Stochastic Analysis on Hilbert Spaces

Thorsten Schmidt

Department of Mathematics, University of Leipzig,
Augustusplatz 10/11 04109 Leipzig Germany
thorsten.schmidt@math.uni-leipzig.de

This version: 4th July 2005

Contents

Chapter 1

Introduction

1.1 Motivation

Recent developments in interest rate theory require models which allow for infinitely many factors. This stems from the idealized assumption that bonds from every maturity are traded in the market. Besides being more general this theory allows for a number of deeper insights and new results. Besides from interest rate theory there are several other places, where infinite dimensional models can be applied. To name just a few, we mention location based models in mortgage risk, credit risk, physics (random motion of a string - therefore infinite dimensional models are sometimes called string models) and population genetics. The introduction in Da Prato and Zabczyk (1992) gives some more examples.

We recall the problem in the interest rate market. This markets consists of bonds. A buyer of a bond is promised to be paid the notional, say 1, at a future time point (the maturity T). The price of the bond at time $t < T$ is denoted by $B(t,T)$. Then the market consists, at time t, of a certain number of bonds. In practice, this number is of course finite. Nevertheless, already the famous model from Heath, Jarrow, and Morton (1992) assumes infinitely many forward rates. So we face a market with bonds with all maturities $t \leq T \leq T^*$. One question to pose is, under which assumptions is the market free of arbitrage. A much more difficult question to answer is about completeness of the model. This is, because in principle, an infinite number of traded assets is available and one has to find suitable trading strategies.

In this framework the question of *consistency* can be posed. The motivation comes from the practice of central banks, which interpolate the several bond prices according to a parametric family to obtain a smooth yield curve. Starting with such a parametric family, a consistent model remains in this family. An excellent introduction to this area is Filipović (2001).

1.2 Preliminaries

1.2.1 Some Basic Spaces

We will mainly consider Hilbert spaces and in some times Banach spaces. For fundamentals in functional analysis we recommend Werner (2000) and Yosida (1971).

We will need some measure of distances in our spaces. Assume X is a $\mathbb{K}$-vector space.

Definition 1.2.1. A mapping $p : X \to [0, \infty)$ is called *norm*, if

(i) $p(c\,x) = |c|p(x)$ for all $c \in \mathbb{K}$ and $x \in X$,

(ii) $p(x + y) \leq p(x) + p(y)$ (triangular inequality),

(iii) $p(x) = 0 \Rightarrow x = 0$.

The couple $(X, \|\cdot\|)$ is called *normed space* if $\|\cdot\|$ is a norm. A norm induces a metric, which is a mapping on $X \times X$, by setting $d(x, y) := p(x - y)$.

Besides this, limits of sequences should not leave the space. A sequence $\{x_n : n \in \mathbb{N}\}$ converges to a limit $x \in X$, iff the distance $d(x_n, x)$ converges to zero, or equivalently, it forms a *Cauchy sequence*, i.e. $\lim_{n,m\to\infty} d(x_n, x_m) = 0$. In a metric space this limit is unique.

Definition 1.2.2. A normed space X is called *complete*, if every Cauchy sequence of elements of X converges to a limit point in X. A normed, complete space is called a *Banach space*.

If we have a normed space which is not complete, it can be imbedded in a larger space (which contains all the limit points additionally) in such a way that the larger space is a Banach space. At heart of this argument lies the Hahn-Banach theorem.

Example 1.2.3. Continuous and differentiable functions: $C^b(T), C^r[a, b]$.

1. Consider a topological space T (or think of subsets of $\mathbb{R}$). Then the space of continuous, bounded functions from T to $\mathbb{R}$, denoted by $C^b(T)$ is a Banach space with the sup-norm, $\|f\|_\infty := \sup_{t\in T} |f(t)|$,

2. Define $C^r[a, b]$ the space of continuous, r-times differentiable functions : $[a, b] \mapsto \mathbb{R}$. $C[a, b]$ is the space of just continuous functions. Then $(C^1[a, b], \|\cdot\|_\infty)$ is not a Banach space. But, setting $|||x||| := \|x\|_\infty + \|x'\|_\infty$, $(C^1[a, b], |||\cdot|||)$ is a Banach space (and similar for $C^r[a, b]$).

It will be convenient to work with *separable* spaces, especially for Hilbert spaces.

Definition 1.2.4. A metric space is called *separable* if it has a countable, dense subset.

In this sense, D is dense in T, if the closure $\overline{D}$ is equal to T. The closure is obtained by adding all limit points.

A *Hilbert space* is essentially a Banach space with a scalar product.

Definition 1.2.5. Consider a $\mathbb{K}$-vector space. A scalar product is a mapping $\langle\cdot,\cdot\rangle : X \times X \mapsto \mathbb{K}$ with

(i) $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle$ for all $x_1, x_2, y \in X$,

(ii) $\langle cx, y\rangle = c\langle x, y\rangle$ for all $x, y \in X$ and $c \in \mathbb{K}$,

(iii) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ for all $x, y \in X$,

(iv) $\langle x, x\rangle \geq 0$ for all $x \in X$,

(v) $\langle x, x\rangle = 0 \Leftrightarrow x = 0$.

If $\mathbb{K} = \mathbb{R}$ the complex conjugate in (iii) is not necessary, and the scalar product is a bilinear mapping. Note that $\|\cdot\| := \langle x, x\rangle^{1/2}$ defines a norm.

Definition 1.2.6. A complete, normed space $(X, \|\cdot\|)$ is called a *Hilbert space*, if there exists a scalar product $\langle\cdot,\cdot\rangle$, such that $\langle x, x\rangle^{1/2} = \|x\|$ for all $x \in X$.

A Hilbert space has an orthonormal basis, such that we can span it by the basis vectors. Any linearly independent set can be made orthonormal by the so-called *Gram-Schmidt* orthogonlization procedure.

Example 1.2.7. We have some important examples

1. For a measurable set $\Omega \subset \mathbb{R}^n$ the space of square integrable functions, $L^2(\Omega)$ has a scalar product defined by (λ denotes the Lebesgue-measure)

$$\langle f, g\rangle = \int_\Omega f\,\bar{g}\,d\lambda. \tag{1.1}$$

   Note that this space is not yet a Hilbert space, as there are different functions which have the norm zero. This is because one can distinguish functions only up to a Lebesgue null set. Therefore one introduces equivalence classes to overcome this problem. More precisely, define

$$\mathscr{L}^p(\Omega) := \left\{ f : \Omega \mapsto \mathbb{K} : f \text{ measurable }, \int_I |f|^p\,d\lambda < \infty \right\}$$

   with norm $\|f\|_p := \left(\int_\Omega |f|^p d\lambda\right)^{1/p}$. Considering $N_p := \{f : f = 0 \text{ almost everywhere}\}$ one defines the quotient vector space

$$L^p(\Omega) := \mathscr{L}^p / N_p.$$

   It can be shown, that $L^p(\Omega)$ is a Banach space for $p \geq 1$, see Werner (2000, p.13 ff). Moreover, $L^2(\Omega)$ is a Hilbert space with the scalar product defined in (1.1).

2. More generally, all spaces $L^2(\mu)$ are Hilbert spaces, where $L^2(\mu)$ denotes the space of square integrable functions w.r.t. some measure μ (of course we always consider the above quotient space).

3. Sobolev spaces are Hilbert spaces. We come back to this later.

Finally, we mention some brief facts on linear operators.

Definition 1.2.8. A continuous and linear mapping between two normed spaces is called continuous *operator*. If we consider $\mathbb{K}$-vector spaces and the image of the operator is $\mathbb{K}$ it is named *functional.*

The following is elementary, but of great importance.

Proposition 1.2.9. *Consider two normed spaces, X and Y and a linear mapping $T : X \to Y$. Then the following are equivalent:*

(i) T is continuous,

(ii) T is continuous at 0,

(iii) there exists $M \geq 0$, s.t. $\| Tx\| \leq M\| x\| \quad \forall x \in X$,

(iv) T is uniformly continuous.

This leads to the norm of a linear mapping.

$$\| T\| := \inf\{M \geq 0 : \| Tx\| \leq M\| x\| \forall x \in X\}.$$

Then we also have that $\| T\| = \sup_{\| x\| =1} \| Tx\|$ and $\| Tx\| \leq \| T\| \, \| x\|$.

Example 1.2.10. Most of this examples are well-known.

1. For $X = Y$ the identity is linear with $\| \operatorname{id} \| = 1$.
2. Integral- and differential operators are continuous and linear, considered in the right norms.
3. Consider $k : [0,1]^2 \mapsto \mathbb{R}$ and $f \in C[0,1]$. Define

$$(T_k f)(s) := \int_0^1 k(s,t) f(t)\, dt.$$

If k is uniformly continuous, then $T_k x$ is continuous, so $T_k : C[0,1] \to C[0,1]$. Furthermore, T_k is continuous w.r.t. the sup-norm.

The spectral representation of a linear operator is an important tool. We need to know what a compact operator is.

Definition 1.2.11. Consider normed spaces X, Y. A linear mapping $T : X \to Y$ is called *compact*, if the closure of $T(B_X)$ is compact, where $B_X := \{x \in X : \sup_{\| x\| \leq 1}\}$.

Then the following holds.

Proposition 1.2.12. *For a compact operator $T : X \to Y$ there exist orthonormal bases $\{e_1, \dots\}$ and $\{f_1, \dots\}$ of X and Y and real numbers $s_1 \geq s_2 \geq \dots$ with $s_k \to 0$, such that*

$$Tx = \sum_{k=1}^{\infty} s_k \langle x, e_k\rangle f_k.$$

The numbers s_k^2 are the eigenvalues of T^*T [1].

[1] Here the operator T^* is the (in Hilbert space sense) adjoint operator of T, i.e. the operator for which $\langle Tf, g\rangle = \langle f, T^* g\rangle$ holds.

If we have a normal operator, this representation gets even more explicit. In the case where $\mathbb{K} = \mathbb{R}$ a selfadjoint operator is already called a normal operator.

Proposition 1.2.13. (Spectral representation for compact operators)
Consider a compact operator $T : H \to H$. Denote by $\{e_1, \dots\}$ the eigenvectors and $\{\lambda_1, \dots\}$ the (distinct) nonzero eigenvalues of T. Then H is the direct sum of the kernel of T and the span of the eigenvectors, i.e. $H = \ker T \oplus \overline{\mathrm{lin}}\{e_1, \dots\}$ and

$$Tx = \sum_{k=1}^{\infty} \lambda_k \langle x, e_k \rangle e_k \qquad \forall x \in H.$$

Furthermore, $\| T \| = \sup_k |\lambda_k|$.

The scalar product has an important property: it is continuous.

Lemma 1.2.14. *The scalar product as a mapping $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ is continuous.*

Proof. Observe, that for $f_1, f_2, g_1, g_2 \in H$

$$\begin{aligned} |\langle f_1, g_1 \rangle - \langle f_2, g_2 \rangle| &= |\langle f_1, g_1 - g_2 \rangle + \langle f_1 - f_2, g_2 | \\ &\leq \| f_1 \| \, \| g_1 - g_2 \| + \| f_1 - f_2 \| \, \| g_2 \| . \end{aligned}$$

The last inequality results from the triangle inequality and the Cauchy-Schwartz inequality. The conclusion follows. ■

The following will be quite useful to us.

