Infinite Dimensional Analysis

Alessandra Lunardi, Michele Miranda, Diego Pallara

2015–16

We collect here the revised edition of the 19^{th} Internet Seminar on "Infinite dimensional analysis". In the lectures, we consider separable infinite dimensional Banach spaces endowed with Gaussian measures and we describe their main properties; in particular we are interested in integration by parts formulae that allow the definition of gradient and divergence operators. Once these tools are introduced, we study Sobolev spaces. In the context of Gaussian analysis the role of the Laplacian ($\Delta = \operatorname{div}\nabla$) is played by the Ornstein-Uhlenbeck operator. We study the realisation of the Ornstein-Uhlenbeck operator and the Ornstein-Uhlenbeck semigroup in spaces of continuous functions and in L^p spaces. In particular, for $p = 2$ the Ornstein-Uhlenbeck operator is self-adjoint and we show that there exists an orthogonal basis consisting of explicit eigenfunctions (the Hermite polynomials) that give raise to the "Wiener Chaos Decomposition".

In the present revision we have taken into account the feedback coming from all the participants to the discussion board and we are grateful for all the contributions, which have corrected many mistakes and improved the presentation. We warmly thank in particular, Prof. Jürgen Voigt and the whole Dresden group for their careful reading and their constructive criticism. We are planning to prepare another revision after the final workshop in Casalmaggiore, where we shall take into account possible further commnets.

Contents

Lecture 1

Preliminaries

We present the basic notions of measure theory, with the aim of fixing the notation and making the exposition self-contained. We deal only with finite measures, even though of course positive infinite measures are the first natural examples. In particular, we assume familiarity with the Lebesgue measure in $\mathbb{R}^d$, which we denote by λ_d. Assuming that the basic notions relative to positive measures are known, we go straight to finite real measures because all measures we are going to discuss are of this type. In the first section we present real measures and the related notions of L^p spaces, absolute continuous and singular measures, the Radon-Nikodym theorem, weak convergence and product measures. In the case of topological spaces we introduce Borel and Radon measures. Next, we introduce characteristic functions (or Fourier transforms) of measures and Gaussian measures in $\mathbb{R}^d$.

1.1 Abstract Measure Theory

We start by introducing measurable spaces, i.e., sets equipped with a σ-algebra.

Definition 1.1.1 (σ-algebras and measurable spaces)**.** *Let X be a nonempty set and let $\mathscr{F}$ be a collection of subsets of X.*

(a) We say that $\mathscr{F}$ is an algebra *in X if $\emptyset \in \mathscr{F}$, $E_1 \cup E_2 \in \mathscr{F}$ and $X \setminus E_1 \in \mathscr{F}$ whenever E_1, $E_2 \in \mathscr{F}$.*

(b) We say that an algebra $\mathscr{F}$ is a σ-algebra *in X if for any sequence $(E_h) \subset \mathscr{F}$ its union $\bigcup_h E_h$ belongs to $\mathscr{F}$.*

(c) For any collection $\mathscr{G}$ of subsets of X, the σ-algebra generated by $\mathscr{G}$ is the smallest σ-algebra containing $\mathscr{G}$. If (X, τ) is a topological space, we denote by $\mathscr{B}(X)$ the σ-algebra of Borel subsets of X, i.e., the σ-algebra generated by the open subsets of X.

(d) If $\mathscr{F}$ is a σ-algebra in X, we call the pair $(X, \mathscr{F})$ a measurable space.

It is obvious by the De Morgan laws that algebras are closed under finite intersections, and σ-algebras are closed under countable intersections. Moreover, since the intersection of any family of σ-algebras is a σ-algebra and the set of all subsets of X is a σ-algebra, the definition of generated σ-algebra is well posed. Once a σ-algebra has been fixed, it is possible to introduce positive measures.

Definition 1.1.2 (Finite measures)**.** *Let $(X, \mathscr{F})$ be a measurable space and let $\mu : \mathscr{F} \to [0, \infty)$. We say that μ is a* positive finite measure *if $\mu(\emptyset) = 0$ and μ is σ-additive on $\mathscr{F}$, i.e., for any sequence (E_h) of pairwise disjoint elements of $\mathscr{F}$ the equality*

$$\mu\left(\bigcup_{h=0}^{\infty} E_h\right) = \sum_{h=0}^{\infty} \mu(E_h) \tag{1.1.1}$$

holds. We say that μ is a probability measure *if $\mu(X) = 1$.*
We say that $\mu : \mathscr{F} \to \mathbb{R}$ is a (finite) real measure *if $\mu = \mu_1 - \mu_2$, where μ_1 and μ_2 are positive finite measures. The triple $(X, \mathscr{F}, \mu)$ is called a* measure space*.*

If μ is a real measure, we define its *total variation* $|\mu|$ for every $E \in \mathscr{F}$ as follows:

$$|\mu|(E) := \sup\left\{\sum_{h=0}^{\infty} |\mu(E_h)| \colon E_h \in \mathscr{F} \text{ pairwise disjoint, } E = \bigcup_{h=0}^{\infty} E_h\right\} \tag{1.1.2}$$

and it turns out to be a positive measure, see Exercise 1.2. If μ is real, then (1.1.1) still holds, and the series converges absolutely, as the union is independent of the order. Notice also that the following equality holds for Borel measures μ on $\mathbb{R}^d$:

$$|\mu|(\mathbb{R}^d) = \sup\Big\{\int_{\mathbb{R}^d} f d\mu : \ f \in C_b(\mathbb{R}^d), \|f\|_\infty \leq 1\Big\}. \tag{1.1.3}$$

See Exercise 1.3.

