and analogous norms for higher derivative orders.

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

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

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

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

We finally need a local version of Sobolev spaces.

Definition 4.3. *For $p \geq 1, s \in \mathbf{N}$ let*

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

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

Aim: Find relationship of Sobolev spaces and "weak derivatives" or "derivatives in distributional sense".

Definition 4.4. *Let $f \in L_{loc}^1(\mathbf{R}^k), a \in \mathbf{R}^k$. Then $u_a \in L_{loc}^1(\mathbf{R}^k)$ is called* weak derivative *of f* in direction *a, if for $\phi \in C_0^\infty(\mathbf{R}^k)$ the equation*

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

is satisfied.

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

Proof
1. **Show:(i)⇒(ii)**
Assume $a \in \mathbf{R}^k, \phi \in C_0^\infty(\mathbf{R}^k)$. Then for $\varepsilon > 0$ by translational invariance of Lebesgue measure

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

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

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

So f possesses weak derivative $d_a f \in L^p(\mathbf{R}^k)$.

2. **Show: (ii)⇒(i)**

Fix $\phi \in C_0^\infty(\mathbf{R}^k), a = (a_1, \cdots, a_k) \in \mathbf{R}^k$. By Taylor's formula with integral remainder term and Fubini's theorem for any $\varepsilon > 0$

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

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

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

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

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

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

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

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

Proof

To show: f weakly differentiable in direction e_i for $1 \le i \le k$, and $d_i f = \lim_{n\to\infty} d_i f_n \in L^p(\mathbf{R}^k)$. So let

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

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

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

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

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

Proof

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

4.2 Gaussian Sobolev spaces in $\mathbf{R}^k$

Now to construct: Sobolev spaces w.r.t. ν_k (finite) instead of λ^k (infinite). According to Theorem 4.1, need *weak derivative, integrability w.r.t. ν_k*; *local Sobolev spaces* important.

For $a \in \mathbf{R}^k, f \in W_{1,loc}^p$ get *weak derivative*: for $\phi \in C_0^\infty(\mathbf{R}^k)$ choose $\psi \in C_0^\infty(\mathbf{R}^k)$ s. th. $\psi = 1$ on a neighborhood of $\operatorname{supp}\phi$. Then

$$\langle f, d_a\phi\rangle = \langle \psi f, d_a\phi\rangle = -\langle d_a(\psi f), \phi\rangle = -\langle u_a, \phi\rangle,$$

with u_a independent on ψ (product rule for ψg with smooth approximation g of f). $d_a f := u_a$ uniquely defined.

For $s \in \mathbf{N}, p \ge 1, 1 \le j_1, \cdots, j_s \le k, f \in W_{s,loc}^p$, with $d_j = d_{e_j}$ for canonical basis $e_1, \cdots, e_k$, define

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

by iteration

So to define Gaussian Sobolev spaces, just have to have integrability of weak derivatives.

Definition 4.5. *Let $s \in \mathbf{N}, p \ge 1$, $1 \le j \le k, f \in W_{s,loc}^p$. Let then*

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

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

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

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

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

This yields notion of Gaussian Sobolev spaces:

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

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

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

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

Remark
$D_s^p(\mathbf{R}^k)$ Banach space. Use arguments as for proving Corollary 4.2.

In the mostly studied Hilbert case $p = 2$, the CONS composed of k-dimensional Hermite polynomials plays central role. **Goal:** compute operators in terms of series expansions w.r.t. this CONS.

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

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

with $c_p(f) \in \mathbf{R}, p \in E_k$. By orthogonality,

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

We write $f \sim (c_p(f))$ in this case. Denote $\mathcal{P}$: linear hull of k-dimensional Hermite polynomials. Plainly, $\mathcal{P} \subset W_{s,loc}^p$ for any $s \in \mathbf{N}, p \geq 1$. By Chapter 3, $\mathcal{P}$ dense in $L^2(\mathbf{R}^k, \nu_k)$. And for functions in $\mathcal{P}$, the generalized derivatives d_j are identical to the usual partial derivatives in direction j, $1 \leq j \leq k$.

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

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

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

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

According to Corollary 4.2 and the calculations above, the domains of the operators extending ∇, δ_j and L beyond $\mathcal{P}$ must be those distributions in $\mathbf{R}^k$ for which formulas just given generate convergent series in the L^2-norm with respect to ν_k. Most important domain: the one of ∇, the Sobolev space $D_1^2(\mathbf{R}^k)$. For $f \sim (c_p(f)) \in \mathcal{P}$ we have

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

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

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

Goal: describe Gaussian Sobolev spaces, domains of principal operators for $p = 2$, using Hermite expansions. Start with $k = 1$.

Theorem 4.2. *Let $r \in \mathbf{N}$, $f \sim (c_p(f)) \in L^2(\mathbf{R}, \nu_1) \cap W^2_{r,loc}$. Denote $f_p = \frac{c_p(f)}{p!} H_p, p \geq 0$. The following are equivalent:*
(i) $\nabla^r f \in L^2(\mathbf{R}, \nu_1)$,
(ii) $\sum_{p\geq 0} p^r ||f_p||_2^2 < \infty$,
(iii) $f \in D_r^2(\mathbf{R})$,
(iv) $\delta^r f \in L^2(\mathbf{R}, \nu_1)$.
In particular, $D_r^2(\mathbf{R})$ is domain of ∇^r, δ^r in $L^2(\mathbf{R}, \nu_1)$. For $f, g \in D_1^2(\mathbf{R})$ we have

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

Proof
1. **Show:** (i) $\Leftrightarrow$ (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 so by iteration

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

Therefore

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

and hence

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

if and only if

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

and this is the case iff

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

2. **Show:** (ii)$\Leftrightarrow$(iv).

Note

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

Hence

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

iff

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

3. (i)$\Leftrightarrow$(iii): by definition.

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

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

whereas

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

•

Differential calculus on Gaussian spaces obeys similar rules as classical differential calculus.

Theorem 4.3. *Let $g \in D_1^4(\mathbf{R}), \mu = \nu_1 \circ g^{-1}$. If $\phi \in L^4(\mathbf{R}, \mu)$, and $\nabla\phi \in L^4(\mathbf{R}, \mu)$, we have*

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

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

Theorem 4.4. *Let $f, g \in D_1^4(\mathbf{R})$. Then $f \cdot g \in D_1^2(\mathbf{R})$ and*

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

Proof
The assertion is clear for $f, g \in \mathcal{P}$. To generalize, approximate in $\mathcal{P}$, use Hölder's inequality. •

Goal: extend to finite dimension k.

Theorem 4.5. *Let $f \sim (c_p(f)) \in L^2(\mathbf{R}^k, \nu_k) \cap W_{1,loc}^2$. Denote $f_p = \frac{c_p(f)}{p!} H_p, p \in E_k$. The following are equivalent:*
(i) $|\nabla f| = [\sum_{j=1}^k (d_j f)^2]^{\frac{1}{2}} \in L^2(\mathbf{R}^k, \nu_k)$,
(ii) $\sum_{p \in E_k} |p| ||f_p||_2^2 < \infty$,
(iii) $f \in D_1^2(\mathbf{R}^k)$.
In particular, $D_1^2(\mathbf{R}^k)$ is domain of ∇ in $L^2(\mathbf{R}^k, \nu_k)$. Analogous results hold for Sobolev spaces of order $(r, 2)$ with $r \in \mathbf{N}$.

Proof
Analogous to proof of Theorem 4.2.•

5 Infinite dimensional Gaussian Sobolev spaces

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

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

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

Let for $n \in \mathbf{N}$ let $\mathbf{C}_n = \sigma(\pi^n) = \sigma(\pi^1, \cdots, \pi^n)$. Then $(\mathbf{C}_n)_{n \in \mathbf{N}}$ filtration on $(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}})$. For $f \in L^p(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu)$ by Lemma 3.1

$$\hat{f}_n = E(f|\mathbf{C}_n) = \int f(\cdot, (x_i)_{i>n}) \prod_{i>n} \nu_1(d(x_i)_{i>n}), \quad n \in \mathbf{N}$$

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

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

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

We show that the martingale property is not destroyed by directional derivative operators.

Lemma 5.1. *Let $p > 1$, $f \in L^p(\mathbf{R}^{\mathbf{N}})$, $(f_n)_{n\in\mathbf{N}}$ sequence according to above remarks. Suppose that $\sup_{n\in\mathbf{N}} ||f_n||_{1,p} < \infty$. Then for $j \in \mathbf{N}$ the sequence $(d_j f_n \circ \pi^n)_{n\in\mathbf{N}}$ converges in $L^p(\mathbf{R}^{\mathbf{N}})$, to a limit denoted by $d_j f$. Corresponding statements hold for higher derivatives.*

Proof
Let $n \in \mathbf{N}, j \in \mathbf{N}$. Then by integrability properties we have

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

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

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

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

So we may define Sobolev spaces in infinite dimensions.