Lemma 1.2.15. *If the series $\sum_k f_k$ of elements in H converges (in H), then*

$$\langle \sum_{k=1}^{\infty} f_k, g \rangle = \sum_{k=1}^{\infty} \langle f_k, g \rangle \qquad \textit{for all } g \in H.$$

Proof. Besides the continuity we just need the linearity of the scalar product to obtain

$$\begin{aligned} \langle \sum_{k=1}^{\infty} f_k, g \rangle &= \langle \lim_{n \to \infty} \sum_{k=1}^{\infty} f_k, g \rangle = \lim_{n \to \infty} \langle \sum_{k=1}^{n} f_k, g \rangle \\ &= \sum_{k=1}^{\infty} \langle f_k, g \rangle . \end{aligned}$$

■

1.2.2 Gaussian Measures on Hilbert Spaces

Consider a separable Hilbert space H with an inner product $\langle \cdot, \cdot \rangle$. The space of linear, continuous mappings from H into itself is denoted by $L(H)$. Note that $L(H)$ is a Banach space, if equipped with the operator norm, because due to Proposition 1.2.9, a linear continuous operator is also bounded. For $D \in L(H)$ and $h \in H$ we often write $D \cdot h$ instead of $D(h)$. The Borel σ-algebra $\mathcal{B}(H)$ is the σ-algebra induced by the norm of H.

We start by defining normality for probability measures in H.

Definition 1.2.16. A probability measure μ on $(H, \mathcal{B}(H))$ is said to be Gaussian, if and only if for any $h \in H$ there exist $p, q \in \mathbb{R}$ with $q \geq 0$, such that

$$\mu\{x \in H, \langle h, x\rangle \in A\} = \mathcal{N}(p,q)(A), \qquad \forall A \in \mathcal{B}(\mathbb{R}).$$

Here, $\mathcal{N}(p,q)$ denotes the Gaussian measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with mean p and variance q.

Second, we generalize the concept of mean and variance.

Definition 1.2.17. For a Gaussian measure μ on $(H, \mathcal{B}(H))$ the element $m \in H$ such that

$$\int_H \langle h, x\rangle \, \mu(dx) = \langle m, h\rangle \qquad \forall h \in H$$

is called the *mean* of μ. The symmetric, nonnegative operator $D \in L(H)$ with

$$\int_H \langle h_1, x\rangle \, \langle h_2, x\rangle \mu(dx) - \langle m, h_1\rangle \, \langle m, h_2\rangle = \langle Dh_1, h_2\rangle \qquad \forall h_1, h_2 \in H$$

is called the *covariance operator* of μ.

Here D is symmetric in the sense, that $\langle Dh_1, h_2\rangle = \langle Dh_2, h_1\rangle$. As for $\mathbb{R}^n$, mean and covariance operator uniquely determine μ. For a random variable with distribution μ we write $\xi \sim \mathcal{N}(m, D)$ and $\mathrm{Cov}(\xi) := D$.

Remark 1.2.18. The connection with Definition 1.2.16 is the following. For $h \in H$ and $p, q \in \mathbb{R}$ such that

$$\mu\{x \in H, \langle h, x\rangle \in A\} = \mathcal{N}(p,q)(A),$$

we have $p = \langle h, m\rangle$ and $q = \langle Dh, h\rangle$.

To see the analogue to the case $H = \mathbb{R}^n$ note that in this case, a measure is Gaussian, iff the characteristic function takes the form

$$\varphi(\boldsymbol{\lambda}) = \exp\big(\mathrm{i}\boldsymbol{\lambda}^\top \mathbf{m} - \frac{1}{2}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}\,\boldsymbol{\lambda}\big), \qquad \boldsymbol{\lambda} \in \mathbb{R}^n,$$

for appropriate $\mathbf{m} \in \mathbb{R}^n$ and $\boldsymbol{\Sigma} \in \mathbb{R}^n \times \mathbb{R}^n$. In this case, we have that

$$\boldsymbol{\lambda}^\top \mathbf{m} = \int \langle \boldsymbol{\lambda}, \mathbf{x}\rangle \, \mu(d\mathbf{x})$$

and

$$\begin{aligned}\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}\,\boldsymbol{\lambda} &= \boldsymbol{\lambda}^\top \Big(\int (\mathbf{x} - \mathbf{m})\,(\mathbf{x} - \mathbf{m})^\top \mu(d\mathbf{x})\Big)\,\boldsymbol{\lambda} \\ &= \int \langle \boldsymbol{\lambda}, \mathbf{x}\rangle\langle \boldsymbol{\lambda}, \mathbf{x}\rangle \, \mu(d\mathbf{x}) - \langle \boldsymbol{\lambda}, \mathbf{m}\rangle^2.\end{aligned}$$

Coming back to linear operators on H, we introduce an elementary but important tool, the *trace*.

Definition 1.2.19. Consider an orthonormal basis $\{e_k : k \in \mathbb{N}\}$ of H. For any linear operator D on H we define the *trace* of D by

$$\operatorname{tr} D := \sum_{j=1}^{\infty} \langle De_j, e_j \rangle,$$

if the above series converges absolutely, and set the trace equal to infinity otherwise. The trace is independent of the chosen basis. If $\sum |\langle De_j, e_j \rangle| < \infty$, the operator D is called *trace-class*. We denote the Banach space of trace-class operators by $L_1(H)$ and its norm by $\| \cdot \|_1$, compare Da Prato and Zabczyk (1992, Appendix C).

It can be shown that the covariance operator of a Gaussian probability measure is a trace class operator, see Da Prato and Zabczyk (1992, Proposition 2.15). Note that for positive D trace-class already follows from $\operatorname{tr} D < \infty$.

In our applications H will be a space of functions $h : \mathbb{R} \mapsto \mathbb{R}$. For a stochastic process $(X(s))_{s \geq 0}$ which takes values in H we set $X(s,t) := X(s)(t)$.

That is, if we consider the process of forward rates $(f(s,t))_{s \geq 0}$, $f(s,t)$ represents the forward rate at time s with maturity t, while $f(s)$ represents the whole term structure at time s.

Definition 1.2.20. For a symmetric, nonnegative trace-class operator $D \in L(H)$ the H-valued process $(X(s))_{s \geq 0}$ is called a *D-Wiener process* if

(i) $X(0) = 0$,

(ii) X has continuous trajectories,

(iii) X has independent increments,

(iv) the distribution of $X(s_2) - X(s_1)$ is a Gaussian measure on H with mean 0 and covariance operator $(s_2 - s_1)D$.

If the considered probability space admits a filtration $(\mathcal{F}_s)_{s \geq 0}$ satisfying the usual conditions[2], $X(s)$ is $\mathcal{F}_s$-measurable and $X(s_2) - X(s_1)$ is independent of $\mathcal{F}_{s_1}$ for all $s_2 > s_1 \geq 0$, we say that X is a D-Wiener process with respect to $(\mathcal{F}_s)_{s \geq 0}$.

Property (iv) specifies the covariance structure of $(X_s)_{s \geq 0}$. The covariance operator of a certain increment, say $X_{s_2} - X_{s_1}$, may be decomposed into a factor which depends only on time, namely $s_2 - s_1$, and an operator D. The operator D refers to the covariance in the Hilbert space. For $s_1 = 0$ and $s_2 = s$ one might think of $\operatorname{Cov}\big(X(s,t_1), X(s,t_2)\big)$, so the second factor describes the covariance w.r.t. the maturity (t_1, t_2 respectively).

Compare to the case $H = \mathbb{R}^n$. A Brownian motion with covariance function $(s \wedge t)\mathbf{\Sigma} = (s \wedge t)\mathbf{A}\mathbf{A}^\top$ is obtained from a Brownian motion $(\mathbf{B}_s)_{s \geq 0}$ with independent components via $(\mathbf{A}\mathbf{B}_s)_{s \geq 0}$. As this procedure can not be transferred to the infinite dimensional case without difficulties, the covariance operator always needs to be specified explicitly, and this is why we speak of a D-Wiener process.

[2] This means, that $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets and the filtration is right-continuous, which is called the usual augmentation of $(\mathcal{F})$, see Revuz and Yor (1994).

We now develop the eigenvalue expansion of a H-valued random variable ξ. This will be crucial if we consider linear operators on $\xi = X(s)$.

Observe that for any orthonormal basis $\{e_k : k \in \mathbb{N}\}$ of H and $f \in H$ we have

$$f = \sum_k \langle f, e_k\rangle\, e_k,$$

where the Fourier-coefficients $\langle f, e_k\rangle$ are real-valued random variables. For $D \in L(H)$ and $f, g \in H$ we obtain

$$\begin{aligned}\langle Df, g\rangle &= \langle D \sum_k e_k \langle f, e_k\rangle, \sum_l e_l \langle g, e_l\rangle\rangle \\ &= \sum_{k,l} \langle f, e_k\rangle\, \langle De_k, e_l\rangle\, \langle g, e_l\rangle.\end{aligned}$$

It is interesting to find a basis which simplifies the above expression. Assume that D is a covariance operator, in particular, D is nonnegative and trace-class. The eigenvectors of D form a complete orthonormal system $\{e_k : k \in \mathbb{N}\}$ while the eigenvalues λ_k form a bounded sequence of nonnegative real numbers, such that[3]

$$D\, e_k = \lambda_k\, e_k. \tag{1.2}$$

Because D is a trace-class operator, we have $\sum_k \lambda_k < \infty$ and obtain

$$\langle Df, g\rangle = \sum_k \lambda_k \langle e_k, f\rangle\, \langle e_k, g\rangle.$$

If we consider $\xi = X(s)$, where $(X(s))$ is a D-Wiener process, this eigenvalue expansion gives a useful representation.

Proposition 1.2.21. *Consider a D-Wiener process $(X(s))_{s\geq 0}$ and denote by $\{e_k : k \in \mathbb{N}\}$ the eigenvectors of D. Define*

$$\beta_k(s) := \langle X(s), e_k\rangle.$$

Then, for $\lambda_k > 0$, $\frac{1}{\sqrt{\lambda_k}}\beta_k(s)$ are mutually independent Brownian motions. Moreover, we have the decomposition

$$X(s) = \sum_{k=1}^{\infty} \beta_k(s) e_k, \tag{1.3}$$

and the series in (1.3) converges in $L^2(\Omega, \mathcal{A}, \mathbb{P})$.