Remark 1.1.3. (Monotonicity) Any positive finite measure μ is monotone with respect to set inclusion and continuous along monotone sequences, i.e., if $(E_h) \subset \mathscr{F}$ is an increasing sequence of sets (resp. a decreasing sequence of sets), then

$$\mu\left(\bigcup_{h=0}^{\infty} E_h\right) = \lim_{h\to\infty} \mu(E_h), \qquad \text{resp.} \quad \mu\left(\bigcap_{h=0}^{\infty} E_h\right) = \lim_{h\to\infty} \mu(E_h),$$

see Exercise 1.1.

We recall the following (unique) extension theorem for σ-additive set functions defined on an algebra. It is a classical result due to K. Carathéodory, we refer to [9], Theorems 3.1.4 and 3.1.10.

Theorem 1.1.4 (Carathéodory Extension Theorem)**.** *Let $\mathscr{G}$ be an algebra of sets of X. If $\mu : \mathscr{G} \to [0, \infty)$ is a finite σ-additive set function then μ can be uniquely extended to $\sigma(\mathscr{G})$ and the extension is a measure on $(X, \sigma(\mathscr{G}))$.*

Definition 1.1.5 (Radon measures)**.** *A real measure μ on the Borel sets of a topological space X is called a* real Radon measure *if for every $B \in \mathscr{B}(X)$ and $\varepsilon > 0$ there is a compact set $K \subset B$ such that $|\mu|(B \setminus K) < \varepsilon$.*
A measure is tight *if the same property holds for $B = X$.*

Proposition 1.1.6. *If (X, d) is a separable complete metric space then every real measure on $(X, \mathscr{B}(X))$ is Radon.*

Proof. Observe that it is enough to prove the result for finite positive measures. The general case follows splitting the given real measure into its positive and negative parts.
Let then μ be a positive finite measure on $(X, \mathscr{B}(X))$. Let us first show that it is a regular measure, i.e., for any $B \in \mathscr{B}(X)$ and for any $\varepsilon > 0$ there are an open set $G \supset B$ and a closed set $F \subset B$ such that $\mu(G \setminus F) < \varepsilon$. Indeed, for a given $\varepsilon > 0$, if $B = F$ is closed it suffices to consider open sets $G_\delta = \{x \in X : \ d(x, F) = \inf_{y \in F} d(x, y) < \delta\}$, getting $F = \bigcap_{\delta > 0} G_\delta$. As $\mu(G_\delta) \to \mu(F)$ as $\delta \to 0$ by Remark 1.1.3, fixed $\varepsilon > 0$, for δ small enough we have $\mu(G_\delta \setminus F) < \varepsilon$. Next, we show that the family $\mathscr{G}$ containing $\emptyset$ and all sets $B \in \mathscr{B}(X)$ such that for any $\varepsilon > 0$ there are an open set $G \supset B$ and a closed set $F \subset B$ such that $\mu(G \setminus F) < \varepsilon$ is a σ-algebra. To this aim, given a sequence $(B_n) \subset \mathscr{G}$, consider open sets G_n and closed sets F_n such that $F_n \subset B_n \subset G_n$ and $\mu(G_n \setminus F_n) < \varepsilon/2^{n+1}$. For $G = \bigcup_{n=1}^\infty G_n$ and $F = \bigcup_{n=1}^N F_n$, with $N \in \mathbb{N}$ such that $\mu(\bigcup_{n=1}^\infty F_n \setminus F) < \varepsilon/2$, we have $F \subset \bigcup_n B_n \subset G$ and $\mu(G \setminus F) < \varepsilon$. Therefore, $\mathscr{G}$ is closed under countable unions, and, since it is closed under complementation as well, it is a σ-algebra.
Since we have proved that all closed sets belong to $\mathscr{G}$, then it coincides with $\mathscr{B}(X)$.

As a consequence, we prove that any positive finite measure on $(X, \mathscr{B}(X))$ is Radon iff it is tight. If μ is Radon then it is tight by definition. Conversely, assuming that μ is tight, for every $\varepsilon > 0$ and every Borel set $B \subset X$, there exists a compact set K_1 such that $\mu(X \setminus K_1) < \varepsilon$ and a closed set $F \subset B$ such that $\mu(B \setminus F) < \varepsilon$. Then, the set $K := K_1 \cap F$ is compact, because it is complete (being closed) and precompact, and it verifies $\mu(B \setminus K) < 2\varepsilon$.