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

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

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

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

This definition makes sense:

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

Proof
We argue for $s = 1$. Let $(f^m)_{m\in\mathbf{N}}$ Cauchy sequence in $D^p_1(\mathbf{R}^{\mathbf{N}})$, and $(f^m_n)_{n,m\in\mathbf{N}}$ finite dimensional functions, according to remarks above. Then for $m, l \in \mathbf{N}$, $n \in \mathbf{N}$ Jensen's inequality gives

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

[Jensen's inequality: f r.v. on $\mathbf{R}^{\mathbf{N}}$, $\phi : \mathbf{R} \to \mathbf{R}$ convex, then $E(\phi(f)|\mathbf{C}_n) \geq \phi(E(f|\mathbf{C}_n))$; hence $E(\phi(f)) \geq E(\phi(\hat{f}_n)$; apply to $\mathbf{R} \ni x \mapsto |x|^p$ convex for $p \geq 1$.]

$D^p_1(\mathbf{R}^n)$ being a Banach space for $n \in \mathbf{N}$, we know that

$$f_n = \lim_{m\to\infty} f^m_n \in D^p_1(\mathbf{R}^n)$$

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

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

Moreover

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

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

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

as $m \to \infty$. •

According to Lemma 5.1, the gradient on infinite dimensional Gaussian Sobolev spaces is defined as follows.

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

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

*(*Malliavin gradient *or* Malliavin derivative*), where for any $j \in \mathbf{N}$ according to Lemma 5.1*

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

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

Remark

∇ being a continuous mapping from $D_1^p(\mathbf{R}^n)$ to $L^p(\mathbf{R}^n, \nu_n)$ for any n, Lemma 5.1 and the definition of the Malliavin gradient imply, that $\nabla : D_1^p(\mathbf{R}^{\mathbf{N}}) \to L^p(\mathbf{R}^{\mathbf{N}}, \nu)$ is continuous.

Goal: For $p = 2$ describe Gaussian Sobolev spaces, using generalized Hermite polynomials.

First of all, suppose $f = \sum_{p\in E} \frac{c_p(f)}{p!} H_p \in L^2(\mathbf{R}^{\mathbf{N}}, \nu)$. We use the notation $f \sim (c_p(f))$. Then for $n \in \mathbf{N}$, we have $\hat{f}_n = \sum_{p\in E} \frac{c_{(p,0)}(f)}{(p,0)!} H_{(p,0)}$, since for $p_i \geq 1$ $\langle H_{p_i}|1\rangle = 0$, where we put $(p, 0) = (p_1, \cdots, p_n, 0, 0, \cdots)$ for $p = (p_1, \cdots, p_n) \in E_n$. Therefore $f_n = \sum_{p\in E_n} \frac{c_{(p,0)}(f)}{p!} H_{(p,0)}$. Let again $\mathcal{P}$: linear hull of generalized Hermite polynomials.

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

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

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

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

For infinite dimensional Gaussian Sobolev spaces of order $(1, 2)$ we get:

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

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

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

6 The canonical Wiener space: multiple integrals

Return to path space $(\Omega, \mathcal{F}, P)$: $\Omega = C(\mathbf{R}_+, \mathbf{R})$, $\mathcal{F}$ (completed) Borel σ-algebra on Ω generated by topology of uniform convergence on compacts in $\mathbf{R}_+$, P Wiener measure on $\mathcal{F}$. Transfer between sequence and path space: isomorphism of Chapter 1. Recall: $(h_n)_{n \in \mathbf{N}}$ enumeration of Haar functions on $\mathbf{R}_+$, where CONS on $[0, 1]$ extended to intervals $[k, k+1], k \in \mathbf{N}$, by translation; then isomorphism

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

Most important tool in differential calculus on Gaussian sequence space: *Hermite expansions* in $L^2(\mathbf{R}^{\mathbf{N}}, \nu)$. On path space, they are interpreted by *multiple Itô integrals.*

By isomorphism, objects corresponding to generalized Hermite polynomials on canonical space are

$$\prod_{i=1}^{\infty} H_{p_i}(\langle h_i, dW \rangle),$$

$p \in E$. **Goal:** description as *multiple Itô integrals.* Use abbreviation $\mathbf{B}_+^1$ for Borel sets of $\mathbf{R}_+$.

Definition 6.1. *For $m \in \mathbf{N}$ let*

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

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

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

let

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

if $J_i =]s_i, t_i], 1 \le i \le n$. By law properties of W

$$\begin{aligned} ||W(f)||_2^2 &= E(\sum_{1\le i,j\le n} a_i a_j (W(t_i) - W(s_i))(W(t_j) - W(s_j)) \\ &= \sum_{1\le i\le n} a_i^2 E((W(t_i) - W(s_i))^2) = \sum_{1\le i\le n} a_i^2 (t_i - s_i) = ||f||_2^2. \end{aligned}$$

Since elementary functions of this form are dense in $L^2(\mathbf{R}_+)$, linear map $f \mapsto W(f)$ extends to $L^2(\mathbf{R}_+)$. In fact, if $(f_n)_{n\in\mathbf{N}}$ Cauchy sequence in $L^2(\mathbf{R}_+)$, then $(W(f_n))_{n\in\mathbf{N}}$ Cauchy sequence in Banach space $L^2(\Omega, \mathcal{F}, P)$. Therefore, for $A \in \mathbf{B}_+^1$ with finite Lebesgue measure, $W(A) = W(1_A)$ is defined. It will be used for definition of multiple stochastic integrals. In this chapter, $\langle \cdot, \cdot \rangle$ will be w. r. t. Lebesgue measure on $\mathbf{R}_+^m$ with unspecified integer m.

Definition 6.2. *For $f = \sum_{i_1,\cdots,i_m=1}^n a_{i_1\cdots i_m} 1_{A_{i_1}\times\cdots\times A_{i_m}} \in \mathcal{E}_m$ let*

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

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

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

Lemma 6.1. *Let $m, q \in \mathbf{N}, f \in \mathcal{E}_m, g \in \mathcal{E}_q$.*
(i) $I_m|\mathcal{E}_m$ linear,
(ii) if $\tilde{f}(t_1, \cdots, t_m) = \frac{1}{m!}\sum_{\sigma\in\mathcal{S}_m} f(t_{\sigma(1)}, \cdots, t_{\sigma(m)})$ (symmetrization of f), then

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

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

Proof

1. (i) follows from additivity of $A \mapsto W(A)$.
2. (ii): in definition of I_m product $W(A_{i_1})\cdots W(A_{i_m})$ invariant under permutations of factors.
3. By (ii), we may assume f, g symmetric. By choosing common subdivisions, we may assume

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

a) $m \ne q$**:** $(A_i)_{1\le i\le n}$ consists of p.d. Borel sets; consider $(j_1, \cdots, j_m), (i_1, \cdots, i_q)$ s. th. $j_k \ne j_l, i_k \ne i_l$ for $k \ne l$; for $q < m$ there is j_l s. th. $j_l \ne i_1, \cdots, i_q$, hence $E(\prod_{k=1}^m W(A_{i_k}) \prod_{l=1}^q W(A_{j_l})) = 0$.

b) $m = q$**:** again by assumptions and symmetry

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

•

To extend I_m beyond $\mathcal{E}_m$, proceed as for $m = 1$.

Lemma 6.2. *$\mathcal{E}_m$ dense in $L^2(\mathbf{R}_+^m)$ for $m \in \mathbf{N}$.*

Proof

Assume $m \geq 2$ (claim clear for $m = 1$.) By standard results of measure theory, **to show:** for $A_1, \cdots, A_m \in \mathbf{B}_+^1$ with finite Lebesgue measure, and $\varepsilon > 0$, there is $f \in \mathcal{E}_m$ s. th.

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

Let $\delta > 0$ (determined later). Choose $B_1, \cdots, B_n \in \mathbf{B}_+^1$ with finite Lebesgue measure, pairwise disjoint, and s. th. for $1 \leq j \leq n$ we have $\lambda(B_j) < \delta$, and s. th. A_i is a finite union of some B_j. Then

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

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

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

and

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

Finally, choose δ small enough. •

Use Lemma 6.2 to extend I_m to $L^2(\mathbf{R}_+^m)$.

Definition 6.3. *The linear and continuous extension of $I_m|\mathcal{E}_m$ to $L^2(\mathbf{R}_+^m)$ according to Lemma 6.2 is called* multiple Wiener-Itô integral *of degree m and also denoted by I_m.*

Properties of elementary integral transfer straightforwardly.

Theorem 6.1. *Let $m, q \in \mathbf{N}, f \in L^2(\mathbf{R}_+^m), g \in L^2(\mathbf{R}_+^q)$. Then*
(i) $I_m|L^2(\mathbf{R}_+^m)$ is linear,
(ii) we have

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

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

Notation
We write

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

Goal: explain relationship between *generalized Hermite polynomials* and *multiple Wiener-Itô integrals*. For this purpose: need recursive relationship between Hermite polynomials of different degrees.

Remark
Recall definition of Hermite polynomials:

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

We may compute for $n \in \mathbf{N}$

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

We need the following *operation of contraction.*

Definition 6.4. *Let $m \in \mathbf{N}$. Suppose $f \in L^2(\mathbf{R}_+^m), g \in L^2(\mathbf{R}_+)$. Then for $t_1, \cdots, t_m, t \in \mathbf{R}_+$*

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

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

Recursion relation for Hermite polynomials will follow from recursion relation between Wiener-Itô integrals.