Proof. Clearly β_k is a continuous, centered Gaussian process. Then

$$\mathbb{E}\Big[\frac{\beta_k(s)\,\beta_l(t)}{\sqrt{\lambda_k \lambda_l}}\Big] = \frac{1}{\sqrt{\lambda_k\lambda_l}}\,\mathbb{E}\Big[\langle X(s), e_k\rangle \langle X(t), e_l\rangle\Big],$$

[3] See Da Prato and Zabczyk (1992, p.86). Note that the system $\{e_k\}$ certainly depends on D. In the following we always refer to this particular $\{e_k\}$ without stressing the dependence on D.

and because $(X(s))_{s\geq 0}$ has independent increments, we obtain

$$\mathbb{E}\Big[\langle X(s), e_k\rangle\langle X(t), e_l\rangle\Big] = \mathbb{E}\Big[\langle X(s\wedge t), e_k\rangle\langle X(s\wedge t), e_l\rangle\Big].$$

According to item (iv) of Definition 1.2.20, the distribution of $X(s\wedge t)$ is Gaussian with mean 0 and covariance operator $(s\wedge t)D$. Therefore

$$\mathbb{E}\big[\langle X(s\wedge t), e_k\rangle\langle X(s\wedge t), e_l\rangle\big] = \int \langle x, e_k\rangle\langle x, e_l\rangle\, \mu_{X_{s\wedge t}}(dx) = \langle (s\wedge t)D\, e_k, e_l\rangle$$

and

$$\begin{aligned}\mathbb{E}\Big[\frac{\beta_k(s)\,\beta_l(t)}{\sqrt{\lambda_k\lambda_l}}\Big] &= \frac{1}{\sqrt{\lambda_k\lambda_l}}(s\wedge t)\langle De_k, e_l\rangle \\ &= \frac{1}{\sqrt{\lambda_k\lambda_l}}(s\wedge t)\,\lambda_k\delta_{kl} = (s\wedge t)\,\delta_{kl},\end{aligned}$$

where δ_{kl} equals one if $k=l$ and zero otherwise. We conclude that the β_k's are independent Brownian motions[4].

Furthermore,

$$\mathbb{E}(\langle X(s), e_k\rangle\langle X(s), e_l\rangle) = s\langle De_k, e_l\rangle = s\lambda_k\delta_{kl},$$

yields

$$\mathbb{E}\Big|\Big|\sum_{k=n}^{m}\beta_k(s)e_k\Big|\Big|^2 = s\sum_{k=n}^{m}\lambda_k$$

and, because D is a trace-class operator, $\sum_k \lambda_k < \infty$. So the $\beta_k(s)e_k$ form a Cauchy-sequence and (1.3) converges in $L^2(\Omega, \mathcal{A}, \mathbb{P})$.

■

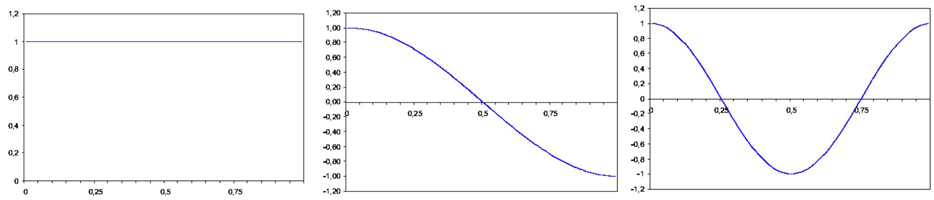
The above result also yields a possibility to construct a Wiener process from a series of independent Brownian motions $(W_k)_{t\geq 0}$: For any orthonormal basis $\{e_k : k\in\mathbb{N}\}$ and positive λ_k such that $\sum_k \lambda_k < \infty$, (1.2) determines a covariance operator, say D. Then a D-Wiener process is obtained by setting

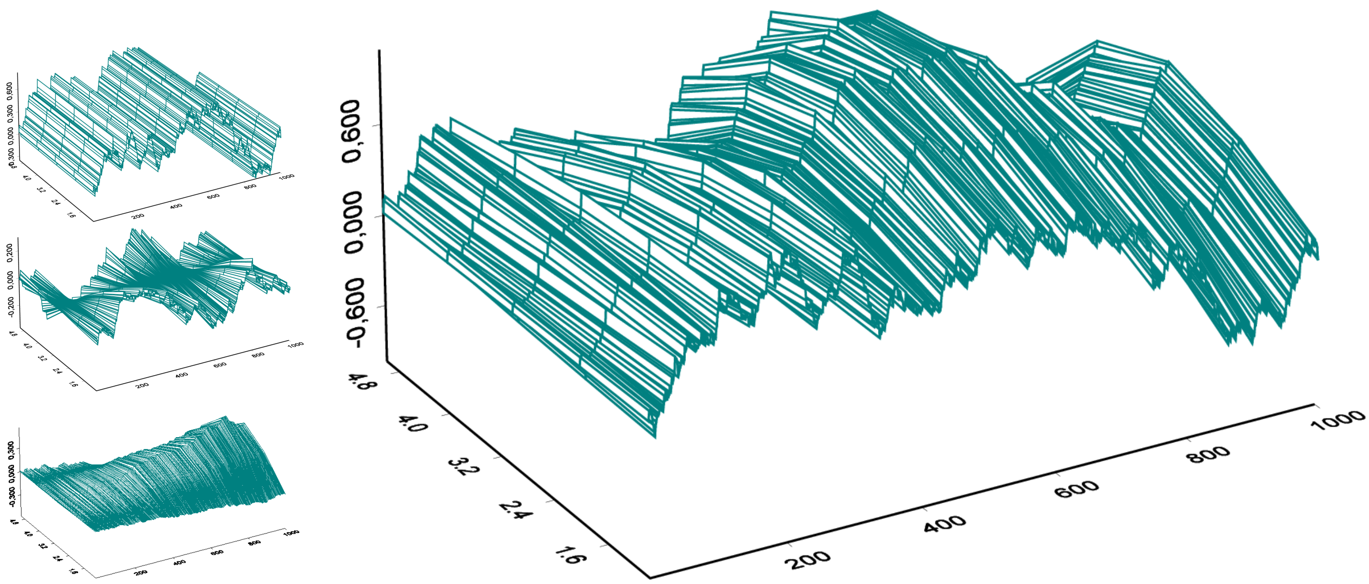
$$X(s) := \sum_{k=1}^{\infty}\sqrt{\lambda_k}W_k(s)e_k.$$

Because of this, X is often called "infinite dimensional Brownian motion".

[4]See, for example, Remark 2.9.2 in Karatzas and Shreve (1988).

A possible basis would be: $e_k = \cos(kx)$ on $[0, \pi]$ (Considering $H = L^2[0, \pi]$):





Chapter 2

SDEs on Hilbert Spaces

For technical reasons we always consider a finite time horizon $T^* \in \mathbb{R}$, i.e., investigate the process $(f(s))_{s\in[0,T^*]}$.

2.1 The Stochastic Integral

In this section we aim to define the stochastic integral with respect to a D-Wiener process on a Hilbert Space H. In the case where $H = \mathbb{R}^n$, the integral w.r.t. an n-dimensional Brownian motion $(B(s))_{s\geq 0}$ with covariance Matrix $\boldsymbol{\Sigma}$ (see page 7) of a $\mathbb{R}^n \times \mathbb{R}^n$- valued stochastic process $(\boldsymbol{\sigma}(s))_{s\geq 0}$ is well known and denoted by

$$\int\limits_0^t \boldsymbol{\sigma}(s)\, dB_s.$$

Note that the matrix $\boldsymbol{\sigma}(s)$ is a linear mapping $\mathbb{R}^n \mapsto \mathbb{R}^n$, operating on $B(s)$.

Imitating this, we consider integrands which take values in the space of linear functions from $H \to H$. As before, for $\Phi \in L(H)$ and $f \in H$ we write $\Phi \cdot f$ for $\Phi(f)$.

Definition 2.1.1. Consider the Hilbert Space H and a process $(\Phi(s))_{s\in[0,T^*]}$, which takes values in $L(H)$. $\Phi(s)$ is called **elementary**, if there exist $0 = t_0 < t_1 < \cdots < t_n = T^*$ and $\Phi_k \in L(H)$, measurable with respect to $\mathcal{F}_{t_k}$, such that $\Phi(0) = 0$ and

$$\Phi(t) = \Phi_k \qquad \text{for } t \in (t_k, t_{k+1}], k = 0, \dots, n-1.$$

In this case we define the stochastic integral for $t \in [0, T^*]$ by

$$\int\limits_0^t \Phi(s) \cdot dX(s) := \sum_{k=0}^{n-1} \Phi_k \cdot \big(X(t_{k+1} \wedge t) - X(t_k \wedge t)\big).$$

Note that the stochastic integral is itself a stochastic process which has values in H. Stochastic integrals prove to be a powerful concept to describe the behavior of martingales.

Denote the normon H by $\| \cdot \|$.

A stochastic process $(X(s))_{s\in[0,T^*]}$ with $E\| X(s)\| < \infty$ for all $s \in [0, T^*]$ is called a *martingale* w.r.t. the filtration $(\mathcal{F}_s)_{s\in[0,T^*]}$, iff

$$\mathbb{E}(X_t|\mathcal{F}_s) = X_s, \qquad \text{for all } 0 \le s < t \le T^*.$$

Usually one considers filtrations of the type $\mathcal{F}_s = \sigma(X_t : 0 \le t \le s)$. This filtration is called the *canonical filtration* of X.

As a consequence of its independent increments, a D-Wiener process $(X(s))_{s\in[0,T^*]}$ w.r.t. to its canonical filtration is a martingale. This leads to the question, under which circumstances this property is inherited by the stochastic integral. The following proposition considers such a case. We first need a small lemma.

Lemma 2.1.2. *Consider the linear operator $\Phi \in L(U, H)$ and a normally distributed random variable ξ with mean zero and covariance operator D. Then*

$$\mathbb{E}\Big(|\Phi \cdot \xi|^2\Big) = \operatorname{tr}\Big[\Big(\Phi D^{\frac{1}{2}}\Big)\Big(\Phi D^{\frac{1}{2}}\Big)^*\Big].$$

Proof. For an expansion of the type $\xi = \sum_k e_k\langle\xi, e_k\rangle$ we would need continuity of Φ. On the other hand, we can directly consider an expansion of $\Phi \cdot \xi$ itself. First, note that for $f \in H$

$$|f|^2 = |\sum_{k=1}^{\infty} e_k\langle f, e_k\rangle|^2 = \Big\langle \sum_{k=1}^{\infty} e_k\langle f, e_k\rangle, \sum_{j=1}^{\infty} e_j\langle f, e_j\rangle \Big\rangle = \sum_{k=1}^{\infty}\langle f, e_k\rangle^2.$$

Applying this to $\Phi \cdot \xi$ yields

$$\mathbb{E}|\Phi \cdot \xi|^2 = \mathbb{E}\sum_{k=1}^{\infty}\langle \Phi \cdot \xi, e_k\rangle^2 = \sum_{k=1}^{\infty}\mathbb{E}\langle \xi, \Phi^* \cdot e_k\rangle\langle \xi, \Phi^* \cdot e_k\rangle,$$

where the last equality follows by dominated convergence. As ξ is normally distributed, we obtain directly that

$$\mathbb{E}|\Phi \cdot \xi|^2 = \sum_{k=1}^{\infty}\langle D\Phi^* \cdot e_k, \Phi^* \cdot e_k\rangle.$$

Basically we have to calculate the trace of $\Phi D\Phi^*$. Recall that D is trace class (so it is compact) and therefore $D = D^{\frac{1}{2}}D^{\frac{1}{2}}$, where $D^{\frac{1}{2}}$ is self-adjoint. Thus, also $\Phi D\Phi^* = (\Phi D^{\frac{1}{2}})(\Phi D^{\frac{1}{2}})^*$ is self-adjoint. This leads to

$$\mathbb{E}|\Phi \cdot \xi|^2 = \sum_{k=1}^{\infty}\Big\langle \big(\Phi D^{\frac{1}{2}}\big)\big(\Phi D^{\frac{1}{2}}\big)^* e_k, e_k\Big\rangle = \operatorname{tr}\Big[\Big(\Phi D^{\frac{1}{2}}\Big)\Big(\Phi D^{\frac{1}{2}}\Big)^*\Big].$$

At a later point we identify the trace as (square of) the Hilbert-Schmidt norm of $\Phi D^{\frac{1}{2}}$. ■

Proposition 2.1.3. *For an elementary stochastic process $(\Phi(s))_{s\in[0,T^*]}$ with values in $L(H)$ and $T^* \in \mathbb{R}$ we have*[1]

$$|||\Phi|||^2_{T^*} := \mathbb{E}\Big(\| \int_0^{T^*} \Phi(s) \cdot dX(s)\|^2\Big) = \mathbb{E}\Big(\int_0^{T^*} \mathrm{tr}[(\Phi(s)D^{\frac{1}{2}})(\Phi(s)D^{\frac{1}{2}})^*]\, ds\Big).$$

Furthermore, if $|||\Phi|||_{T^*} < \infty$, then the stochastic integral

$$\int_0^t \Phi(s) \cdot dX(s)$$

is a square-integrable martingale for all $t \leq T^*$.