Therefore, to prove our statement it suffices to show that every positive finite Borel measure on X is tight. Let (x_n) be a dense sequence and notice that $X \subset \bigcup_{n=1}^\infty \overline{B}(x_n, 1/k)$ for every $k \in \mathbb{N}$. Then, given $\varepsilon > 0$, for every $k \in \mathbb{N}$ there is $N_k \in \mathbb{N}$ such that

$$\mu\Big(\bigcup_{n=1}^{N_k} \overline{B}(x_n, 1/k)\Big) > \mu(X) - \varepsilon/2^k.$$

Then, the set

$$K := \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{N_k} \overline{B}(x_n, 1/k)$$

is closed and totally bounded. It verifies $\mu(K) > \mu(X) - \varepsilon$. $\square$

Let us come to measurable functions.

Definition 1.1.7 (Measurable functions)**.** *Let $(X, \mathscr{F}, \mu)$ be a measure space and let $(Y, \mathscr{G})$ be a measurable space. A function $f : X \to Y$ is said to be $(\mathscr{F} - \mathscr{G})$*-measurable *if*

*$f^{-1}(A) \in \mathscr{F}$ for every $A \in \mathscr{G}$. If Y is a topological space, a function $f : X \to Y$ is said to be $\mathscr{F}$-*measurable *(or μ-measurable) if $f^{-1}(A) \in \mathscr{F}$ for every open set $A \subset Y$.*

In particular, if Y is a topological space and f is $\mathscr{F}$-measurable then $f^{-1}(B) \in \mathscr{F}$ for every $B \in \mathscr{B}(Y)$. For $E \subset X$ we define the *indicator* (or *characteristic*) *function* of E, denoted by $\mathbb{1}_E$, by

$$\mathbb{1}_E(x) := \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

and we say that $f : X \to \mathbb{R}$ is a *simple function* if the image of f is finite, i.e., if f belongs to the vector space generated by the indicator functions of measurable sets. We assume that the readers are familiar with the usual notion of integral of a measurable function. For μ positive, we define the L^p (semi)-norms and spaces as follows,

$$\|u\|_{L^p(X,\mu)} := \left(\int_X |u|^p \, d\mu \right)^{1/p}$$

if $1 \le p < \infty$, and

$$\|u\|_{L^\infty(X,\mu)} := \inf \{ C \in [0, \infty] \colon |u(x)| \le C \text{ for } \mu\text{-a.e. } x \in X \}$$

for every $\mathscr{F}$-measurable $u : X \to \mathbb{R}$. We define the space $L^p(X, \mu)$ as the space of equivalence classes of functions u agreeing μ-a.e. such that $\|u\|_{L^p(X,\mu)} < \infty$. In this way, $\|\cdot\|_{L^p(X,\mu)}$ is a norm and $L^p(X, \mu)$ is a Banach space, see e.g. [9, Theorem 5.2.1]. When there is no risk of confusion, we use the shorter notation $\|\cdot\|_p$.

We assume that the reader is familiar also with the properties of integrals, measurable functions and L^p spaces as well as the main convergence theorems of Levi, Fatou, Lebesgue, see e.g. [9, Section 4.3]. We just recall the convergence in measure and the Lebesgue-Vitali theorem on uniformly integrable sequences, see Exercise 1.4.

Definition 1.1.8. *Let $(X, \mathscr{F}, \mu)$ be a measure space and let (f_k) be a sequence of real-valued $\mathscr{F}$-measurable functions. We say thet (f_k) converges in measure to the function $f : X \to \mathbb{R}$ if*

$$\lim_{k \to \infty} \mu(\{x \in X : \ |f_k(x) - f(x)| > \varepsilon\}) = 0 \qquad \text{for every } \varepsilon > 0 \ . \tag{1.1.4}$$

Theorem 1.1.9 (Lebesgue-Vitali Convergence Theorem)**.** *Let $(X, \mathscr{F})$ be a measurable space, let μ be a positive finite measure on it and let (f_k) be a sequence of measurable functions such that*

$$\lim_{M \to \infty} \sup_{k \in \mathbb{N}} \int_{\{|f_k| > M\}} |f_k| \, d\mu = 0.$$

If $f_k \to f$ in measure then $f \in L^1(X, \mu)$ and $\lim_{k \to \infty} \int_X |f - f_k| d\mu = 0$.

Given a σ-algebra on a set X, we have defined the class of real-valued measurable functions. Conversely, given a family of real-valued functions defined on X, it is possible to define a suitable σ-algebra.