Lemma 6.3. *Let $m \in \mathbf{N}, f \in L^2(\mathbf{R}_+^m), g \in L^2(\mathbf{R}_+)$. Then*

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

Proof

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

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

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

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

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

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

a) **Show:** first term on r. h. s. close to $I_{m+1}(f \otimes g)$. In fact, let

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

Then

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

b) **Show:** second term on r. h. s. negligeable as $\varepsilon \to 0$. In fact, denote it by R_ε. 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 last term, note

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

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. •

Here is the relationship between Hermite polynomials and multiple Itô integrals.

Theorem 6.2. *Let $m \in \mathbf{N}$, $h \in L^2(\mathbf{R}_+)$ be s. th. $||h||_2 = 1$. Denote $h^{\otimes m}$ the m-fold tensor product of h with itself. Then*

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

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

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

In particular, for $F \in L^2(\Omega, \mathcal{F}, P)$, $m \geq 0$ there is $f_m \in L^2(\mathbf{R}_+^m)$, s. th.

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

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

Proof

1. **To prove:**

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

Induction on m: For $m = 1$, the formula is clear from $H_1 = x, I_1(h) = W(h)$. Assume it holds for m. Then Lemma 6.3 and the recursion formula for Hermite polynomials given above combine to yield, remembering that $||h||_2 = 1$,

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

2. Let $L_s^2(\mathbf{R}_+^m)$ be the linear space of symmetric functions in $L^2(\mathbf{R}_+^m)$. Then by Theorem 6.1

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

hence $\mathcal{H}_m = I_m(L_s^2(\mathbf{R}_+^m))$ closed. Orthogonality: consequence of Theorem 6.1.

3. Let $(g_i)_{i\in\mathbf{N}}$ be CONS of $L^2(\mathbf{R}_+)$, $F \in L^2(\Omega, \mathcal{F}, P)$. Let $f \in L^2(\mathbf{R}^{\mathbf{N}}, \mathbf{B}^{\mathbf{N}}, \nu)$ be s. th. $T(f) = F$. Assume $f \sim (c_p(f))$, according to the notation of Chapter 4. For $m \geq 0$, define

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

considered as function of m variables. Then $f_m \in L^2(\mathbf{R}_+^m)$ and by a slight generalization of Lemma 6.3 (contractions between $g_i^{\otimes p_i}$ and g_j vanish for $i \neq j$ by orthogonality)

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

Summing this expression over m yields the desired

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

The remaining claims are obvious.•

7 The canonical Wiener space: Malliavin's derivative

In this Chapter: investigate analogue of gradient in differential calculus on sequence space. Fix CONS $(h_i)_{i\in\mathbf{N}}$ of Haar functions in $L^2(\mathbf{R}_+)$, recall the isomorphism

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

where $W(h_i) = \langle h_i, dW\rangle, i \in \mathbf{N}$.

Pb: define Malliavin's derivative on canonical space in consistent and basis independent way?

For $f(W(h_i)) = f(\langle h_i, dW\rangle)$ with $f = \mathrm{id}$: plausible to get $h_i \in L^2(\mathbf{R}_+)$ as derivative. Also plausible: *chain rule*. So we set:

Definition 7.1. *For $n \in \mathbf{N}$ let $C_p^\infty(\mathbf{R}^n)$: set of smooth functions with partial derivatives of polynomial growth. Let*

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

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

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

Claim: good candidate for definition of Malliavin gradient on path space.

a): independence on representation of functionals

Let $h_1, \cdots, h_n \in L^2(\mathbf{R}_+)$, and $g_1, \cdots, g_m \in L^2(\mathbf{R}_+)$ orthonormal, s. th. linear hulls are identical, i. e. s. th. for $1 \leq i \leq n$

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

Then, denoting

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

for $f \in C_p^\infty(\mathbf{R}^n)$ we have $g = f \circ \Gamma \in C_p^\infty(\mathbf{R}^m)$ and

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

b): $T((\nabla f)^2)=\langle DF,DF\rangle$**:**
For $h\in L^2(\mathbf{R}_+)$ by definition

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

for $i\in\mathbf{N}$

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

So $\langle D.F,g_i\rangle$ can be seen as directional derivative in direction g_i, and by Parseval's identity

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

Analogously, higher derivatives are related. So transfer of notions of Gaussian Sobolev spaces from sequence to path space possible.

Definition 7.2. *Let $p\geq 2, s\in\mathbf{N}$. For $F\in L^2(\Omega,\mathcal{F},P)$ denote by $f\in L^2(\mathbf{R}^{\mathbf{N}},\mathbf{B}^{\mathbf{N}},\nu)$ the function for which we have $F=f((W(g_i)_{i\in\mathbf{N}})$. Then let*

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

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

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

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

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

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

Remark
D_s^p Banach space with norm $||\cdot||_{s,p}$ for $p \geq 2, s \in \mathbf{N}$. For $F \in D_s^p$

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

D_s^p is the closure of $\mathcal{S}$ w. r. t. the norm $||\cdot||_{s,p}$.

Goal: describe D_1^2 by Wiener-Itô integrals
D_1^2 Hilbert space w. r. t. scalar product

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

D is closed, and $D : D_1^2 \to L^2(\Omega, \mathcal{F}, P)$ continuous.

Theorem 7.1. *Let $F = \sum_{m=0}^{\infty} I_m(f_m) \in L^2(\Omega, \mathcal{F}, P)$, with f_m symmetric for $m \geq 0$. Then*

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

In this case

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

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

Proof
1. Suppose that w. r. t. a CONS $(g_i)_{i\in\mathbf{N}}$ of $L^2(\mathbf{R}_+)$ we have $F = T(f)$ with $f \sim (c_p(f))$. As before, for $m \geq 0$ let

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

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

2. Let now $p \in E$ such that $|p| = m$, and let $t \geq 0$. Then (Theorem 6.2)

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

Summation over p s. th. $|p| = m$ gives the formula

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

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

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

By closedness of D and remarks above,

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

Now by 2.

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

Let $n, m \in \mathbf{N}, n \geq m$. Then (Theorem 6.1)

$$E(\langle D(F_n - F_m), D(F_n - F_m)\rangle) = \sum_{k=m+1}^{n} k^2 \int_{\mathbf{R}_+} E(I_k(f_k(.,t))^2)dt$$

$$= \sum_{k=m+1}^{n} k^2 \int_{\mathbf{R}_+} (k-1)!\langle f_k(\cdot,t), f_k(\cdot,t)\rangle dt = \sum_{k=m+1}^{n} k^2 (k-1)!||f_k||_2^2.$$

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

Some rules for operating with the Malliavin gradient D:

Theorem 7.2. *Let $p \geq 2, d \in \mathbf{N}$, $\phi \in C^1(\mathbf{R}^d)$ with bounded partial derivatives, let $F = (F^1, \cdots, F^d) \in (D_1^p)^d$. Then $\phi \circ F \in D_1^p$ and*

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

Proof
Definition of D, denseness of $\mathcal{S}$, approximation, see proof of Theorem 4.3. •

The following leads to study of dual operator of D.

Theorem 7.3. *Let $F \in \mathcal{S}, h \in L^2(\mathbf{R}_+)$. Then*

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

Proof
We may assume $F = f(W(g_1), \cdots, W(g_n))$ and $h = g_0$ or $h = g_1$ w. r. t. ONS $g_0, g_1, \cdots, g_n \in L^2(\mathbf{R}_+)$.
Case 1: $h = g_0$. Then $F, W(h)$ independent, and by orthogonality

$$E(\langle DF, h\rangle) = 0 = E(F)E(W(h)) = E(FW(h)).$$

Case 2: $h = g_1$. By duality of d_1 and δ_1

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

•

Theorem 7.4. *Let $F, G \in \mathcal{S}, h \in L^2(\mathbf{R}_+)$. Then*

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

Proof

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

•

8 The canonical Wiener space: Skorokhod's integral

In this Chapter: study of dual operator of Malliavin gradient in path space. This operator turns out to be stochastic integral.

We know:

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

densely defined and linear.

Definition 8.1. *Let*

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

For $u \in \mathrm{dom}(\delta)$ the map $F \mapsto E(\langle DF, u\rangle)$ extends to continuous linear functional. By Riesz' representation we find $\delta(u) \in L^2(\Omega)$ s. th.

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

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

Definition 8.2. *For $u \in dom(\delta)$ the uniquely determined random variable $\delta(u) \in L^2(\Omega)$ is called* Skorokhod integral *of u.*

Notation
$\delta(u) = \int_{\mathbf{R}_+} u_t \delta W(t)$.

Pb: Why *integral*? First see how elementary processes 'integrate'.