Proof. Enhance the partition by $t =: t_m$. Setting $\Delta_k X = X(t_{k+1}) - X(t_k)$ we obtain

$$\begin{aligned}
\mathbb{E}\| \int_0^t \Phi(s)\, dX(s)\|^2 &= \mathbb{E}\| \sum_{k=0}^{m-1} \Phi_k \cdot \Delta_k X\|^2 \\
&= \mathbb{E}\Big(\sum_{k=0}^{m-1} \| \Phi_k \cdot \Delta_k X\|^2\Big) + \mathbb{E}\Big(2 \sum_{i,j=0,i<j}^{m-1} \langle \Phi_i \cdot \Delta_i X, \Phi_j \cdot \Delta_j X\rangle\Big) \\
&=: (1) + (2).
\end{aligned}$$

First, we directly have

$$\begin{aligned}
(2) &= 2 \sum_{i,j=1,i<j}^{m-1} \mathbb{E}\Big(\langle \Phi_i \cdot \Delta_i X, \Phi_j \cdot \Delta_j X\rangle\Big) \\
&= 2 \sum_{i,j=1,i<j}^{m-1} \mathbb{E}\Big[\mathbb{E}\Big(\langle \Phi_i \cdot \Delta_i X, \Phi_j \cdot \Delta_j X\rangle \Big| \mathcal{F}_{t_{i+1}}\Big)\Big] = 0,
\end{aligned}$$

because X has independent increments and zero mean.

On the other side, we can apply Lemma 2.1.2 for (1). Using iterated expectations we find

$$\begin{aligned}
(1) &= \sum_{k=0}^{m-1} \mathbb{E}\Big(|\Phi_k \cdot \Delta_k X|^2\Big) = \sum_{k=0}^{m-1} (t_{k+1} - t_k)\mathbb{E}\Big(\mathrm{tr}\Big[\Big(\Phi_k D^{\frac{1}{2}}\Big)\Big(\Phi_k D^{\frac{1}{2}}\Big)^*\Big]\Big) \\
&= \mathbb{E}\Big(\int_0^t \mathrm{tr}\, \big[(\Phi(s)D^{\frac{1}{2}})(\Phi(s)D^{\frac{1}{2}})^*\big]\, ds\Big).
\end{aligned}$$

[1] Using positivity and the eigenvalue expansion of D, we define for compact and linear D $D^{\frac{1}{2}}(x) := \sum_k \sqrt{\lambda_k}\langle x, e_k\rangle e_k$, see Werner (2000, p. 244). Furthermore, for $T \in L(H)$ we denote its Hilbert space adjoint by T^*, see Werner (2000, p. 208). That is, $a, b \in H$ yields $\langle Ta, b\rangle = \langle a, T^* b\rangle$. Note that $D^{\frac{1}{2}}$ is self-adjoint, i.e. $(D^{\frac{1}{2}})^* = D^{\frac{1}{2}}$.

For the martingale property we enhance the partition further by $s =: t_{\tilde{m}} < t$. Then

$$\begin{aligned}\mathbb{E}\Big[\int_0^t \Phi(u)\,dX(u)\Big|\mathcal{F}_s\Big] &= \int_0^s \Phi(u)\,dX(u) + \mathbb{E}\Big[\sum_{j=\tilde{m}}^{m-1} \Phi_j \cdot \Delta_j X\Big|\mathcal{F}_s\Big] \\ &= \int_0^s \Phi(u)\,dX(u),\end{aligned}$$

again, because of independent increments and zero means of $(X(s))_{s\in[0,T^*]}$. ■

As a next step we want to extend the stochastic integral to more general functions Φ. Therefore we look for a class of processes which can be approximated by elementary functions, such that at the same time the martingale property of the integral is preserved. It turns out that the proper class is formed by certain Hilbert-Schmidt operators.

First, consider the space $H_0 := D^{\frac{1}{2}}(H)$, which, endowed with the inner product[2]

$$\langle u,v\rangle_0 := \sum_k \frac{1}{\lambda_k}\langle u,e_k\rangle\langle v,e_k\rangle = \langle D^{-\frac{1}{2}}u, D^{-\frac{1}{2}}v\rangle$$

is a Hilbert space. For an orthonormal basis $\{e_k : k\in\mathbb{N}\}$ of H, setting $e_k^0 := D^{\frac{1}{2}}e_k$ yields an orthonormal basis $\{e_k^0 : k\in\mathbb{N}\}$ of H_0. This representation can be simplified a bit if we choose a clever basis. Denote the eigenvectors and eigenvalues by $\{\tilde{e}_k : k\in\mathbb{N}\}$ and $\{\tilde{e}_k : k\in\mathbb{N}\}$, respectively, we have $e_k^0 = \sqrt{\lambda_k}\tilde{e}_k$.

Then, denote by $L_2(H_0,H)$ the space of all *Hilbert-Schmidt* operators from H_0 into H, that is, linear operators T with

$$\sum_{k=1}^{\infty} \|\,Te_k^0\|^2 \quad < \infty$$

for the orthonormal basis $\{e_k^0 : k\in\mathbb{N}\}$ of H^0 and the norm $\|\cdot\|$ of H. Note that the inner product

$$\langle S,T\rangle_2 := \sum_{k=1}^{\infty}\langle Se_k^0, Te_k^0\rangle$$

induces the Hilbert-Schmidt norm

$$\|\,T\|_2^2 := \sum_{k=1}^{\infty}\langle Te_k^0, Te_k^0\rangle = \sum_{k=1}^{\infty}\lambda_k\|\,Te_k\|^2$$

and $L_2(H_0,H)$ is again a Hilbert space. Also, $\|\,T\|_2^2 = \operatorname{tr}\big(T^*T\big)$, i.e. T is Hilbert-Schmidt, iff T^*T is trace-class. The norm and the scalar product is independent of the chosen basis. Then, $\|\,T\|_2^2 = \sum_{k=1}^{\infty} s_k^2$, where $s_1 \le s_2 \le \dots$ are the eigenvalues of T appearing as often as their multiplicity, compare Proposition 1.2.12.

[2]Similar to $D^{\frac{1}{2}}$, we define $D^{-\frac{1}{2}}(x) := \sum_{k=1}^{\infty}\frac{1}{\sqrt{\lambda_k}}\langle x,e_k\rangle e_k 1_{\{\lambda_k>0\}}$.

With the above notations we define for $T \in \mathbb{R}$ the norm $|||\Phi|||_T$ also for non-elementary processes with values in $L_2(H_0, H)$ by

$$|||\Phi|||_T \quad := \quad \Big[\mathbb{E}\int_0^T \|\,\Phi(s)\|_2^2\,ds\Big]^{\frac{1}{2}} = \Big[\mathbb{E}\int_0^T \sum_{k=1}^{\infty}\langle\Phi(s)e_k^0, \Phi(s)e_k^0\rangle\,ds\Big]^{\frac{1}{2}}.$$

Because $e_k^0 = D^{\frac{1}{2}}e_k$ we obtain

$$\begin{aligned} |||\Phi|||_T \quad &= \quad \Big[\mathbb{E}\int_0^T \sum_{k=1}^{\infty}\langle\Phi(s)D^{\frac{1}{2}}e_k, \Phi(s)D^{\frac{1}{2}}e_k\rangle\,ds\Big]^{\frac{1}{2}} \\ &= \quad \Big[\mathbb{E}\int_0^T \sum_{k=1}^{\infty}\langle\Phi(s)D^{\frac{1}{2}}\big(\Phi(s)D^{\frac{1}{2}}\big)^* e_k, e_k\rangle\,ds\Big]^{\frac{1}{2}} \\ &= \quad \Big[\mathbb{E}\int_0^T \operatorname{tr}\big((\Phi(s)D^{\frac{1}{2}})(\Phi(s)D^{\frac{1}{2}})^*\big)\,ds\Big]^{\frac{1}{2}} \\ &= \quad \Big[\mathbb{E}\int_0^T \|\,\Phi(s)D^{\frac{1}{2}}\|_2^2\,ds\Big]^{\frac{1}{2}}. \end{aligned}$$

We then have the following

Lemma 2.1.4. *For a predictable process $(\Phi(s))_{s\in[0,T^*]}$ with values in $L_2(H_0, H)$ and $|||\Phi|||_{T^*} < \infty$, there exists a sequence of elementary processes Φ_n, such that*

$$|||\Phi - \Phi_n|||_{T^*} \longrightarrow 0 \qquad \textit{as } n \longrightarrow \infty.$$

Proof. See Da Prato and Zabczyk (1992), Lemma 4.7. Essentially, the proof is similar to the well-known approximation in the real-valued case. ■

Now we define for predictable $(\Phi(s))_{s\in[0,T^*]}$ with values in $L_2(H_0, H)$ and $|||\Phi|||_{T^*} < \infty$

$$\int_0^t \Phi(s)\cdot dX(s) := \lim_{n\to\infty}\int_0^t \Phi_n(s)\cdot dX(s). \tag{2.1}$$

It can be shown that this stochastic integral is well-defined and furthermore a martingale if $|||\Phi|||_{T^*} < \infty$ (see Proposition 2.1.3).

So we finally found the class of suitable integrands which ensure that the stochastic integrals inherit the martingale property, namely predictable processes with values in $L_2(H_0, H)$ which satisfy $|||\Phi|||_{T^*} < \infty$.