Definition 1.1.10. *Given a family F of functions $f : X \to \mathbb{R}$, let us define the σ-algebra $\mathscr{E}(X, F)$ generated by F on X as the smallest σ-algebra such that all the functions $f \in F$ are measurable, i.e., the σ-algebra generated by the sets $\{f < t\}$, with $f \in F$ and $t \in \mathbb{R}$.*

Given a metric space, the set of real Borel measures μ is a vector space in an obvious way. All continuous and bounded functions are in $L^1(X, \mu)$ and we define the *weak convergence of measures* by

$$\mu_j \to \mu \quad \Longleftrightarrow \quad \int_X f\, d\mu_j \to \int_X f\, d\mu \qquad \forall f \in C_b(X). \tag{1.1.5}$$

Let us now introduce the notions of absolute continuity and singularity of measures. Let μ be a positive finite measure and let ν be a real measure on the measurable space $(X, \mathscr{F})$. We say that ν is *absolutely continuous* with respect to μ, and we write $\nu \ll \mu$, if $\mu(B) = 0 \implies |\nu|(B) = 0$ for every $B \in \mathscr{F}$. If μ, ν are real measures, we say that they are *mutually singular*, and we write $\nu \perp \mu$, if there exists $E \in \mathscr{F}$ such that $|\mu|(E) = 0$ and $|\nu|(X \setminus E) = 0$. Notice that for mutually singular measures μ, ν the equality $|\mu + \nu| = |\mu| + |\nu|$ holds. If $\mu \ll \nu$ and $\nu \ll \mu$ we say that μ and ν are *equivalent* and we write $\mu \approx \nu$. If μ is a positive measure and $f \in L^1(X, \mu)$, then the measure $\nu := f\mu$ defined below is absolutely continuous with respect to μ and the following integral representations hold, see Exercise 1.2:

$$\nu(B) := \int_B f\, d\mu, \quad |\nu|(B) = \int_B |f|\, d\mu \qquad \forall B \in \mathscr{F}. \tag{1.1.6}$$

In the following classical theorem we see that if a real measure ν is absolutely continuous with respect to μ, then the above integral representation holds, with a suitable f.

Theorem 1.1.11 (Radon-Nikodym)**.** *Let μ be a positive finite measure and let ν be a real measureon the same measurable space $(X, \mathscr{F})$. Then there is a unique pair of real measures ν^a, ν^s such that $\nu^a \ll \mu$, $\nu^s \perp \mu$ and $\nu = \nu^a + \nu^s$. Moreover, there is a unique function $f \in L^1(X, \mu)$ such that $\nu^a = f\mu$. The function f is called the density (or Radon-Nikodym derivative) of ν with respect to μ and it is denoted by $d\nu/d\mu$.*

Since trivially each real measure μ is absolutely continuous with respect to $|\mu|$, from the Radon-Nikodym theorem the *polar decomposition* of μ follows: there exists a unique real valued function $f \in L^1(X, |\mu|)$ such that $\mu = f|\mu|$ and $|f| = 1$ $|\mu|$-a.e.

The following result is a useful criterion of mutual singularity.

Theorem 1.1.12 (Hellinger)**.** *Let μ, ν be two probability measures on a measurable space $(X, \mathscr{F})$, and let λ be a positive measure such that $\mu \ll \lambda$, $\nu \ll \lambda$. Then the integral*

$$H(\mu, \nu) := \int_X \sqrt{\frac{d\mu}{d\lambda} \frac{d\nu}{d\lambda}}\, d\lambda$$

is independent of λ and

$$2(1 - H(\mu, \nu)) \le |\mu - \nu|(X) \le 2\sqrt{1 - H(\mu, \nu)^2}. \tag{1.1.7}$$

Proof. Let us first take $\lambda = \mu+\nu$ and notice that $\mu, \nu \ll \lambda$. Then, setting $f := d\mu/d\lambda$ and $g := d\nu/d\lambda$, i.e., $\mu = f\lambda$ and $\nu = g\lambda$, we have $|\mu - \nu|(X) = \|f - g\|_{L^1(X,\lambda)}$ and integrating the inequalities

$$(\sqrt{f} - \sqrt{g})^2 \leq |f - g| = |\sqrt{f} - \sqrt{g}|\,|\sqrt{f} + \sqrt{g}|$$

we get

$$\begin{aligned}
\int_X (\sqrt{f} - \sqrt{g})^2\, d\lambda = 2(1 - H(\mu, \nu)) \leq \int_X |f - g|\, d\lambda &= |\mu - \nu|(X) \\
&= \int_X |\sqrt{f} - \sqrt{g}|\,|\sqrt{f} + \sqrt{g}|\, d\lambda \\
&\leq \Big(\int_X |\sqrt{f} - \sqrt{g}|^2\, d\lambda\Big)^{1/2} \Big(\int_X |\sqrt{f} + \sqrt{g}|^2\, d\lambda\Big)^{1/2} \\
&= (2 - 2H(\mu, \nu))^{1/2}(2 + 2H(\mu, \nu))^{1/2} = 2\sqrt{1 - H(\mu, \nu)^2},
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality. If λ' is another measure such that $\mu = f'\lambda' \ll \lambda'$ and $\nu = g'\lambda' \ll \lambda'$, then $\lambda \ll \lambda'$: setting $\phi := \frac{d\lambda}{d\lambda'}$, we have $f' = \phi f,\ g' = \phi g$ and then

$$\int_X \sqrt{\frac{d\mu}{d\lambda}\frac{d\nu}{d\lambda}}\, d\lambda = \int_X \sqrt{fg}\, d\lambda = \int_X \sqrt{fg}\phi\, d\lambda' = \int_X \sqrt{f'g'}\, d\lambda' = \int_X \sqrt{\frac{d\mu}{d\lambda'}\frac{d\nu}{d\lambda'}}\, d\lambda'.$$