Definition 8.3. *Let*

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

Lemma 8.1. *Let* $u = \sum_{i=1}^n F_i h_i \in \mathcal{S}_{L^2(\mathbf{R}_+)}$. *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 $u = Fh$ with $F \in \mathcal{S}$ and $h \in L^2(\mathbf{R}_+)$. Then for $G \in \mathcal{S}$ by Theorem 7.4

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

Hence

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

•

Recall standard Wiener filtration $(\mathcal{F}_t)_{t \geq 0}$; for $t \geq 0$ given by P-completion $\mathcal{F}_t$ of $\sigma(W(s) : s \leq t)$. Lemma 8.1 provides elementary Itô integral, if F_i is $\mathcal{F}_{t_i}$-measurable, $h_i = 1_{]t_i, t_{i+1}]}$, where $0 = t_0 < t_1 < \cdots < t_n$, if $\langle DF_i, h_i \rangle = 0, 1 \leq i \leq n-1$. This is the case:

Lemma 8.2. *Let* $F \in D_1^2, A \in \mathbf{B}_+^1, \mathcal{F}_A = \sigma(W(1_B) : B \subset A, \lambda(B) < \infty)$. *Then*

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

and

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

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

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

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

Let

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

Then by independence of $\mathcal{F}_A$ and the vector $(W(h_1 1_{A^c}), \cdots, W(h_n 1_{A^c}))$ (recall that for $A, B \in \mathbf{B}_+^1$ s. th. $A \cap B = \emptyset$ $W(A)$ and $W(B)$ are Gaussian and uncorrelated by Theorem 6.1, hence independent) we have

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

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

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

2. Now approximate $F \in D_1^2$ by standard arguments. •

Theorem 8.1. *Let $u \in L^2(\Omega \times \mathbf{R}_+)$ be $(\mathcal{F}_t)$-adapted. Then*

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

Proof

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

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

a) Let first $F \in \mathcal{S}$. Then by Lemma 8.1 and 8.2

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

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

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

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

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

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

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

b) Given u as in the claim, choose sequence $(u^n)_{n\in\mathbf{N}}$ of processes as in a) s. th. $||u^n - u||_2 \to 0$ in $L^2(\Omega \times \mathbf{R}_+)$. Use closedness of δ and definition of Itô integral to get

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

•

Goal: study operation of Skorokhod integral on Wiener-Itô expansion.

Lemma 8.3. *Let $u \in L^2(\Omega \times \mathbf{R}_+)$. Then for $m \geq 0$ there exist $f_m \in L^2(\mathbf{R}_+^{m+1})$ s. th. f_m is symmetric in its first m variables, and s. th.*

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

We have

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

Proof

Choose sequence of elementary $(u^n)_{n\in\mathbf{N}} \subset L^2(\Omega \times \mathbf{R}_+)$ s. th.

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

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

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

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

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

Then

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

Define

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

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

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

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

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

Hence also $u = \sum_{m=0}^\infty I_m(f_m(\cdot,\cdot))$ in $L^2(\Omega \times \mathbf{R}_+)$, and so $P \otimes \lambda$-a.e. •

Theorem 8.2. *Let $u \in L^2(\Omega \times \mathbf{R}_+)$, $u = \sum_{m=0}^\infty I_m(f_m(\cdot,\cdot))$ according to Lemma 8.3. Then*

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

In this case

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

Proof

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

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

By Theorem 7.1

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

2. **Show: "⇒"**

Let $u \in \mathrm{dom}(\delta), G = I_n(g) \in \mathcal{H}_n, n \in \mathbf{N}$. Then by 1. and duality

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

By extending to $G^n = \sum_{k=0}^n I_k(g_k)$ with $g_k \in L^2(\mathbf{R}_+^k)$ and linearity

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

3. **Show: "⇐"**

Assume $\sum_{n=0}^\infty (n+1)!||\tilde{f}_n||_2^2 < \infty$. Let $V = \sum_{n=0}^\infty I_{n+1}(f_n)$ (well defined), and $G^n = \sum_{k=0}^n I_k(g_k), n \in \mathbf{N}$, with $g_k \in L^2(\mathbf{R}_+^k), k \in \mathbf{N}$. Then by 1.

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

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

$\mathrm{dom}(\delta)$ analytically hardly accessible. For simpler treatment of calculus of Skorokhod's integral, following subspace more practical.

Definition 8.4. *Let*

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

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

Goal: description of $\mathbf{L}_1^2$ by Wiener-Itô decompositions.

Remark
Let $u \in L^2(\Omega \times \mathbf{R}_+)$, $u = \sum_{m=0}^{\infty} I_m(f_m(\cdot,\cdot))$ according to Lemma 8.3. In these terms state conditions of definition of $\mathbf{L}_1^2$.

For $t \geq 0$ by Theorem 7.1 $u_t \in D_1^2$ means

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

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

$$\sum_{m=0}^{\infty}\int_{\mathbf{R}_+}\int_{\mathbf{R}_+} m^2 E(I_{m-1}(f_m(\cdot,s,t))^2) ds dt = \sum_{m=0}^{\infty} m\, m! \int_{\mathbf{R}_+} ||f_m(\cdot,t)||_2^2 dt = \sum_{m=0}^{\infty} m\, m! ||f_m||_2^2 < \infty.$$

The latter holds iff

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

Compare this with condition obtained in Theorem 8.2. Since for $m \geq 0$ $||\tilde{f}_m||_2 \leq ||f_m||_2$, obviously

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

Transfer of Itô's isometry to Skorokhod integral:

Theorem 8.3. *Let $u, v \in \mathbf{L}_1^2$. Then*

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

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

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

Compare Theorem 3.1. By linearity, we may assume $u = Fh$, where $F \in \mathcal{S}, h \in L^2(\mathbf{R}_+)$. Then by Lemma 8.1 $\delta(u) = FW(h) - \langle DF, h\rangle$, and so

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

2. Let still $u \in \mathcal{S}_{L^2(\mathbf{R}_+)}$. Then duality, the calculation in 1. and $v \in \mathbf{L}_1^2$ lead to

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

3. Now approximate u by functions in $\mathcal{S}_{L^2(\mathbf{R}_+)}$, and use that convergence in $\mathbf{L}_1^2$ implies convergence for all three terms in formula. •

Goal: formula for Malliavin differentiation of Itô integrals. Needed in applications to backward stochastic differential equations (BSDE).

Theorem 8.4. *Let $u \in L^2(\Omega \times [0,1])$ be adapted, $X_t = \int_0^t u_s dW(s), 0 \le t \le 1$, its Itô integral process. Then*

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

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

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

and

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

Proof

For simplicity: $T = 1$. Use Wiener-Itô decompositions for approximations.

1. Use extension of representation of Lemma 8.3 to adapted u. Let $u = \sum_{m=0}^\infty I_m(f_m(\cdot,\cdot))$ by Lemma 8.3. Here for $m \ge 0$ the function $f_m \in L^2(\mathbf{R}_+^{m+1})$ is symmetric in its first m variables. **Show:**

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

(i):

In fact, resume notation of proof of Lemma 8.3, to specialize to adapted case. Approximate u in $L^2(\Omega \times [0,1])$ by

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

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

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

of F_k^n, since $I_m(f_m^{k,n}) = E(I_m(f_m^{k,n})|\mathcal{F}_{t_k^n})$, by Lemma 8.2 we have for $t_1, \cdots, t_m \in \mathbf{R}_+$

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

Hence

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

fulfill

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

For $m \geq 0$ recall (proof of Lemma 8.3) that $f_m^n \to f_m$ symmetric in its first m variables in $L^2(\mathbf{R}_+^{m+1})$, hence (via choosing a subsequence for a.e. convergence)

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

(ii):
Note that by validity of (i)

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

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

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

as claimed.

2. **Show:** $u \in \mathbf{L}_1^2$ iff $X_1 \in D_1^2$.
As shown in remark above, $u \in \mathbf{L}_1^2$ translates into

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

Moreover $X_1 = \delta(u) = \sum_{m=0}^{\infty} I_{m+1}(f_m)$, and by Theorem 7.1 $X_1 \in D_1^2$ iff

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

But by first part of 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.

3. **Show:** differentiation formula.
By Lemma 8.2

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

Moreover, for $t \leq s \leq 1$ according to Theorem 7.1

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

Since $Du \in L^2(\Omega \times [0,1]^2)$ by definition of $\mathbf{L}_1^2$, and since for fixed t the process $(D_t u_s)_{t \leq s \leq 1}$ is adapted, Theorem 8.1 implies $(D_t u_s)_{t \leq s \leq 1}$ is Itô integrable and Itô and Skorokhod integrals are identical. More formally,

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

Finally, $X_1 = \delta(u) \in D_1^2$. For $t \in [0,1]$ compute Malliavin derivative (in the sense of equality in $L^2(\Omega \times [0,1])$)

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

Last line identifies with $u_t + \int_t^1 D_t u_s dW(s)$, by using expansions of expressions given above. Norm equation follows from Itô's isometry. •

9 Utility maximization, martingale optimality and backward stochastic differential equations

Backward stochastic differential equations (BSDE): stochastic (control) tool for utility indifference based hedging on financial markets. In this Chapter: motivate this tool, look at typical optimal investment problem of small agent in stochastic finance, express *utility maximization* for optimal hedging by *martingale optimality*, derive important

class of BSDE to provide stochastic solution. We follow Hu, Imkeller, Müller [8]. Further references in this paper.