Finally the stochastic integral may be, by use of a localization procedure, extended to stochastically integrable processes, i.e., processes, for which

$$\mathbb{P}\Big(\int_0^{T^*} \|\,\Phi(s)\|_2^2\,ds < \infty\Big) = 1.$$

Note that, in doing so, the martingale property is lost, but stochastic integrals still remain local martingales. For a full treatment see Da Prato and Zabczyk (1992, p. 94 p.p.).

2.2 Covariances of Stochastic Integrals

In this section we consider covariances of the previously defined stochastic integrals. The following definition is in analogy to Definition 1.2.17.

Definition 2.2.1. For two H-valued random variables X_i with mean m_i, $i = 1, 2$, the symmetric operator $D \in L(H)$, such that

$$\mathbb{E}\big[\langle X_1, f\rangle\langle X_2, g\rangle\big] - \langle m_1, f\rangle\langle m_2, g\rangle = \langle Df, g\rangle$$

is called the covariance of X_1 and X_2 and denoted by $\mathrm{Cov}(X_1, X_2)$.

First, we need a small lemma giving the connections between Hilbert-Schmidt and trace-class operators.

Lemma 2.2.2. *Consider separable Hilbert spaces H_1, H_2. If $S \in L_2(H_2, H_1)$ and $T \in L_2(H_1, H_2)$, then the product ST is in $L_1(H_1, H_1)$ and*

$$\| ST\|_1 \leq \| S\|_2 \| T\|_2.$$

Proof. By definition we have for a basis $\{e_k : k \in \mathbb{N}\}$ of H_1 that

$$\begin{aligned}
\| ST\|_1 &= \sum_k \Big|\langle STe_k, e_k\rangle\Big| = \sum_k \Big|\langle Te_k, S^* e_k\rangle\Big| \\
&\leq \sum_k \| Te_k\| \, \| S^* e_k\| \leq \Big(\sum_k \| Te_k\|^2\Big)^{\frac{1}{2}} \Big(\sum_k \| S^* e_k\|^2\Big)^{\frac{1}{2}} \\
&= \| T\|_2 \| S^*\|_2 = \| T\|_2 \| S\|_2.
\end{aligned}$$

■

We obtain

Proposition 2.2.3. *Assume $(\Phi_1(s))_{s\in[0,T^*]}, (\Phi_2(s))_{s\in[0,T^*]}$ are predictable processes with values in $L_2(H_0, H)$, $|||\Phi_1|||_{T^*} < \infty$ and $|||\Phi_2|||_{T^*} < \infty$. Then, for all $t \in [0, T^*]$*

$$\mathbb{E}\int_0^t \Phi_i(s) \cdot dX(s) = 0, \qquad \mathbb{E}\| \int_0^t \Phi_i(s) \cdot dX(s)\|^2 < \infty, \qquad i = 1, 2$$

and for all $t, s \in [0, T^*]$ the covariance operator equals

$$\mathrm{Cov}\Big(\int_0^t \Phi_1(u) \cdot dX(u), \int_0^s \Phi_2(v) \cdot dX(v)\Big) \;=\; \mathbb{E}\int_0^{t\wedge s} \big(\Phi_2(u) D^{\frac{1}{2}}\big) \big(\Phi_1(u) D^{\frac{1}{2}}\big)^* \, du. \tag{2.2}$$

Furthermore,

$$\mathbb{E}\langle \int_0^t \Phi_1(u) \cdot dX(u), \int_0^s \Phi_2(u) \cdot dX(u)\rangle = \mathbb{E}\int_0^{t\wedge s} \mathrm{tr}[(\Phi_2(u) D^{\frac{1}{2}})(\Phi_1(u) D^{\frac{1}{2}})^*] \, du. \tag{2.3}$$

Proof. As $(\Phi_2(u)D^{\frac{1}{2}})$ and $(\Phi_1(u)D^{\frac{1}{2}})^*$ are $L_2(H)$ valued processes, the process $(\Phi_2(u)D^{\frac{1}{2}})(\Phi_1(u)D^{\frac{1}{2}})^*$ takes values[3] in $L_1(H)$. We first show the inequality

$$\mathbb{E}\int_0^{T^*} \| (\Phi_2(u)D^{\frac{1}{2}})(\Phi_1(u)D^{\frac{1}{2}})^* \|_1 \, du \leq |||\Phi_1|||_{T^*} |||\Phi_2|||_{T^*}. \tag{2.4}$$

Observe that, by the preceding lemma we have that

$$\| (\Phi_2(u)D^{\frac{1}{2}})(\Phi_1(u)D^{\frac{1}{2}})^* \|_1 \leq \| \Phi_2(u)D^{\frac{1}{2}} \|_2 \, \| \Phi_1(u)D^{\frac{1}{2}} \|_2.$$

Applying Cauchy-Schwartz several times yields immediately (2.4) and existence of the integral in (2.2) is ensured. Note that $||T||_2 = \mathrm{tr}(T^*T)$ and $\Phi D \Phi$ is self-adjoint.

Further on, consider elementary processes Φ_1 and Φ_2. We proceed similarly to the proof of Proposition 2.1.3. Assume w.l.o.g. that $s \leq t$ and enhance the partition by t and s at the points t_m and $t_{\tilde{m}}$, say. Then

$$\begin{aligned} &\mathbb{E}\Big(\langle \int_0^t \Phi_1(u)\, dX(u), a\rangle \langle \int_0^s \Phi_2(v) dX(v), b\rangle\Big) \qquad (2.5)\\ &\quad = \sum_{i,j=0, i\neq j}^{\tilde{m}} \mathbb{E}\Big(\langle \Phi_1(t_i)\Delta_i X, a\rangle \, \langle \Phi_2(t_j)\Delta_j X, b\rangle\Big) \\ &\qquad + \sum_{i=0}^{\tilde{m}} \sum_{j=\tilde{m}+1}^{m} \mathbb{E}\Big(\langle \Phi_1(t_i)\Delta_i X, a\rangle \, \langle \Phi_2(t_j)\Delta_j X, b\rangle\Big) \\ &\qquad + \sum_{i=0}^{\tilde{m}} \mathbb{E}\Big(\langle \Phi_1(t_i)\Delta_i X, a\rangle \, \langle \Phi_2(t_i)\Delta_i X, b\rangle\Big). \end{aligned}$$

As for $i \neq j$

$$\mathbb{E}\Big(\langle \Phi_1(t_i)\Delta_i X, a\rangle \, \langle \Phi_2(t_j)\Delta_j X, b\rangle\Big) = 0$$

the first two sums vanish. Furthermore, $\Delta_i X \sim \mathcal{N}(0, (t_{i+1} - t_i)D)$ and again using iterated expectations we obtain

$$\begin{aligned} \mathbb{E}\Big(\langle \Phi_1(t_i)\Delta_i X, a\rangle \, \langle \Phi_2(t_i)\Delta_i X, b\rangle\Big) &= \mathbb{E}\Big(\langle \Delta_i X, \Phi_1(t_i)^* a\rangle \, \langle \Delta_i X, \Phi_2(t_i)^* b\rangle\Big) \\ &= (t_{i-1} - t_i)\, \mathbb{E}\Big(\langle D\Phi_1(t_i)^* a, \Phi_2(t_i)^* b\rangle\Big) \\ &= (t_{i-1} - t_i)\, \mathbb{E}\Big(\Big\langle \big(\Phi_2(t_i)D^{\frac{1}{2}}\big)\big(D^{\frac{1}{2}}\Phi_1(t_i)\big)^* a, b\Big\rangle\Big) \qquad (2.6) \end{aligned}$$

As we consider only finite sums this directly yields

$$(2.5) \;=\; \mathbb{E}\langle \int_0^{t\wedge s} (\Phi_2(u)D^{1/2})(\Phi_1(u)D^{1/2})^* \, du \, a, b\rangle.$$

[3]Here, $L_1(H)$ is the Banach space of all trace-class operators in $L(H)$, see Page 7.

So the conclusion holds for elementary processes. With the bound (2.4) the general conclusion follows from an appropriate approximation through elementary processes.

Observe that equation (2.6) yields

$$(2.5) = \sum_{i=0}^{\tilde{m}} \mathbb{E}\Big[\langle (\Phi_2(t_i)D^{\frac{1}{2}})(\Phi_1(t_i)D^{\frac{1}{2}})^* a \,, b\rangle\Big](t_{i+1} - t_i)$$

and (2.3) follows immediately. ■

2.3 Itô's formula

The formula of Itô (1946) yields the chain rule for functions of diffusion processes. In comparison to the fundamental theorem of calculus there appears an unexpected second term. As the formula mainly relies on the Taylor formula this is a result of the non vanishing second-order term and leads to interesting probabilistic interpretations. The reason for its appearance is due to infinite variation of the Brownian motion. Interestingly, there is a close analogue to processes in Hilbert spaces which is derived in this chapter.

We only cite the Taylor formula for Hilbert spaces. A detailed treatment may be found in Dieudonné (1969). Consider a Hilbert space H and an open subset $A \subset H$. It may be recalled that, if the derivative of a continuous mapping $f : A \mapsto H$ denoted by Df exists, it is a continuous and linear mapping form H into H and therefore an element of the Banach space $L(H)$.

Furthermore, if the second derivative $D^2 f$ exists, it is an element of $L(H; L(H))$ and a symmetric[4] mapping. The space $L(H; L(H))$ can be identified[5] with the space of continuous bilinear mappings of $H \times H$ into H, denoted by $L(H, H; H)$.

As a result of the mean value theorem we obtain **Taylor's formula**:

Theorem 2.3.1. *Assume f is a twice continuously differentiable mapping of A into H. If $x + \theta t \in A$ for $x, t \in H$ and all $\theta \in [0, 1]$, we have*

$$f(x + t) = f(x) + Df(x) \cdot t + \frac{1}{2} D^2 f(x + \zeta t) \cdot (t, t),$$

where ζ is an element of $[0, 1]$.

For the Itô-formula on Hilbert spaces we consider a D-Wiener process $(X(s))_{s\geq 0}$ on H and a predictable process $(\Phi(s))_{s\geq 0}$ with values in $L_2(H_0, H)$, such that $|||\Phi|||_{T^*} < \infty$. Then the stochastic process

$$S(t) = S(0) + \int_0^t \Phi(s) \cdot dX(s) \tag{2.7}$$

is a square-integrable martingale, as already mentioned in the previous section.

[4] $D^2 f$ is symmetric in the sense that $D^2 f \cdot (f, g) = D^2 f \cdot (g, f)$.

[5] By $h \cdot (s, t) \simeq (h \cdot s) \cdot t$.

Theorem 2.3.2. *For an open subset A of the Hilbert space H, let $f : A \mapsto H$ be a function, whose first and second derivative is uniformly continuous on bounded subsets of A. For $(S(t))_{t\in[0,T^*]}$, as in (2.7) we have for all $t \in [0,T^*]$ $\mathbb{P}$-a.s.*

$$f(S(t)) = f(S(0)) + \int_0^t Df(S(u)) \cdot dS(u) + \frac{1}{2}\int_0^t \sum_{k=1}^{\infty} \lambda_k D^2 f(S(u)) \cdot (\Phi(u)\cdot e_k, \Phi(u)\cdot e_k)\, du.$$

Note that the first integral equals

$$\int_0^t Df(S(u)) \cdot \Phi(u) \ \cdot dX(u).$$

Proof. By a localization procedure we can restrict ourselves to bounded $(X(s))_{s\in[0,T^*]}$ and $(\Phi(s))_{s\in[0,T^*]}$, see Da Prato and Zabczyk (1992, p. 106).