□

Corollary 1.1.13. *If μ and ν are probability measures, then $\mu \perp \nu$ iff $H(\mu, \nu) = 0$ iff $|\mu - \nu|(X) = 2$.*

Proof. It is obvious from Hellinger's theorem that $|\mu-\nu|(X) = 2$ if and only if $H(\mu, \nu) = 0$. Let us show that this is equivalent to $\mu \perp \nu$. Using the notation in the proof of Theorem 1.1.12, notice that $H(\mu, \nu) = 0$ if and only if the set F defined by $F := \{fg \neq 0\}$ verifies $\lambda(F) = 0$ (hence also $\mu(F) = \nu(F) = 0$). Therefore, for the measurable set $E = \{f = 0, g > 0\}$ we have $\mu(E) = \nu(X \setminus E) = 0$ and the assertion follows. □

We recall the notions of *push-forward* of a measure (or *image measure*) and the constructions and main properties of *product measure*. The push-forward of a measure generalises the classical change of variable formula.

Definition 1.1.14 (Push-forward)**.** *Let $(X, \mathscr{F})$ and $(Y, \mathscr{G})$ be measurable spaces, and let $f : X \to Y$ be $(\mathscr{F}, \mathscr{G})$-measurable, i.e., such that $f^{-1}(F) \in \mathscr{F}$ whenever $F \in \mathscr{G}$. For any positive or real measure μ on $(X, \mathscr{F})$ we define the push-forward or image measure of μ under f, that is the measure $\mu \circ f^{-1}$ in $(Y, \mathscr{G})$ by*

$$\mu \circ f^{-1}(F) := \mu\left(f^{-1}(F)\right) \qquad \forall F \in \mathscr{G}.$$

Sometimes $\mu \circ f^{-1}$ is denoted by $f_{\#}\mu$.

The *change of variables formula* immediately follows from the previous definition. If $u \in L^1(Y, \mu \circ f^{-1})$, then $u \circ f \in L^1(X, \mu)$ and we have the equality

$$\int_Y u \, d(\mu \circ f^{-1}) = \int_X (u \circ f) \, d\mu. \tag{1.1.8}$$

The above equality is nothing but the definition for simple functions f, and it is immediately extended to the whole of L^1 by density.

We consider now two measure spaces and we describe the natural resulting structure on their cartesian product.

Definition 1.1.15 (Product σ-algebra)**.** *Let $(X_1, \mathscr{F}_1)$ and $(X_2, \mathscr{F}_2)$ be measurable spaces. The* product σ-algebra *of $\mathscr{F}_1$ and $\mathscr{F}_2$, denoted by $\mathscr{F}_1 \times \mathscr{F}_2$, is the σ-algebra generated in $X_1 \times X_2$ by*

$$\mathscr{G} = \{E_1 \times E_2 \colon E_1 \in \mathscr{F}_1, E_2 \in \mathscr{F}_2\}.$$

Remark 1.1.16. Let $E \in \mathscr{F}_1 \times \mathscr{F}_2$; then for every $x \in X_1$ the section $E_x := \{y \in X_2 \colon (x, y) \in E\}$ belongs to $\mathscr{F}_2$, and for every $y \in X_2$ the section $E^y := \{x \in X_1 \colon (x, y) \in E\}$ belongs to $\mathscr{F}_1$. In fact, it is easily checked that the families

$$\mathscr{G}_x := \{F \in \mathscr{F}_1 \times \mathscr{F}_2 \colon F_x \in \mathscr{F}_2\}, \quad \mathscr{G}^y := \{F \in \mathscr{F}_1 \times \mathscr{F}_2 \colon F^y \in \mathscr{F}_1\}$$

are σ-algebras in $X_1 \times X_2$ and contain $\mathscr{G}$, see Exercise 1.5.

Theorem 1.1.17 (Fubini)**.** *Let $(X_1, \mathscr{F}_1, \mu_1)$, $(X_2, \mathscr{F}_2, \mu_2)$ be measure spaces with μ_1, μ_2 positive and finite. Then, there is a unique positive finite measure μ on $(X_1 \times X_2, \mathscr{F}_1 \times \mathscr{F}_2)$, denoted by $\mu_1 \otimes \mu_2$, such that*

$$\mu(E_1 \times E_2) = \mu_1(E_1) \cdot \mu_2(E_2) \qquad \forall E_1 \in \mathscr{F}_1, \, \forall E_2 \in \mathscr{F}_2.$$