Work in setting of following financial market model: $(\Omega, \mathcal{F}, P)$ probability space with m-dim. Brownian motion $W = (W^1, \cdots, W^m)$, T finite time horizon; $(\mathcal{F}_t)_{t\geq 0}$ natural filtration.

asset price processes

trivial *bond* $S^0 \equiv 1$; *stocks* $1 \leq i \leq d$ following SDE

$$\begin{aligned} \frac{dS_t^i}{S_t^i} &= \sum_{j=1}^m \sigma_t^{ij} dW^j(t) + b_t^i dt \\ &= \sigma_t^i dW(t) + b_t^i dt \\ &= \sigma_t^i [dW(t) + \theta_t dt], \quad t \in [0,T] \end{aligned}$$

$$\theta = \sigma^*[\sigma\sigma^*]^{-1} b, \quad \underline{\epsilon} I^d \leq [\sigma\sigma^*] \leq \overline{\epsilon} I^d \quad \text{with} \quad \underline{\epsilon} > 0$$

θ: *price of risk*

Small trader seeks to optimize income; uses investment strategies under *non-convex constraints* described by set:

$\tilde{C} \subset \mathbb{R}^d$ closed (e. g. $\tilde{C} = \mathbb{Z}^d$); $\tilde{\mathcal{A}}$: strategies $\pi = (\pi^1, \cdots, \pi^d)$ s. th.
$\pi \in \tilde{C} \quad P \otimes \lambda$-a.s. ($\lambda$ Lebesgue measure)
$\{\exp(-\alpha \int_0^\tau \pi_s \frac{dS_s}{S_s}) : \tau$ stopping time in $[0,T]\}$ uniformly integrable (for technical reasons; see utility function)

wealth process

$$X_t^\pi = x + \sum_{i=1}^d \int_0^t \pi_s^i \frac{dS_s^i}{S_s^i} = x + \int_0^t \pi_s \sigma_s [dW(s) + \theta_s ds]$$

preferences of small agent measured by **exponential utility** from terminal wealth

utility function

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

Small agent subject to liabilities originating outside of market: *incomplete financial market*

F $\mathcal{F}_T$-measurable *liability* at time T.

Example 1:

Insurer sells *weather derivative,* price K, strike time T:

T_i minimum temperature on day i at specific location
heating degree days $HDD_i = \max(0, 18 - T_i)$
cumulative heating degree days $cHDD_t = \sum_{i=1}^{30} HDD_{t-i}$.

insurer has obligation to pay

$$F = (cHDD_{30} - K)^+ \quad \mathcal{F}_T - \text{measurable}$$

market incompleteness: S^i driven by Wiener process components $W^j, 1 \leq j \leq k < m$.

So get following optimization problem:

First formulation:

Find

$$V(x) = \sup_{\pi \in \tilde{\mathcal{A}}} E(U(X_T^\pi - F)) = \sup_{\pi \in \tilde{\mathcal{A}}} E(U(x + \int_0^T \pi_s \sigma_s [dW(s) + \theta_s ds] - F)).$$

For simplicity: define

$$\begin{array}{rcl} p & = & \pi\sigma, \\ C & = & \tilde{C}\sigma, \\ \mathcal{A} & = & \tilde{\mathcal{A}}\sigma. \end{array}$$

Then

$$X_t^p = x + \int_0^t p_s [dW(s) + \theta_s ds], \quad t \in [0, T].$$

Second formulation:

Find

$$V(x) = \sup_{p \in \mathcal{A}} E(U(X_T^p - F)) = \sup_{p \in \mathcal{A}} E(-\exp(-\alpha(x + \int_0^T p_s [dW(s) + \theta_s ds] - F))).$$

Now concept of martingale optimality enters the scene:

Goal: Construct family of processes $R^{(p)}$ s. th.

$$\textbf{form 1} \qquad \begin{array}{rcl} R_0^{(p)} & = & \text{constant}, \\ R_T^{(p)} & = & -\exp(-\alpha(X_T^p - F)), \\ R^{(p)} & & \text{supermartingale}, \quad p \in \mathcal{A}, \\ R^{(p^*)} & & \text{martingale, for (exactly) one} \quad p^* \in \mathcal{A}. \end{array}$$

Then

$$\begin{array}{rcl} E(-\exp(-\alpha[X_T^p - F])) & = & E(R_T^{(p)}) \\ & \leq & E(R_0^p) \\ & = & V(x) \\ & = & E(R_0^{(p^*)}) \\ & = & E(-\exp(-\alpha[X_T^{(p^*)} - F])). \end{array}$$

Hence p^* optimal strategy.

Now introduce BSDE into the solution algorithm: find *generator* f of **BSDE**

$$Y_t = F - \int_t^T Z_s dW(s) - \int_t^T f(s, Z_s) ds, \quad \text{eq.} \quad \begin{array}{l} Y_t = Y_0 + \int_0^t Z_s dW(s) + \int_0^t f(s, Z_s) ds, \\ Y_T = F, \end{array}$$

s. th. with

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

we have

$$\begin{aligned} R_0^{(p)} &= -\exp(-\alpha(x - Y_0)) \\ &= \text{constant}, \qquad \text{(fulfilled)} \end{aligned}$$

$$\textbf{form 2} \qquad R_T^{(p)} = -\exp(-\alpha(X_T^p - F)) \qquad \text{(fulfilled)}$$

$$\begin{aligned} &R^{(p)} && \text{supermartingale}, \quad p \in \mathcal{A}, \\ &R^{(p^*)} && \text{martingale, for (exactly) one} \quad p^* \in \mathcal{A}. \end{aligned}$$

This gives solution of valuation problem.

Goal: determine generator f:

Suppose f is generator of BSDE. Then

$$\begin{aligned} R_t^{(p)} &= -\exp(-\alpha[X_T^p - Y_t]) \\ &= -\exp(-\alpha[x - Y_0]) \cdot \exp(-\alpha[\int_0^t (p_s - Z_s)dW(s) - \int_0^t [f(s, Z_s) - p_s\theta_s]ds]) \\ &= \exp(-\alpha[x - Y_0]) \cdot \exp(-\alpha \int_0^t (p_s - Z_s)dW(s) - \frac{\alpha^2}{2}\int_0^t (p_s - Z_s)^2 ds) \\ &\quad \cdot -\exp(\int_0^t [\alpha f(s, Z_s) - \alpha p_s\theta_s + \frac{\alpha^2}{2}(p_s - Z_s)^2]ds) \\ &= M_t^{(p)} \cdot A_t^{(p)}, \end{aligned}$$

with $M^{(p)}$ nonnegative martingale. $R^{(p)}$ satisfies **(form 2)** iff for

$$q(\cdot, p, z) = f(\cdot, z) - p\theta + \frac{\alpha}{2}(p - z)^2, \quad p \in \mathcal{A}, z \in \mathbb{R},$$

we have

$$\textbf{form 3} \qquad \begin{aligned} q(\cdot, p, z) &\geq 0, \quad p \in C \quad \text{(supermartingale cond.)} \\ q(\cdot, p^*, z) &= 0, \quad \text{for (exactly) one} \quad p^* \in C \quad \text{(martingale cond.)}. \end{aligned}$$

Now

$$\begin{aligned} q(\cdot, p, z) &= f(\cdot, z) - p\theta + \frac{\alpha}{2}(p - z)^2 \\ &= f(\cdot, z) + \frac{\alpha}{2}(p - z)^2 - (p - z) \cdot \theta + \frac{1}{2\alpha}\theta^2 \quad - z\theta - \frac{1}{2\alpha}\theta^2 \\ &= f(\cdot, z) + \frac{\alpha}{2}[p - (z + \frac{1}{\alpha}\theta)]^2 \quad - z\theta - \frac{1}{2\alpha}\theta^2. \end{aligned}$$

Under *non-convex constraint* $p \in C$:

$$[p - (z + \frac{1}{\alpha}\theta)]^2 \geq d^2(C, z + \frac{1}{\alpha}\theta).$$

with **equality** for at least one possible choice of p^* by **closedness** of C. Hence **(form 3)** is solved by the choice

$$\textbf{form 4} \qquad \begin{array}{rcl} f(\cdot,z) & = & -\frac{\alpha}{2}d^2(C, z+\frac{1}{\alpha}\theta) + z\cdot\theta + \frac{1}{2\alpha}\theta^2 \quad \text{(supermartingale cond.)} \\ p^* & & \text{s. th.} \quad d(C, z+\frac{1}{\alpha}\theta) = d(p^*, z+\frac{1}{\alpha}\theta) \quad \text{(martingale cond.)}. \end{array}$$

So we end in **Pb:** Let

$$\pi_C(v) = \{p \in \mathbb{R}^d : d(C,v) = d(p,v)\}.$$

Find measurable selection p_t^* from $\pi_{C_t}(Z_t + \frac{1}{\alpha}\theta_t)$. This is solved by classical *measurable selection method.*

So we obtain

Theorem 9.1. *Let (Y,Z) be unique solution of BSDE*

$$Y_t = F - \int_t^T Z_s dW(s) - \int_t^T f(s, Z_s)ds, \quad t \in [0,T],$$

with

$$f(t, Z_t) = -\frac{\alpha}{2}d^2(C_t, Z_t + \frac{1}{\alpha}\theta_t) + Z_t \cdot \theta_t + \frac{1}{2\alpha}\theta_t^2.$$

Then value function of utility optimization problem under constraint $p \in C$ given by

$$V(x) = -\exp(-\alpha[x - Y_0]).$$

There exists an optimal trading strategy $p^* \in \mathcal{A}$ s. th.