Further on, consider a partition $\Pi = \{t_0, t_1, \dots, t_n\}$ of $[0,t]$ with $0 = t_0 < t_1 < \cdots < t_n$ and denote its mesh by $\|\Pi\| := \max_{1\le i\le n}(t_i - t_{i-1})$.

Using the Taylor formula on Banach spaces and writing $(t)^{(2)}$ for (t,t), we obtain

$$\begin{aligned} f(S_t) - f(S_0) &= \sum_{j=0}^{n-1} f(S(t_{j+1})) - f(S(t_j)) \\ &= \sum_{j=0}^{n-1} Df(S(t_j)) \cdot (S(t_{j+1}) - S(t_j)) + \frac{1}{2} D^2 f(\tilde{S}_j) \cdot (S(t_{j+1}) - S(t_j))^{(2)} \\ &= \sum_{j=0}^{n-1} Df(S(t_j)) \cdot \Delta_j S + \frac{1}{2} D^2 f(S(t_j)) \cdot (\Delta_j S)^{(2)} \\ &\quad + \frac{1}{2}\big[D^2 f(\tilde{S}_j) - D^2 f(S(t_j))\big] \cdot (\Delta_j S)^{(2)} \\ &= I + II + III, \end{aligned}$$

where we set $\tilde{S}_j := S(t_j) + \zeta_j(S(t_{j+1}) - S(t_j))$ with $\zeta_j = \zeta_j(\omega) \in [0,1]$.

Considering I, we intend to approximate $Y_s := Df(S(s))$ by the elementary process

$$Y_s^n := Df(S(0))1_{\{0\}}(s) + \sum_{j=0}^{n-1} Df(S(t_j))1_{(t_j,t_{j+1}]}(s).$$

Indeed, uniform continuity of the derivative and the bounded convergence theorem yield

$$|||Y - Y^n|||_{T^*} = \mathbb{E}\int_0^{T^*} \|Y(u) - Y^n(u)\|_2^2\, du \longrightarrow 0$$

as the mesh of the partition tends to zero.

Then, by Definition (2.1), we have $\mathbb{P}$-a.s. for $\|\Pi\| \to 0$,

$$I \longrightarrow \int_0^t Df(S(s)) \cdot dS(s).$$

The third summand, III, converges to 0 $\mathbb{P}$-a.s. for $\|\Pi\| \to 0$, because of continuity of the derivative, using a similar argument.

Consider the second term, II. We calculate the conditional expectation of the summands

$$\begin{aligned} &\mathbb{E}\Big(D^2 f(S(t_j)) \cdot (\Delta_j S)^{(2)} \Big| \mathcal{F}_{t_j}\Big) \\ &\quad = \mathbb{E}\Big(D^2 f(S(t_j)) \cdot (\Phi(t_j) \cdot \Delta_j X)^{(2)} \Big| \mathcal{F}_{t_j}\Big). \end{aligned} \tag{2.8}$$

Using the eigenvalue expansion of X, we obtain

$$\begin{aligned} (2.8) &= \sum_{k,l=1}^{\infty} \mathbb{E}\Big(\Delta_j \beta_k \, \Delta_j \beta_l \Big| \mathcal{F}_{t_j}\Big) D^2 f(S(t_j)) \cdot (\Phi(t_j) \cdot e_k, \Phi(t_j) \cdot e_l) \\ &= \sum_{k=1}^{\infty} (t_{j+1} - t_j) \lambda_k \, D^2 f(S(t_j)) \cdot (\Phi(t_j) \cdot e_k, \Phi(t_j) \cdot e_k) \\ &=: (t_{j+1} - t_j) \, J(t_j). \end{aligned} \tag{2.9}$$

To show L^2-convergence it suffices to prove that the following expectation converges to zero:

$$\begin{aligned} &\mathbb{E}\Big[\sum_{j=0}^{n-1} \Big(D^2 f(S(t_j))(\Delta_j S)^{(2)} - (t_{j+1} - t_j) \, J(t_j)\Big)\Big]^2 \\ &\quad = \sum_{j=0}^{n-1} \Big(\mathbb{E}\big[D^2 f(S(t_j))(\Delta_j S)^{(2)}\big]^2 - (t_{j+1} - t_j)^2 \mathbb{E}\big[J(t_j)\big]^2\Big). \end{aligned} \tag{2.10}$$

In the last step we used the fact that the summands are independent from each other and, due to (2.9), have zero mean.

If we expand the first summand via (1.3) and denote $D^2 f(S(t_j))(\Phi(t_j) e_k, \Phi(t_j) e_l) =: \xi^j_{k,l}$ we get

$$\begin{aligned} &\sum_{j=0}^{n-1} \mathbb{E}\Big[D^2 f(S(t_j))(\Delta_j S)^{(2)}\Big]^2 \\ &\quad = \sum_{j=0}^{n-1} \mathbb{E}\Big\{\sum_{k,l,m,n=1}^{\infty} \mathbb{E}\Big[\Delta_j \beta_k \, \Delta_j \beta_l \, \Delta_j \beta_m \, \Delta_j \beta_n \Big| \mathcal{F}_{t_j}\Big] \xi^j_{k,l} \, \xi^j_{m,n}\Big\} \\ &\quad = \sum_{j=0}^{n-1} \sum_{k,m=1}^{\infty} \lambda_k \lambda_m (t_{j+1} - t_j)^2 \mathbb{E}(\xi^j_{k,k} \xi^j_{m,m}) + \sum_{j=0}^{n-1} \sum_{k=1}^{\infty} \lambda_k^3 (t_{j+1} - t_j)^2 \mathbb{E}\big((\xi^j_{k,k})^2\big). \end{aligned}$$

Second moments of $\xi^j_{k,l}$ are bounded for any j,k,l because $D^2 f$ itself is bounded by assumption. So the last sum converges to zero as $\|\Pi\| \to 0$. For the second summand of (2.10) we conclude

$$\sum_{j=0}^{n-1}(t_{j+1}-t_j)^2\mathbb{E}\Big[\sum_{k=1}^{\infty}\lambda_k\, D^2 f(S(t_j))\cdot(\Phi(t_j)\cdot e_k,\Phi(t_j)\cdot e_k)\Big]^2$$
$$= \sum_{j=0}^{n-1}(t_{j+1}-t_j)^2\mathbb{E}\Big[\sum_{k=1}^{\infty}\lambda_k\xi^j_{k,k}\Big]^2$$
$$= \sum_{j=0}^{n-1}(t_{j+1}-t_j)^2\sum_{k,m=1}^{\infty}\lambda_k\lambda_m\mathbb{E}\big[\xi^j_{k,k}\xi^j_{m,m}\big],$$

which also converges to zero as $\sup_j(t_{j+1}-t_j)\to 0$.

Up to now we obtained convergence in L^2. Considering a subsequence of $\{\Pi^{(n)}\}_{n=1}^{\infty}$ yields the desired $\mathbb{P}$-a.s. convergence, c.f. Karatzas and Shreve (1988, p.152). ■

The Itô-formula can be extended to processes of the type

$$S(t) = S(0) + \int_0^t \mu(s)\,ds + \int_0^t \Phi(s)\,dX(s),$$

where $(\mu(s))_{s\in[0,T^*]}$ is an adapted, H-valued process. Also the function f might be time-dependent, see Da Prato and Zabczyk (1992, p. 105-108)

$$f(t,S(t)) = f(0,S(0)) + \int_0^t \frac{\partial}{\partial t}f(u,S(u))\,ds + \int_0^t Df(u,S(u))\cdot dS(u)$$
$$+\frac{1}{2}\int_0^t\sum_{k=1}^{\infty}\lambda_k D^2 f(S(u))\cdot(\Phi(u)\cdot e_k,\Phi(u)\cdot e_k)\,du.$$

2.3.1 The Heath-Jarrow-Morton Model

In this section we want to apply the previously developed techniques to the famous approach of Heath, Jarrow, and Morton (1992) (henceforth HJM). The HJM model is a interest rate model. Primary assets in the considered interest rate market are bonds. A bond is a contract, which promises the payoff 1 at a future time-point, the maturity T. Considering an equivalent martingale measure Q, the price of a bond equals

$$B(t,T) = \mathbb{E}^Q\Big[\exp\Big(-\int_t^T r_u\,du\Big)\Big|\mathcal{F}_t\Big], \qquad t\le T.$$

The idea of HJM was to model the forward rates instead of the bond prices. The forward rates are defined by

$$f(t,T) = -\frac{\partial}{\partial T}\ln B(t,T),$$

such that we obtain the valuation formula $B(t,T) = \exp(-\int_t^T f(t,u)\,du)$. HJM proposed a dynamics for the forward rates and obtained a condition on the drift of the forward rates to guarantee that the market is free of arbitrage.

At some point, Musiela (1993) realized the connections to stochastic differential equations on Hilbert spaces. The key observation is that $(f(t,\cdot))_{t\geq 0}$ is a process in a function space. The condition $t \leq T$ poses some kind of difficulties, such that Musiela proposed to consider $f_t(x) := f(t, t+x)$ instead. Now the process takes values in a fixed space and the condition $x \geq 0$ of $x \in [0, T^*]$, respectively, is sufficient. In this section we will examine this model in more detail.

It is very important to distinguish between the bond dynamics in Musiela parametrization and the canonical dynamics. To clarify this, we consider the dynamics of the bond price in both representation. Define $B_t(T-t) := B(t,T)$, where the first expression is the bond price in the Musiela parametrization. Now $dB(t,T)$ is considered as a system of SDEs, each for fixed T. This is not the case if we would like to consider $dB_t(T-t)$, as the second coordinate cannot be fixed for different t. One rather has to think of $B_t(T-t) = \mathcal{S}_{-t}B_t(T)$, where $\mathcal{S}$ is the shift-operator, i.e. $\mathcal{S}_y f(x) = f(x+y)$, and $\mathcal{S}$ acts on a functional space, i.e. a Hilbert space in our case. Applying the Itô-formula yields

$$\begin{aligned} dB(t,T) = d\mathcal{S}_{-t}B_t(T) &= -\frac{\partial}{\partial x}\mathcal{S}_{-t}B_t(T)\,dt + d\mathcal{S}_{-t}B_t(T) \\ &= -\frac{\partial}{\partial x}B(t, T-t)\,dt + dB(t, T-t). \end{aligned}$$

However, this small calculation does not yet reveal the full consequences of this approach. As is intuitively clear, the same affects appears with the forward rates. If we want to include the classical HJM-model as a special case, we need to find a clever formulation for the dynamics of the forward rates in the Musiela parametrization. The following turns out to be precisely the right one.