Furthermore, for any μ-measurable function $u : X_1 \times X_2 \to [0, \infty)$ the functions

$$x \mapsto \int_{X_2} u(x, y) \, \mu_2(dy) \quad \textit{and} \quad y \mapsto \int_{X_1} u(x, y) \, \mu_1(dx)$$

are respectively μ_1-measurable and μ_2-measurable and

$$\begin{aligned} \int_{X_1 \times X_2} u \, d\mu &= \int_{X_1} \left(\int_{X_2} u(x, y) \, \mu_2(dy) \right) \mu_1(dx) \\ &= \int_{X_2} \left(\int_{X_1} u(x, y) \, \mu_1(dx) \right) \mu_2(dy). \end{aligned}$$

Remark 1.1.18. More generally, it is possible to construct a product measure on infinite cartesian products. If I is a set of indices, typically $I = [0, 1]$ or $I = \mathbb{N}$, and $(X_t, \mathscr{F}_t, \mu_t)$, $t \in I$, is a family of probability spaces, the product σ-algebra is that generated by the family of sets of the type

$$B = B_1 \times \cdots \times B_n \times \bigtimes_{t \in I \setminus \{t_1, \ldots, t_n\}} X_t, \quad B_k \in \mathscr{F}_{t_k},$$

whose measure is $\mu(B) := \mu_{t_1}(B_1) \cdots \mu_{t_n}(B_n)$.

In the sequel we shall sometimes encounter some ideas coming from probability theory and stochastic analysis. In order to simplify several computations concerning probability measures on $\mathbb{R}^d$, it is often useful to use *characteristic functions* of measures. This is the probabilistic counterpart of Fourier transform. Indeed, given a finite real measure μ on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, we define its characteristic function by setting

$$\hat{\mu}(\xi) := \int_{\mathbb{R}^d} e^{ix\cdot\xi}\,\mu(dx), \qquad \xi \in \mathbb{R}^d. \tag{1.1.9}$$

We list the main elementary properties of characteristic functions, whose proofs are in Exercise 1.6.

1. $\hat{\mu}$ is uniformly continuous on $\mathbb{R}^d$;
2. $\hat{\mu}(0) = \mu(X)$;
3. if $\hat{\mu}_1 = \hat{\mu}_2$ then $\mu_1 = \mu_2$;
4. if $\mu_j \to \mu$ in the sense of (1.1.5), then $\hat{\mu}_j \to \hat{\mu}$ uniformly on compacts;
5. if (μ_j) is a sequence of probability measures and there is $\phi : \mathbb{R}^d \to \mathbb{C}$ continuous in $\xi = 0$ such that $\hat{\mu}_j \to \phi$ pointwise, then there is a probability measure μ such that $\hat{\mu} = \phi$.

1.2 Gaussian measures

Gaussian (probability) measures are the main reference measures we shall encounter in the Lectures. Let us start from the finite dimensional case. We recall the following elementary equality

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} \exp\Big\{-\frac{(x-a)^2}{2\sigma^2}\Big\}\,dx = 1, \tag{1.2.1}$$

that holds for all $a \in \mathbb{R}$ and $\sigma > 0$. An easy way to prove (1.2.1) is to compute the double integral

$$\int_{\mathbb{R}^2} \exp\{-(x^2+y^2)\}\,dxdy$$

using polar coordinates and the change of variables formula and apply Fubini Theorem 1.1.17.

Definition 1.2.1 (Gaussian measures on $\mathbb{R}$)**.** *A probability measure γ on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is called* Gaussian *if it is either a Dirac measure δ_a at a point a (in this case, we put $\sigma = 0$), or a measure absolutely continuous with respect to the Lebesgue measure λ_1 with density*

$$\frac{1}{\sigma\sqrt{2\pi}} \exp\Big\{-\frac{(x-a)^2}{2\sigma^2}\Big\},$$

with $a in \mathbb{R}$ and $\sigma > 0$. In this case we call a the mean*, σ the* mean-square deviation *and σ^2 the* variance *of γ and we say that γ is* centred *or* symmetric *if $a = 0$ and* standard *if in addition $\sigma = 1$.*

By elementary computations we get

$$a = \int_{\mathbb{R}} x\,\gamma(dx), \qquad \sigma^2 = \int_{\mathbb{R}} (x-a)^2\,\gamma(dx).$$

Remark 1.2.2. For every $a, \sigma \in \mathbb{R}$ we have $\hat{\gamma}(\xi) = e^{ia\xi - \frac{1}{2}\sigma^2\xi^2}$, see Exercise 1.7. Conversely, by property 3 of characteristic functions, a probability measure on $\mathbb{R}$ is Gaussian iff its characteristic function has this form. Therefore, it is easy to recognise a Gaussian measure from its characteristic function. This is true in $\mathbb{R}^d$, as we are going to see in Proposition 1.2.4, and also in infinite dimensions, as we shall see in the next lecture.

Let us come to Gaussian measures in $\mathbb{R}^d$.

Definition 1.2.3 (Gaussian measures on $\mathbb{R}^d$)**.** *A Borel probability measure γ on $\mathbb{R}^d$ is said to be Gaussian if for every linear functional ℓ on $\mathbb{R}^d$ the measure $\gamma \circ \ell^{-1}$ is Gaussian on $\mathbb{R}$.*

The first example of Gaussian measure in $\mathbb{R}^d$ is $\gamma_d := (2\pi)^{-d/2} e^{-|x|^2/2}\lambda_d$, that is called *standard Gaussian measure.* We denote by G_d the *standard Gaussian density* in $\mathbb{R}^d$, i.e., the density of γ_d with respect to λ_d. Notice also that if $d = h + k$ then $\gamma_d = \gamma_h \otimes \gamma_k$.