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

Proof:
We need
- existence and uniqueness for BSDE with locally Lipschitz driver f in z,
- measurable selection theorem for $\Pi_{C_t}(Z_t + \frac{1}{\alpha}\theta_t)$
- BMO properties of the martingales $\int Z_s dW(s), \int p_s^* dW(s)$
  for uniform integrability of exponentials (regularity of coefficients)

For details, see [8].•

Goal: existence and uniqueness for BSDE with *globally Lipschitz drivers*, for simplicity in dimension 1.

10 Backward stochastic differential equations

Work on the following spaces:

Time horizon $T > 0$, $(\Omega, \mathcal{F}, P)$ canonical one dim. Wiener space, Wiener process W; $(\mathcal{F}_t)_{t\geq 0}$ natural filtration, completed.

L^2 space of real valued $\mathcal{F}_T$-measurable r.v., norm $E(|X|^2)^{\frac{1}{2}}$.

H^2 space of $(\mathcal{F}_t)_{0\le t\le T}$-adapted processes $X : \Omega \times [0,T] \to \mathbf{R}$, norm $||X||_2 = E(\int_0^T |X_t|^2 dt)^{\frac{1}{2}}$.

For $\beta > 0$, $X \in H^2$ let

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

$H^{2,\beta}$ space H^2, normed by $||\cdot||_{2,\beta}$.

The following are basic *parameters* of a BSDE:

terminal condition: $\xi \in L^2$.

generator: $f : \Omega \times \mathbf{R}_+ \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ product measurable, adapted in t,

$$\text{(H1)} \qquad f(\cdot, 0, 0) \in H^2,$$

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

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

Definition 10.1. (f, ξ) *fulfilling above measurability requirements, (H1), (H2) are called* standard parameters.

Goal: find pair of $(\mathcal{F}_t)_{0\le t\le T}$-adapted processes (Y, Z) s. th. for $0 \le t \le T$

$$\begin{aligned} Y_t &= \xi - \int_t^T Z_s dW(s) - \int_t^T f(\cdot, s, Y_s, Z_s) ds, \quad \text{or} \qquad (1) \\ Y_t &= Y_0 + \int_0^t Z_s dW(s) + \int_0^t f(\cdot, s, Y_s, Z_s) ds, \quad Y_T = \xi. \end{aligned}$$

We call (Y, Z) a *solution of the BSDE* (1). For this purpose, use contraction on suitable Banach space, based on *a priori inequalities.*

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

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

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

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

Proof:

1.Let $(Y,Z) \in H^2 \times H^2$ be solution of (1) with (ξ, f). **Show:**

$$\sup_{0\le t\le T} |Y_t| \in L^2.$$

By (1)

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

and by Doob's inequality

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

By (H1), (H2) $|\xi| + \int_0^T |f(\cdot, s, Y_s, Z_s)|ds \in L^2$. Hence we get

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

2. **Show:** preliminary bound.
We have for $0 \le t \le T$

$$\delta Y_t = \delta Y_0 + \int_0^t \delta Z_s dW(s) + \int_0^t [f^1(., s, Y_s^1, Z_s^1) - f^2(., s, Y_s^2, Z_s^2)]ds.$$

Apply Itô's formula to semimartingale $(e^{\beta s}|\delta Y_s|^2)_{0\le s\le T}$ (integrability of stochastic integral term by 1.). For $0 \le t \le T$

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

Reorder to get

$$\begin{aligned} e^{\beta t}|\delta Y_t|^2 \; + \; & \beta \int_t^T e^{\beta s}|\delta Y_s|^2 ds + \int_t^T e^{\beta s}|\delta Z_s|^2 ds \\ = \; & e^{\beta T}|\delta Y_T|^2 - 2\int_t^T e^{\beta s}\delta Y_s \delta Z_s dW(s) \\ & -2\int_t^T e^{\beta s}\delta Y_s[f^1(\cdot, s, Y_s^1, Z_s^1) - f^2(\cdot, s, Y_s^2, Z_s^2)]ds. \end{aligned}$$

3. For $0 \le t \le T$ **show**:

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

Integrate in the inequality of 2., to get with $\mu, \lambda > 0$, noting that for $a, b \in \mathbf{R}$ we have $2|ab| \leq \mu^2 a^2 + \frac{1}{\mu^2} b^2$

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

Now take term 1 and 2 in the first line of (2) to the r.h.s. to get

$$
\begin{aligned}
E(e^{\beta t}|\delta Y_t|^2) \;\leq\; & E(e^{\beta T}|\delta Y_T|^2) + E(\int_t^T e^{\beta s}|\delta Y_s|^2 ds)[-\beta + C(2+\lambda^2) + \mu^2] \qquad (3) \\
& + E(\int_t^T e^{\beta s}|\delta Z_s|^2 ds)[-1 + \frac{C}{\lambda^2}] + \frac{1}{\mu^2}E(\int_t^T e^{\beta s}|\delta_2 f_s|^2 ds) \\
\leq\; & E(e^{\beta T}|\delta Y_T|^2) + \frac{C - \lambda^2}{\lambda^2}E(\int_t^T e^{\beta s}|\delta Z_s|^2 ds) + \frac{1}{\mu^2}E(\int_t^T e^{\beta s}|\delta_2 f_s|^2 ds).
\end{aligned}
$$

Take second term from the r.h.s. to the left to get claimed inequality.

4. First inequality: integrate inequality from 3. in t, omitting second term on l.h.s.

5. Second inequality: omit first term on l.h.s in 3., set $t = 0$. •

We can deduce existence and uniqueness for the BSDE:

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

$$
Y_t = \xi - \int_t^T Z_s dW(s) - \int_t^T f(\cdot, s, Y_s, Z_s)ds, \quad 0 \leq t \leq T.
$$

Proof:

Consider

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

where (Y, Z) is solution of

$$
Y_t = \xi - \int_t^T Z_s dW(s) - \int_t^T f(\cdot, s, y_s, z_s)ds, \quad 0 \leq t \leq T. \qquad (4)
$$

1. **Show:** (Y, Z) well defined.
By assumptions

$$Y_t := E(\xi - \int_t^T f(\cdot, s, y_s, z_s)ds|\mathcal{F}_t) \in L^2(\Omega), \quad 0 \le t \le T.$$

And with

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

we have

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

M is a square integrable continuous martingale, since we are on Wiener filtration. Martingale representation provides (unique) $Z \in H^2$ s. th.

$$M_t = M_0 + \int_0^t Z_s dW(s), \quad 0 \le t \le T.$$

And we may further write

$$Y_t = M_0 + \int_0^t Z_s dW(s) + \int_0^t f(\cdot, s, y_s, z_s)ds.$$

Hence

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

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

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

This is the claimed equation (4).

2. **Prove:** For $\beta > 2(1+T)C^2$ Γ is contraction.

Let $(y^1, z^1), (y^2, z^2) \in H^{2,\beta} \times H^{2,\beta}$, $(Y^1, Z^1), (Y^2, Z^2)$ solutions of (4) by 1. We apply Lemma 1 with $C = 0, \beta = \mu^2$, and $f^i = f(\cdot, y^i, z^i)$. We obtain

$$\begin{aligned} ||\delta Y||^2_{2,\beta} &\le \frac{T}{\beta} E(\int_0^T e^{\beta s}|f(\cdot, s, y^1_s, z^1_s) - f(\cdot, s, y^2_s, z^2_s)|^2 ds), \\ ||\delta Z||^2_{2,\beta} &\le \frac{1}{\beta} E(\int_0^T e^{\beta s}|f(\cdot, s, y^1_s, z^1_s) - f(\cdot, s, y^2_s, z^2_s)|^2 ds). \end{aligned}$$

By Lipschitz continuity of f

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

We summarize to obtain

$$||\delta Y||_{2,\beta}^2 + ||\delta Z||_{2,\beta}^2 \leq \frac{2C^2(T+1)}{\beta}[||\delta y||_{2,\beta}^2 + ||\delta z||_{2,\beta}^2]. \tag{5}$$

By choice of β, Γ is contraction.

3. Let $(\overline{Y}, \overline{Z})$ be fixed point of Γ given by 2. Then $(\overline{Y}, \overline{Z})$ solves BSDE.

4. Uniqueness: contractivity of Γ, uniqueness of fixed point. •

Goal: recursive approach of solution by a priori inequality.

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

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

by preceding proof. Then $((Y^k, Z^k))_{k\geq 0}$ converges in $H^{2,\beta} \times H^{2,\beta}$ to the solution (Y, Z) of (1).

Proof

(5) in proof of Theorem 1 yields

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

with $\varepsilon = \frac{2C^2(T+1)}{\beta} < 1$. Hence for $m \leq l$

$$[||Y^l - Y^m||_{2,\beta}^2 + ||Z^l - Z^m||_{2,\beta}^2]^{\frac{1}{2}} \leq \sum_{k=m+1}^{l} \epsilon^{\frac{k}{2}}[||Y^1 - Y^0||_{2,\beta}^2 + ||Z^1 - Z^0||_{2,\beta}^2]^{\frac{1}{2}},$$

and $(Y^k, Z^k)_{k\in\mathbf{N}}$ is Cauchy in $H^{2,\beta} \times H^{2,\beta}$. Its limit solves (1). •

Results can be easily extended to $\mathbf{R}^n$ (for Y) and $\mathbf{R}^{m\times n}$ (for Z).