Consider stochastic processes $\alpha : [0,T^*] \times \Omega \mapsto H$ and $\sigma : [0,T^*] \times \Omega \mapsto L_2(H_0; H)$, both predictable w.r.t. $(\mathcal{F}_t)_{t\geq 0}$, satisfying $\mathbb{P}(\int_0^{T^*} \alpha(s)\,ds < \infty) = 1$ and $|||\sigma|||_{T^*} < \infty$. Further on, assume that $(X(t))_{t\geq 0}$ is a D-Wiener process. Assume the forward rate dynamics to follow

$$r_t = S(t)r_0 + \int_0^t S(t-u)\alpha_u\,du + \int_0^t S(t-u)\sigma_u \cdot dX_u. \tag{2.11}$$

We then get the following

Theorem 2.3.3. *Set* $\sigma_k^*(t,x) := \int_0^x \sigma_k(u,v)\,dv$. *Then all discounted bond prices are martingales iff*

$$\alpha_t(x) = \sum_{k=1}^{\infty} \lambda_k\,\sigma_k^*(t,x)\,\sigma_k(t,x) \qquad \forall t \in [0,T^*], x \in [0,T^{**}]. \tag{2.12}$$

Proof. In the Musiela parameterization, the bond price equals

$$B(t,T) = \exp\Big(-\int_0^{T-t} r_t(v)\,dv\Big).$$

As we want to consider $\int_0^{T-t} r_t(v)\,dv$, we introduce

$$y(t,x) = F(r_t, x) := \int_0^x r_t(v)\,dv,$$

where $F : H \times \mathbb{R}^+ \mapsto H$. Consider x as fixed. Then, as F is linear, we immediately obtain that the Frèchet-derivative of F is F itself. Of course, the second derivative is zero. Thus, applying the Itô-formula (2.3.2) yields

$$F(r_t, x) = F(r_0, x) + \int_0^t DF \cdot \frac{\partial}{\partial x} r_u\,du + \int_0^t DF \cdot \alpha_u\,du + \int_0^t DF \cdot \sigma_u \cdot dX_u.$$

We suppress the dependence of the derivative on x. For example, $DF \cdot \alpha_u = \int_0^x \alpha_u(v)\,dv$. Defining $\Phi : [0, T^*]^2 \times \Omega \mapsto L(H; H)$ by

$$\Phi(u,x) \cdot f := \int_0^x [\sigma_u \cdot f](v)\,dv,$$

and $\alpha^*(t,x) := F(\alpha_t, x) = \int_0^x \alpha_t(v)\,dv$, we obtain the dynamics of $F(r_t, x)$ as

$$dF(r_t, x) = \Big[r_t(x) - r_t(0) + \alpha^*(t,x) \Big]\,dt + \Phi(t,x) \cdot dX_t.$$

The second step is to derive the dynamics of the bond price $B(t,T) = \exp(y(t,T))$. To apply Itô's formula, we define

$$\tilde{F} : A \mapsto H, \qquad g(\cdot) \to \exp(g(\cdot)).$$

Here A is chosen in a way, such that $\exp(g(\cdot))$, defined by $x \mapsto \exp(g(x))$, for all $x \in \mathbb{R}$, is again an element of H. Then we have $B(t,\cdot) = [\tilde{F}(y(t))](\cdot)$ or $B(t) = \tilde{F}(y(t))$, respectively.

We calculate the first and second derivative of $\tilde{F}$. First, define for $f, g \in H$ the product of f and g by $(f \times g)(\cdot) := f(\cdot)g(\cdot)$. Then $\tilde{F}(g(\cdot)) = \sum_{k=1}^\infty \frac{g(\cdot)^k}{k!}$. The derivative of g^2 is

$$D(g^2)(x) = 2g(x) \times \mathrm{id},$$

where id is the identity on H. The derivative of g^n is easily obtained by induction and we conclude

$$D\tilde{F}(g) = \tilde{F}(g) \times \mathrm{id} \quad \text{and} \quad D^2\tilde{F}(g) = \tilde{F}(g) \times \mathrm{id} \times \mathrm{id}\,.$$

Therefore, applying Itô's formula yields for the canonical dynamics

$$\begin{aligned} dB(t,T) = D\tilde{F}\big(B(u)\big) \cdot \Big[-\big(r_t(T-t) - r_t(0) + \alpha^*(t,T-t)\big)\,dt - \Phi(t,T-t) \cdot dX_t \Big] \\ + \frac{1}{2} \sum_{k=1}^\infty \lambda_k D^2\tilde{F}\big(B(t)\big) \big(\Phi(t,T-t) \cdot e_k, \Phi(t,T-t) \cdot e_k\big)\,dt + r_t(T-t)B(t,T-t) \\ = B(t,T)\Big[\big(r_t(0) - \alpha^*(t,T-t)\big)\,dt - \Phi(t,T-t) \cdot dX_t \\ + \frac{1}{2} \sum_{k=1}^\infty \lambda_k [\sigma_k^*(t,T-t)]^2\,dt \Big]. \end{aligned}$$

Define the discounting process $D_t := \exp\big(-\int_0^t r_u\,du\big)$. Note that D_t is of finite variation. Applying the common Itô-formula therefore yields

$$\begin{aligned} d[D_t B(t,T)] &= D_t B(t,T)\Big\{\Big[-\alpha^*(t,T-t)+\frac{1}{2}\sum_{k=1}^{\infty}\lambda_k[\sigma_k^*(t,T-t)]^2\Big]\,dt \\ &\qquad -\sum_{k=1}^{\infty}\sigma_k^*(t,T-t)\,d\beta_k(t)\Big\}. \end{aligned} \tag{2.13}$$

Consequently the discounted bond price is a martingale under $|||\sigma|||_{T^*} < \infty$, iff

$$\alpha^*(t,T-t) = \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k[\sigma_k^*(t,T-t)]^2\,dt, \qquad \forall T \in [t, t+T^{**}],$$

which is equivalent to

$$\int_0^x \alpha_t(v)\,dv = \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k\Big[\int_0^x \sigma_k(t,v)\,du\Big]^2\,dv\,dt.$$

Taking the partial derivative w.r.t. x, we get (2.12). ■

2.4 The Fubini Theorem

The Fubini theorem is just stated for convenience. For a proof, see Da Prato and Zabczyk (1992, p. 109 p.p.).

Let $(E,\mathcal{E})$ denote a measurable space and μ be a finite, positive measure on $(E,\mathcal{E})$. Furthermore, consider a predictable, measurable mapping[6]

$$\Phi(t,\omega,x) : ([0,T]\times\omega\times E) \to L_2(H_0,H).$$

Theorem 2.4.1. *Assume that*

$$\int_E |||\Phi(\cdot,\omega,x)|||_T\,\mu(dx) < \infty, \quad \textit{for } \mathbb{P}\textit{-almost all } \omega.$$

Then it follows that

$$\int_E \Big[\int_0^T \Phi(t,x)\cdot dX(t)\Big]\mu(dx) = \int_0^T \Big[\int_E \Phi(t,x)\,\mu(dx)\Big]\cdot dX(t), \quad \mathbb{P}\textit{-a.s.}$$

[6]For details on measurability and predictability in this case, see Da Prato and Zabczyk (1992, p. 109).

2.5 The Girsanov Theorem

Recall that we set $H_0 := D^{\frac{1}{2}}(H)$ with inner product $\langle \cdot, \cdot \rangle_0$

Theorem 2.5.1. *Consider a predictable process $(\mu(s))_{s\in[0,T]}$ with values in H_0 and set $\Phi(s)(\cdot) := \langle \mu(s), \cdot \rangle_0$. Define*

$$L_T := \exp\Big(\int_0^T \Phi(s) \cdot dX(s) - \frac{1}{2} \int_0^T |\mu(s)|_0^2 \, ds \Big)$$

and assume that $\mathbb{E}(L_T) = 1$. Here $(X(s))_{s\geqslant 0}$ is a D-Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$. Define the measure

$$d\tilde{P} := L_T d\mathbb{P}.$$

Then

$$\tilde{X}(t) := X(t) - \int_0^t \mu(s) ds, \quad t \in [0, T]$$

is a D-Wiener process under $\tilde{P}$.

Proof The proof has 3 steps. First, we assume that Ψ is bounded and proof $\mathbb{E}(L_T) = 1$. Then we show that $\tilde{X}$ is a Wiener process under $\tilde{P}$ and finally we extend the result to general Ψ by an approximation argument.

For the first step we will use a small Lemma, which clarifies the term $\int \Phi \cdot dX$ we will use a small Lemma, which clarifies the term $\int \Phi \cdot dX$.

Lemma 2.5.2. *Assume that Ψ a predictable H_0-valued process satisfies $\mathbb{P}(\int_0^T |\Psi(s)|_0^2 ds < \infty) = 1$. Then there exists a real-valued standard BM$(\beta(t))_{t\in[0,T]}$ (w.r.t. $\mathcal{F}_t$) s.t.*

$$\int_0^t \Phi(s) \cdot dX(s) = \int_0^t \langle \Psi(s), dX(s) \rangle_0$$
$$= \int_0^t |\Psi(s)|_0 \, d\beta(s) \quad \mathbb{P}\text{-a.s.}$$

Proof. Note that $\int \Phi \cdot dX$ is a continuous, real-valued process with quadratic variation $\int |\Psi|_0^2 \, ds$. Therefore

$$\beta(t) := \int_0^t 1_{\{|\Psi(s)|_0 \neq 0\}} \frac{1}{|\Psi(s)|_0} \Phi(s) \cdot dX(s)$$

is also a square integrable martingal with $< \beta >_t = t$, hence by Lévy's characterization theorem a standard BM.

Clearly

$$\int_0^t |\Psi(s)|_0 d\beta_s = \int_0^t \Phi(s) dX(s)$$

■

Proof. (continued) First assume that $|\Psi(t)|_0 \leqslant K$ for all $t \in [0,T]$ and some $K > 0$. By the Lemma it is clear that

$$\mathbb{E}\Big(\exp\Big[\int_0^T \Phi(s) dX_s = \frac{1}{2}\int_0^T |\Psi(s)|_0^2 \, ds\Big]\Big) = 1.$$

The next step is to show that $\tilde{W}$ is a $\tilde{P}$-Wiener process.

It is sufficient to show that $\int_0^T \langle g(t), d(t)\rangle_0$ is a Gaussian (with the correct covariance/quadr. variation) and that it has independent increments. Therefore, consider

$$\begin{aligned}
&\widetilde{\mathbb{E}}\Big(\exp\Big(\int_0^T < g(s), d\widetilde{X}(s) >_0 \Big)\Big) \\
&= \mathbb{E}\Big(\exp\Big(\int_0^T \langle g(s), dX(s)\rangle_0 + \int_0^T < g(s), \Psi(s) >_0 ds\Big) \\
&\qquad \cdot \exp\Big(\int_0^T \langle \Psi(s), dX(s)\rangle_0 - \frac{1}{2}\int_0^T \mid \Psi(s) \mid_0^2 ds\Big)\Big) \\
&= \mathbb{E}\Big(\exp\Big(\int_0^T \langle g + \Psi, dX\rangle - \frac{1}{2}\int_0^T \mid g + \Psi \mid_0^2 ds\Big)\Big) \cdot \exp\frac{1}{2}\int_0^T \mid g(s) \mid_0^2 ds
\end{aligned}$$

The expectation is 1 by the previous Lemma. We obtain the Laplace transform

$$\tilde{\mathbb{E}}\Big(\exp\Big(\lambda \int_0^T \langle g(s), d\tilde{X}(s)\rangle_0\Big)\Big) = \exp\Big(\frac{1}{2}\lambda^2 \int_0^T |g(s)|_0^2 \, ds\Big)$$

and so $\tilde{X}$ is a Gaussian process. Similar, we get

$$\tilde{\mathbb{E}}\Big(\exp\Big(\lambda \int_0^T \langle g, d\tilde{X}\rangle_0\Big) 1_F\Big) = \tilde{\mathbb{P}}(F)\, e^{\frac{1}{2}\lambda^2 \int_0^T |g(s)|_0^2}, \qquad \forall F \in \mathcal{F}_t$$

such that we have independent increments.