The following result gives a useful characterisation of a Gaussian measure through its characteristic function.

Proposition 1.2.4. *A measure γ on $\mathbb{R}^d$ is Gaussian if and only if its characteristic function is*

$$\hat{\gamma}(\xi) = \exp\Big\{ia \cdot \xi - \frac{1}{2} Q\xi \cdot \xi\Big\} \tag{1.2.2}$$

for some $a \in \mathbb{R}^d$ and Q nonnegative symmetric $d \times d$ matrix. Moreover, γ is absolutely continuous with respect to the Lebesgue measure λ_d if and only if Q is nondegenerate. In this case, the density of γ is

$$\frac{1}{\sqrt{(2\pi)^d \det Q}} \exp\Big\{-\frac{1}{2}\big(Q^{-1}(x-a)\cdot(x-a)\big)\Big\}. \tag{1.2.3}$$

Proof. Let γ be a measure such that (1.2.2) holds. Then, for every linear funtional $\ell : \mathbb{R}^d \to \mathbb{R}$ (here we identify ℓ with the vector in $\mathbb{R}^d$ such that $\ell(x) = \ell \cdot x$) we compute the characteristic function of the measure $\mu_\ell := \gamma \circ \ell^{-1}$ on $\mathbb{R}$,

$$\widehat{\mu_\ell}(\tau) = \int_{\mathbb{R}} e^{i\tau t}\,\mu_\ell(dt) = \int_{\mathbb{R}^d} e^{i\tau\ell(x)}\,\gamma(dx) = \hat{\gamma}(\tau\ell) = \exp\Big\{i\tau a\cdot\ell - \frac{\tau^2}{2}Q\ell\cdot\ell\Big\}$$

by (1.2.2). Therefore, by Remark 1.2.2 μ_ℓ is a Gaussian measure with mean $a_\ell = a \cdot \ell$ and variance $\sigma_\ell^2 = Q\ell \cdot \ell$, and also γ is a Gaussian measure by the arbitrariness of ℓ.

Conversely, assume that μ_ℓ is Gaussian for every ℓ as above. Its mean a_ℓ and its variance σ_ℓ^2 are given by

$$a_\ell := \int_{\mathbb{R}} t\,\mu_\ell(dt) = \int_{\mathbb{R}^d} \ell(x)\,\gamma(dx) = \ell \cdot \Big(\int_{\mathbb{R}^d} x\,\gamma(dx)\Big) \tag{1.2.4}$$

$$\sigma_\ell^2 := \int_{\mathbb{R}} (t - a_\ell)^2\,\mu_\ell(dt) = \int_{\mathbb{R}^d} (\ell(x) - a_\ell)^2\,\gamma(dx). \tag{1.2.5}$$

These formulas show that the map $\ell \mapsto a_\ell$ is linear and the map $\ell \mapsto \sigma_\ell^2$ is a nonnegative quadratic form. Therefore, there are a vector $a \in \mathbb{R}^d$ and a nonnegative definite symmetric matrix $Q = (Q_{ij})$ such that $a_\ell = a \cdot \ell$ and $\sigma_\ell^2 = Q\ell \cdot \ell$, whence (1.2.2) follows. Notice that

$$a = \int_{\mathbb{R}^d} x\,\gamma(dx), \qquad Q_{ij} = \int_{\mathbb{R}^d} (x_i - a_i)(x_j - a_j)\,\gamma(dx).$$

To prove the last part of the statement, let us assume that $\gamma \ll \lambda_d$, i.e. $\gamma = f\lambda_d$ for some $f \in L^1(\mathbb{R}^d, \lambda_d)$. We want to show that $Q\ell \cdot \ell = 0$ iff $\ell = 0$. From (1.2.4), (1.2.5) we have

$$Q\ell \cdot \ell = \int_{\mathbb{R}^d} (\ell \cdot (x - a))^2 f(x) dx,$$

then $Q\ell \cdot \ell = 0$ iff $(\ell \cdot (x - a))^2 = 0$ for a.e. $x \in \mathbb{R}^d$, i.e. iff $\ell = 0$, as $f \not\equiv 0$. Hence Q is nondegenerate.