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

In this chapter: establish connection between Malliavin's calculus and BSDE theory. **Goal:** prove that control component Z given by *Malliavin trace* of Y.

For this purpose: need versions of $\mathbf{L}_1^2$ and other spaces with integrability parameter $p \geq 2$.

Definition 11.1. *Let $p \geq 2$, and*

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

For $u \in \mathbf{L}_1^p$ define

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

Denote L^p: space of $(\mathcal{F}_T)$-measurable r. v. X, with norm $||X||_p = E(|X|^p)^{\frac{1}{p}}$, S^p: space of measurable $(\mathcal{F}_t)$-adapted continuous processes $X : \Omega \times [0,T] \to \mathbf{R}$, with norm $||X||_{S^p} = E(\sup_{0\le t\le T} |X_t|^p)^{\frac{1}{p}}$, H^p: space of measurable $(\mathcal{F}_t)$-adapted processes $X : \Omega \times [0,T] \to \mathbf{R}$ with norm $||X||_p = E([\int_0^T X_t^2 dt]^{\frac{p}{2}})^{\frac{1}{p}}$. Let $B^p = S^p \times H^p$, with norm $||(Y,Z)||_p = [||Y||_{S^p}^p + ||Z||_p^p]^{\frac{1}{p}}$.

Here is our main result.

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

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

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

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

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

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

For the unique solution (Y,Z) of the BSDE (1) we suppose

$$\int_0^T ||D_s f(\cdot,t,Y_t,Z_t)||_2^2 ds < \infty.$$

Then:

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

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

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

Proof

1. a): Use notation of proof of Theorem 10.1. **We show:** Γ well defined on B^p. Use appropriate *a priori inequalities.*

Lemma 11.1. *Let $p \geq 2$, assume $f(\cdot, 0, 0) \in H^p$, and define*

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

where (Y, Z) is solution of BSDE

$$Y_t = \xi - \int_t^T Z_s dW(s) + \int_t^T f(\cdot, s, y_s, z_s) ds. \tag{6}$$

For $i = 1, 2$ let (Y^i, Z^i) be solutions corresponding to (y^i, z^i) in (6), and

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

Then there is a constant C_p not depending on (y, z, Y, Z) s. th.

$$\begin{aligned}
(i) \quad ||Y||^p_{S^p} &\leq C_p E([|\xi|^p) + T^{\frac{p}{2}}(\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds)^{\frac{p}{2}}]), \\
(ii) \quad ||Z||^p_p &\leq C_p E([|\xi|^p) + T^{\frac{p}{2}}(\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds)^{\frac{p}{2}}]), \\
(iii) \quad ||(\delta Y, \delta Z)||^p_p &\leq C_p T^{\frac{p}{2}} ||(\delta y, \delta z)||^p_p.
\end{aligned}$$

Proof

(i):

Recall for $0 \leq t \leq T$

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

We have

$$\sup_{0 \leq t \leq T} |Y_t| \leq E(|\xi| + \int_0^T |f(\cdot, s, y_s, z_s)| ds | \mathcal{F}_t).$$

hence Doob's and Cauchy-Schwarz' inequalities provide C^1_p s. th.

$$\begin{aligned}
||Y||^p_{S^p} &\leq C^1_p E([|\xi| + \int_0^T |f(\cdot, s, y_s, z_s)| ds]^p) \\
&\leq C^1_p E([|\xi| + T^{\frac{1}{2}}[\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds]^{\frac{1}{2}}]^p) \\
&\leq 2^p C^1_p E([|\xi|^p + T^{\frac{p}{2}}[\int_0^T |f(\cdot, s, y_s, z_s)|^2 ds]^{\frac{p}{2}}]).
\end{aligned}$$

Use $|f(\cdot, y, z)| \leq |f(\cdot, 0, 0)| + C(|y| + |z|)$, and invoke $f(\cdot, 0, 0) \in H^p$, and $(y, z) \in B^p$, to see that r. h. s. of preceding inequality is finite.

(ii):

We use inequality of Burkholder (for martingales $(M_t)_{0 \leq t \leq T}$ with quadratic variation

$(\langle M\rangle_t)_{0\le t\le T}$ and $p\ge 1$ there are $c_p, C_p \in \mathbf{R}$ s. th. $c_pE([\langle M\rangle_T]^{\frac{p}{2}}) \le E(|M_T|^p) \le C_pE([\langle M\rangle_T]^{\frac{p}{2}})$. Together with (i) it provides universal $C_p^2, \cdot, C_p^4$ s. th.

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

Hence $Z\in H^p$, and therefore $(Y,Z)\in B^p$.

(iii):

The solution $(\delta Y,\delta Z)$ belongs to generator $f(\cdot,t,y_t^1,z_t^1)-f(\cdot,t,y_t^2,z_t^2)$, and $\xi=0$. Hence (i) and (ii), as well as the Lipschitz condition, give, with universal C_p^5, C_p^6

$$\begin{aligned} ||(\delta Y,\delta Z)||_p^p &\le C_p^5T^{\frac{p}{2}}E([\int_0^T |f(\cdot,t,y_t^1,z_t^1)-f(\cdot,t,y_t^2,z_t^2)|^2dt]^{\frac{p}{2}}) \\ &\le C_p^6T^{\frac{p}{2}}[||\delta y||_{S^p}^p + ||\delta z||_p^p] \\ &= C_p^6T^{\frac{p}{2}}||(\delta y,\delta z)||_p^p. \end{aligned}$$

•

b): Now return to proof of Theorem 11.1. Define recursive approximations of solution of BSDE: for $k\ge 0, 0\le t\le T$ let

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

Show:

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

Recall universal constant C_4 from Lemma 11.1, (iii). Modulo repeating argument finitely often on successive small subintervals of $[0,T]$, assume that $\epsilon := [C_4T^2]^{\frac{1}{4}} < 1$. With this, Lemma 11.1 implies that Γ is contraction, and solution (Y,Z) of BSDE (1) its unique fixed point in B^4. Moreover, $(Y^k,Z^k)_{k\in\mathbf{N}}$ Cauchy sequence in B^4: for $k\le l$

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

2. **Prove by recursion on k:**

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

This is trivial for $k=0$. Let it be guaranteed for k. By hypotheses concerning standard parameters: 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

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

To see this: discretize Lebesgue integral to Riemann sums, use chain rule, approximate using boundedness of partial derivatives, (H4) and closedness of D. Hence Lemma 8.2 yields for $0 \le t \le T$

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

as well. Consequently, also

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

Now Theorem 8.4 implies $Z^{k+1} \in \mathbf{L}_1^2$ and in $L^2(\Omega \times [0,T]^2)$ we have the equation

$$D_s \int_t^T Z_u^{k+1} dW(u) = 1_{[t,T]}(s) Z_s^{k+1} + \int_t^T D_s Z_u^{k+1} dW(u).$$

All stated differentiabilities combine with square integrability of Malliavin derivatives in all variables. This means

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

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

$$\begin{aligned} D_s Y_t^{k+1} &= D_s\xi - \int_t^T D_s Z_u^{k+1} dW(u) \\ &- \int_t^T [\frac{\partial}{\partial y} f(\cdot, u, Y_u^k, Z_u^k)\, D_s Y_u^k + \frac{\partial}{\partial z} f(\cdot, u, Y_u^k, Z_u^k)\, D_s Z_u^k \\ &\qquad + D_s f(\cdot, u, Y_u^k, Z_u^k)]du. \end{aligned} \tag{7}$$

3. **Show:**

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

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

$$\begin{aligned} Y_t^s &= D_s\xi - \int_t^T Z_u^s dW(u) - \int_t^T [\frac{\partial}{\partial y} f(\cdot, u, Y_u, Z_u)\, Y_u^s + \frac{\partial}{\partial z} f(\cdot, u, Y_u, Z_u)\, Z_u^s \\ &\qquad + D_s f(\cdot, u, Y_u, Z_u)]du, \quad 0 \le s \le t \le T, \\ Y_t^s &= Z_t^s = 0, \quad 0 \le t < s \le T. \end{aligned} \tag{8}$$

By hypotheses, at least for λ-a.e. $0 \le s \le T$ the parameters $(F^s, D_s\xi)$ with

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

$0 \le t \le T, y, z \in \mathbf{R}$, are standard. Hence (Y^s, Z^s) well defined (and set trivial where parameters fail to be standard).

a) Show: $\int_0^T ||(Y^s,Z^s)||_2^2 ds < \infty$, hence $(Y^s,Z^s) \in B^2$ for λ-a.a. $0 \le s \le T$. Again our arguments are based on *a priori inequalities.*

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

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

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

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

$0 \le t \le T$. Then for T small enough there is $C_{p,T}$ s. th.