Finally, note that for $\Psi_N \xrightarrow{\mathcal{L}^2} \Psi$ we obtain a.s. convergence of the stochastic integrals and hence stochastic convergence of terms like

$$\exp\Big(\int \langle \Psi_N, dX\rangle - \frac{1}{2}\int |\Psi_N|_0^2 \, ds\Big).$$

As they all have expectation 1, we have uniform integrability and hence convergence of their expectations. ■

2.6 Some Examples

(i) Consider the differential equation (in $\mathbb{R}$, or R^n)

$$dy_t = f(y_t)dt + b(y_t)d\beta_t, y_0 = \zeta \in \mathbb{R}. \tag{2.14}$$

Then every solution depends on the initial value ζ. We therefore could consider the process (y) as $(y_t(\zeta))$, thus a new process $Y_t := y_t(\cdot)$ takes values in a functional space, of mappings $K \to R$ with $K \subset \mathbb{R}$ closed. This space E is contained in $C(K, \mathbb{R})$ and we assume it is a Hilbert space. Then

$$d\mathcal{Y}(t) = F(Y)(\zeta) := f\left(\mathcal{Y}(\zeta)\right)dt + B\left(> (t)\right)d\beta_t, \mathcal{Y}(0) = y \in E$$

with $F(\mathcal{Y})(\zeta) := f\left(\mathcal{Y}(\zeta)\right)$ and $B(\mathcal{Y})(\zeta) = b\left(\mathcal{Y}(\zeta)\right)$ is called the *lift* of (2.14).

If the identity belongs to t and the corresponding solution exists, then

$$y(t, \zeta) => (t, \mathrm{Id})(\zeta)$$

defines a version of the solution of (2.14) which continuously depends on ζ. This version is called a *stochastic flow.*

So results for stochastic flows can be obtained from results about Hilbert-space valued SDEs.

(ii) Delay equations

Consider the equation

$$dy_t = \left[\int_{-r}^{0} y(t+\theta)\mu(d\theta) + f(y_t)\right]dt + bd\beta_t, \tag{2.15}$$
$$y_0 = x_0, \quad y(\theta) = x_1(\theta), \quad \theta \in [-r, 0]$$

where μ is a finite measure on $[-r, 0]$. The solution of this SDE will not be Markovian. However, typically, if we enlarge the filtration we will get a Markov-process. Therefore, consider

$$X(t) := \begin{pmatrix} y(t) \\ (y(t+\theta)) \end{pmatrix}_{\theta \in [-r,0]}.$$

Then it can be shown that X solves

$$dX = \left(AX + F(X)\right)dt + B(X)d\beta, X_0 = \begin{pmatrix} X_0 \\ X_1(\cdot) \end{pmatrix}$$

with properly defined generator A and F, B.

In the converse, under fairly general conditions the first coordinate of X will solve (2.15). This type of lifting is called "Markovian lifting".

(iii) Random motion of a string

Consider a system of N particles, which are under the influence of 3 kinds of forces:

(a) *elastic forces* between neighbours

(b) *external force* f

(c) *random forces* of the white noise type

$$dy_k = \left[\frac{K}{2}N^2\left(y_{k+1} + y_{k-1} - 2y_k\right) + f(y_k)\right]_{dt} + \sqrt{N}\; b(y_k)dW_k$$

This leads (under Lipschitz conditions) to a unique process, with initial conditions $y_k(0)$ and $y_0(t), y_{N+1}(t)$.

Now the question is how to obtain a notion for the whole string. Define

$$X_N(t,z) := y_k(t) + \frac{z - z_k}{z_{k+1} - z_k} y_{k+1}(t) \text{ for } z \in [z_k, z_{k+1}), k = 1, \ldots, N-1$$

Consider the function space $E = L^2\left([0,1], \mathbb{R}^{(d)}\right)$ and

$$E^2 = \{X \in W^{2,2}([0,1], \mathbb{R}) : \frac{d^2x}{dz^2} \in E\}$$

Then (Funaki 1993) $X_N \xrightarrow{\mathcal{L}} X$ with

$$dX_t = \left(\frac{\partial^2}{\partial z^2} X_t + F(X_t)\right) dt + B(z)dW_t,$$

where $F(x)(z) = f\left(x(z)\right)$ and $(B(x)u) = b\left(X(z)\right)u(z), z \in [0,1]$ and W is a cylindrical Wiener process on $U = L^2\left([0,1], \mathbb{R}\right)$.

Nomenclature

H_0 .... $:= D^{\frac{1}{2}}H$, where D is a covariance operator, p. 14
$L(H)$ .... space of linear, continuous ($\Rightarrow$ bounded) mappings $H \to H$, p. 5
$L_2(H_0, H)$ .... space of Hilbert-Schmidt operators $H_0 \to H$, p. 14

Appendix A

The Taylor formula on Banach spaces

In this section we develop the derivative in a context of a Banach space. Observe that a Hilbert space is itself a Banach space, so this applies very well to our setting. For a detailed treatment of derivatives in Banach spaces see Dieudonné (1969).

Consider two Banach Spaces U, H and an open subset $A \subset U$ and a continuous function $f : A \mapsto H$. The derivative of f at a point x_0 is a linear function which locally approximates f. This local approximation is made more precise by the notion "tangent".

Definition A.1.1. We call f and g *tangent* at point $x_0 \in A$ if

$$\lim_{x \to x_0, x \neq x_0} \frac{||f(x) - g(x)||}{||x - x_0||} = 0.$$

For example consider two linear functions on $\mathbb{R}$ which meet in x_0. Then they are tangent iff they have the same leverage. In this sense it can be shown, that if there is a linear function tangent to f at x_0 it is unique and we call it derivative.

Definition A.1.2. The continuous mapping $f : A \mapsto H$ is said to be **differentiable** at $x_0 \in A$ if there is a linear mapping $u : U \mapsto H$, such that $x \to f(x_0) + u(x - x_0)$ is tangent to f at x_0. This mapping is unique and called the **derivative** or (total derivative) of f at x_0, written $Df(x_0)$.

Note that the derivative is continuous and a linear mapping from U to H, thus element of the Hilbert space $L(U; H)$. For $u \in L(U; H)$ we often write $u \cdot t$ instead of $u(t)$.

For this derivative formal rules, like the chain rule, linearity and rules for the computation of the derivative of an inverse mapping hold true and can be found in Dieudonné (1969).

If the derivative Df, itself a continuous mapping of A into $L(U; H)$ is differentiable at a point $x_0 \in A$, we say that f is **twice differentiable** at x_0 and write for the second derivative $D^2 f(x_0)$. Not that it is an element of $L(U; L(U; H))$ which can be naturally identified[1] with the space of continuous bilinear mappings of $U \times U$ into H, written $L(U, U; H)$. The second derivative is a symmetric mapping.

[1]Identify $u \cdot (s, t)$ with $(u \cdot s) \cdot t$.

As a result of the mean value theorem we obtain **Taylor's formula**. For our purpose we state it only for $p = 2$ while the result is valid for any p.

Theorem A.1.3. *For a twice continuously differentiable mapping of A into H and if $x + \theta t \in A$ for all $\theta \in [0, 1]$, we have*

$$f(x + t) = f(x) + Df(x) \cdot t + \frac{1}{2!} D^2 f(x + \zeta t) \cdot t^{(2)}.$$

Here $t^{(2)}$ stands for (t, t) and ζ is an element of $[0, 1]$.

Consider a functional space, which consists of real-valued functions on $[0, T]$ and is a Hilbert space with innerproduct $\langle \cdot, \cdot \rangle$. We define $F : A \mapsto H$ by

$$F(g)(\cdot) = \exp(g(\cdot)),$$

where $A = \{g \in H : F(g) \in H\}$. Then the Taylor formula applies to F. Denote the bilinear mapping $\times : H \times H \mapsto H$ by

$$(f \times g)(\cdot) = f(\cdot) g(\cdot).$$

For $F(g) = g \times g = g^2$, we have

$$DF(x) = 2x \times \mathrm{id},$$

where id is the identity on H. This is true, because

$$\begin{aligned} F(x) - F(x_0) - 2x_0 \times (x - x_0) &= x \cdot x - x_0 \cdot x_0 - 2x_0 \cdot (x - x_0) \\ &= (x - x_0)^2 \end{aligned}$$

and, therefore

$$\begin{aligned} \lim_{x \to x_0} \frac{||F(x) - F(x_0) - 2x_0 \times (x - x_0)||}{||x - x_0||} &= \lim_{x \to x_0} \frac{||(x - x_0)^2||}{||x - x_0||} \\ &\leq \lim_{x \to x_0} \frac{||x - x_0||}{||x - x_0||} \sup_{t \in [0,T]} |x(t) - x_0(t)| = 0. \end{aligned}$$

By induction, this holds also for $F(g) = g^n$ and we therefore conclude for $F(g)(\cdot) = \exp(g(\cdot)) = \sum_k \frac{g^k}{k!}$

$$DF(x) = F(x) \times \mathrm{id}\,.$$

References

Bogachev, V. I. (1991). *Gaussian Measures*. American Mathematical Society.

Da Prato, G. and J. Zabczyk (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge University Press.

Dieudonné, J. (1969). *Foundations of Modern Analysis*. Academic Press.

Filipović, D. (2001). *Consistency Problems for Heath-Jarrow-Morton Interest Rate Models*, Volume 1760 of *Lecture Notes in Mathematics*. Springer Verlag. Berlin Heidelberg New York.

Heath, D., R. A. Jarrow, and A. J. Morton (1992). Bond pricing and the term structure of interest rates. *Econometrica 60*, 77–105.

Itô, K. (1946). On a stochastic integral equation. *Proc. Imperial Acad Tokyo 22*, 32–35.

Karatzas, I. and S. E. Shreve (1988). *Brownian Motion and Stochastic Calculus*. Springer Verlag. Berlin Heidelberg New York.

Musiela, M. (1993). Stochastic PDEs and term structure models. *Journées Internationales de France, IGR-AFFI, La Baule*.

Revuz, D. and M. Yor (1994). *Continuous Martingales and Brownian Motion*. Springer Verlag. Berlin Heidelberg New York.

Werner, D. (2000). *Funktionalanalysis*. Springer.

Yosida, K. (1971). *Functional Analysis*. Springer Verlag. Berlin Heidelberg New York.