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

Proof

With calculations analogous to the ones proving Lemma 10.1, (i) and (ii) in Lemma 11.1, and with Doob's and Cauchy-Schwarz' inequalities, we get with universal constants $C_p^1, \cdot, C_p^3$

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

Now choose T small to ensure $C_p^3(T^p + T^{\frac{p}{2}}) < 1$. Take the last two expressions in the inequality from the r.h.s. to the l.h.s., to obtain the desired estimate. •

Now fix $0 \le s \le T$, and apply Lemma 11.2 with

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

$0 \le t \le T, y, z \in \mathbf{R}$. Then

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

With some universal constant C we get the inequality

$$||(Y^s, Z^s)||_2^2 \le C[||D_s\xi||_2^2 + ||D_s f(\cdot, Y, Z)||_2^2], \text{ and so } \int_0^T ||(Y^s, Z^s)||_2^2 ds < \infty$$

by hypotheses.

b) Show: estimates for differences of (DY^k, DZ^k) and $(Y^\cdot, Z^\cdot)$.
Fix s as in a), $k \in \mathbf{N}$, and apply Lemma 11.2 to the following parameters

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

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

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

The Lemma yields

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

with a universal constant C_1. Let us further estimate the r. h. s.. We have with a universal C_2

$$E((\int_s^T |\delta_t^k| dt)^2) \le C_2[A_k^s(T) + B_k^s(T) + C_k^s(T)],$$

where

$$\begin{aligned}
A_k^s(T) &= E([\int_s^T |D_s f(\cdot, t, Y_t, Z_t) - D_s f(\cdot, t, Y_t^k, Z_t^k)| dt]^2), \\
B_k^s(T) &= E([\int_s^T |\frac{\partial}{\partial y} f(\cdot, t, Y_t^k, Z_t^k)(Y_t^s - D_s Y_t^k)| dt]^2) \\
&+ E([\int_s^T |\frac{\partial}{\partial z} f(\cdot, t, Y_t^k, Z_t^k)(Z_t^s - D_s Z_t^k)| dt]^2), \\
C_k^s(T) &= E([\int_s^T |\frac{\partial}{\partial y} f(\cdot, t, Y_t, Z_t) - \frac{\partial}{\partial y} f(\cdot, t, Y_t^k, Z_t^k)|\, |Y_t^s| dt]^2) \\
&+ E([\int_s^T |\frac{\partial}{\partial z} f(\cdot, t, Y_t, Z_t) - \frac{\partial}{\partial z} f(\cdot, t, Y_t^k, Z_t^k)|\, |Z_t^s| dt]^2).
\end{aligned}$$

ba): $A_k^{\cdot}(T)$:
With further universal constants C_3, C_4 we deduce, using (H4)

$$\begin{aligned} A_k^s(T) &\leq E([\int_s^T K_s(t)[|Y_t - Y_t^k| + |Z_t - Z_t^k|]dt]^2) \\ &\leq C_3 E(\int_s^T K_s(t)^2 dt[\int_s^T |Y_t - Y_t^k|^2 dt + \int_s^T |Z_t - Z_t^k|^2 dt]) \\ &\leq C_4[||K_s||_4^2][||Y - Y^k||_4^2 + ||Z - Z^k||_4^2]. \end{aligned}$$

Hence by part 1. and (H4)

$$\lim_{k\to\infty} \int_0^T A_k^s(T) ds = 0.$$

bb): $C_k^{\cdot}(T)$:
Since the partial derivatives of f with respect to y, z are bounded and continuous, and since $\int_0^T ||(Y^s, Z^s)||_2^2 ds < \infty$, by dominated convergence

$$\lim_{k\to\infty} \int_0^T C_k^s(T) ds = 0.$$

bc): $B_k^{\cdot}(T)$:
By boundedness of partial derivatives of f with a universal constant C_5

$$B_k^s(T) \leq C_5 T^2 ||D_s Y^k - Y^s, D_s Z^k - Z^s)||_2^2.$$

Choose T small to ensure $\alpha = C_5 T^2 < 1$. Let $\varepsilon > 0$. By ba) and bb) there is $N \in \mathbf{N}$ s. th. for $k \geq N$ we have

$$\begin{aligned} &\int_0^T ||(D_s Y^{k+1} - Y^s, D_s Z^{k+1} - Z^s)||^2 ds \\ &\quad \leq \varepsilon + \alpha \int_0^T ||(D_s Y^k - Y^s, D_s Z^k - Z^s)||^2 ds. \end{aligned}$$

By recursion for $k \geq N$

$$\begin{aligned} \int_0^T ||(D_s Y^k - Y^s, D_s Z^k - Z^s)||^2 ds &\leq \varepsilon(1 + \alpha + \alpha^2 + \cdots + \alpha^{k-N-1}) \\ &\quad + \alpha^{k-N} \int_0^T ||(D_s Y^N - Y^s, D_s Z^N - Z^s)||^2 ds \\ &\leq \frac{\varepsilon}{1-\alpha} + \alpha^{k-N} \int_0^T ||(D_s Y^N - Y^s, D_s Z^N - Z^s)||^2 ds. \end{aligned}$$

Now let $k \to \infty$. Since ε is arbitrary, we conclude

$$\lim_{k\to\infty} \int_0^T B_k^s(T) ds = 0.$$

In summary, $(Y^k, Z^k) \to (Y^\cdot, Z^\cdot)$ in $L^2(\Omega \times [0,T]^2)$.

3. Since $\mathbf{L_1^2}$ Hilbert space, and D closed operator, $(Y,Z) \in \mathbf{L}_1^2 \times \mathbf{L}_1^2$, and $(Y^s, Z^s)_{0 \le s \le T}$ is version of $(D_s Y, D_s Z)_{0 \le s \le T}$ in the usual sense.

4. **We show:**

$$(D_t Y_t)_{0 \le t \le T} \text{ is version of } (Z_t)_{0 \le t \le T}.$$

For $u < t$ we have

$$Y_t = Y_u + \int_u^t Z_r dW(r) + \int_u^t f(\cdot, r, Y_r, Z_r) dr.$$

Hence by parts 2. and 3. for $u < s \le t$

$$\begin{aligned} D_s Y_t &= Z_s + \int_s^t D_s Z_r dW(r) \\ &\quad + \int_s^t [\frac{\partial}{\partial y} f(\cdot, r, Y_r, Z_r)\, D_s Y_r + \frac{\partial}{\partial z} f(\cdot, r, Y_r, Z_r)\, D_s Z_r \\ &\qquad + D_s f(\cdot, r, Y_r, Z_r)] dr. \end{aligned}$$

By continuity in t of Y^s we may choose $t = s$, to obtain the claimed identity. •

Literatur

[1] S. Ankirchner, P. Imkeller, G. Reis *Classical and Variational Differentiability of BSDEs with Quadratic Growth.* Electron. J. Probab. 12 (2007), 1418–1453 (electronic).

[2] S. Ankirchner, P. Imkeller, G. Reis *Pricing and hedging of derivatives based on non-tradable underlyings.* Math. Finance (2010), Vol. 20, No. 2, 289–312.

[3] Bell, D. *The Malliavin calculus.* Pitman Monographs and Surveys in Pure and Applied Math. 34. Longman and Wiley: 1987.

[4] Bell, D. *Degenerate stochastic differential equations and hypoellipticity.* Pitman Monographs and Surveys in Pure and Applied Mathematics. 79. Harlow, Essex: Longman (1995).

[5] Bismut, J. M. *Martingales, the Malliavin Calculus and Hypoellipticity under General Hörmander's Conditions.* Z. Wahrscheinlichkeitstheorie verw. Geb. 56 (1981), 469-505.

[6] Bismut, J. M. *Large deviations and the Malliavin calculus.* Progress in Math. 45, Birkhäuser, Basel 1984.

[7] El Karoui, N.; Peng, S.; Quenez, M.C. Backward stochastic differential equations in finance. Math. Finance 7, No.1, 1-71 (1997).

[8] Hu, Y., Imkeller, P., Müller, M. *Utility maximization in incomplete markets.* Ann. Appl. Probab. 15 (2005), No.3, 1691-1712.

[9] Janson, Svante *Gaussian Hilbert spaces.* Cambridge Tracts in Mathematics. 129. Cambridge: Cambridge University Press (1997).

[10] Malliavin, P. Stochastic analysis. Grundlehren der Mathematischen Wissenschaften. 313. Berlin: Springer (1997).

[11] Malliavin, P. Integration and probability. Graduate Texts in Mathematics. 157. New York, NY: Springer (1995).

[12] Malliavin, P. *Stochastic calculus of variation and hypoelliptic operators.* Proc. Intern. Symp. SDE Kyoto, 1976, 195-263. Kinokynia: Tokyo 1978.

[13] Malliavin, P. C^k-*hypoellipticity with degeneracy.* Stochastic Analysis. ed: A. Friedman and M. Pinsky, 199-214, 327-340. Acad. Press: New York 1978.

[14] **Malliavin, P., Thalmaier, A. Stochastic calculus of variations in mathematical finance. Springer: Berlin 2006.**

[15] Nualart, D. The Malliavin calculus and related topics. Springer: Berlin 1995.

[16] Revuz, D., Yor, M. *Continuous martingales and Brownian motion.* Springer, Berlin 1991.

[17] Stroock, D. W. *The Malliavin Calculus. Functional Analytic Approach.* J. Funct. Anal. 44 (1981), 212-257.