Conversely, if Q is nondegenerate, we consider the measure $\nu = f\lambda_d$ with f given by (1.2.3) and we compute its characteristic function. Using the change of variable $z = Q^{-1/2}(x-a)$, since $\gamma_d = \otimes_{j=1}^d \gamma_1$, we have:

$$\begin{aligned}
\hat\nu(\xi) &= \int_{\mathbb{R}^d} \exp\{i\xi \cdot x\} f(x) dx = \frac{1}{\sqrt{(2\pi)^d \det Q}} \int_{\mathbb{R}^d} \exp\Big\{ i\xi \cdot x - \frac{1}{2}\big(Q^{-1}(x-a)\cdot(x-a)\big)\Big\} dx \\
&= \exp\{i\xi \cdot a\} \int_{\mathbb{R}^d} \exp\{iQ^{1/2}\xi \cdot z\}\gamma_d(dz) = \exp\{i\xi \cdot a\} \prod_{j=1}^d \int_{\mathbb{R}} \exp\{i(Q^{1/2}\xi)_j t\}\gamma_1(dt) \\
&= \exp\{i\xi \cdot a\} \prod_{j=1}^d \hat\gamma_1((Q^{1/2}\xi)_j) = \exp(i\xi \cdot a) \prod_{j=1}^d \exp\left\{ -\frac{1}{2}((Q^{1/2}\xi)_j)^2 \right\} \\
&= \exp\left\{ i\xi \cdot a - \frac{1}{2} Q\xi \cdot \xi \right\} = \hat\gamma(\xi).
\end{aligned}$$

Hence, by property 3 of the characteristic function, $\gamma = \nu$. □

Remark 1.2.5. If γ is a Gaussian measure and (1.2.2) holds, we call a the *mean* and Q the *covariance* of γ, and we write $\gamma = \mathscr{N}(a, Q)$ when it is useful to emphasise the relevant parameters. If $a = 0$ we say that γ is centred. If the matrix Q is invertible then the Gaussian measure $\gamma = \mathscr{N}(a, Q)$ is said to be *nondegenerate*. Its density, given by (1.2.3), is denoted $G_{a,Q}$. The nondegeneracy is equivalent to the fact that $\mu_\ell \ll \lambda_1$ for every $\ell \in \mathbb{R}^d$.

Proposition 1.2.6. *Let γ centred Gaussian measure in $\mathbb{R}^d$ and for every $\theta \in \mathbb{R}$ define the map $\phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ by $\phi(x, y) := x \sin\theta + y \cos\theta$. Then, the image measure $(\gamma \otimes \gamma) \circ \phi^{-1}$ in $\mathbb{R}^d$ is γ.*

Proof. We use characteristic functions and Proposition 1.2.4. Indeed, the characteristic function of γ is $\exp\{-\frac{1}{2}Q\xi \cdot \xi\}$ for some nonnegative $d \times d$ matrix Q. Then we may compute the characteristic function of $\mu := (\gamma \otimes \gamma) \circ \phi^{-1}$ as follows:

$$\begin{aligned}
\hat{\mu}(\xi) &= \int_{\mathbb{R}^d} e^{iz\xi}\, \mu(dz) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{i(x \sin\theta + y\cos\theta)\xi}\, \gamma \otimes \gamma(d(x,y)) \\
&= \int_{\mathbb{R}^d} e^{i(x\sin\theta)\xi}\, \gamma(dx) \int_{\mathbb{R}^d} e^{i(y\cos\theta)\xi}\, \gamma(dy) \\
&= \exp\{-\frac{1}{2}\sin^2\theta\, Q\xi\cdot\xi\} \exp\{-\frac{1}{2}\cos^2\, \theta Q\xi\cdot\xi\} \\
&= \exp\{-\frac{1}{2}Q\xi\cdot\xi\},
\end{aligned}$$

and the assertion follows from property 3 of characteristic functions. □

Remark 1.2.7. We point out that the property stated in Proposition 1.2.6 is not the invariance of γ under rotations in $\mathbb{R}^d$. Indeed, rotation invariance holds iff the covariance of γ is a positive multiple of an orthogonal matrix.

1.3 Exercises

Exercise 1.1. Let μ be a positive finite measure on $(X, \mathscr{F})$. Prove the monotonicity properties stated in Remark 1.1.3.

Exercise 1.2. Prove that the set function $|\mu|$ defined in (1.1.2) is a positive finite measure and that the integral representation for $|\nu| = |f\mu|$ in (1.1.6) holds. Prove also that if $\mu \perp \nu$ then $|\mu + \nu| = |\mu| + |\nu|$.

Exercise 1.3. Prove the equality (1.1.3).

Exercise 1.4. Prove the Vitali–Lebesgue Theorem 1.1.9.

Exercise 1.5. Prove that the families $\mathscr{G}_x$ and $\mathscr{G}^y$ defined in Remark 1.1.16 are σ-algebras.

Exercise 1.6. Prove the properties of characteristic functions listed in Section 1.1.

Exercise 1.7. Prove the equality $\hat{\gamma}(\xi) = e^{ia\xi - \frac{1}{2}\sigma^2\xi^2}$ stated in Remark 1.2.2.

Exercise 1.8. (Layer cake formula) Prove that if μ is a positive finite measure on $(X, \mathscr{F})$ and $0 \le f \in L^1(X, \mu)$ then

$$\int_X f\, d\mu = \int_0^\infty \mu\big(\{x \in X :\ f(x) > t\}\big)\, dt.$$

Lecture 2

Gaussian measures in infinite dimension

In this lecture, after recalling a few notions about σ-algebras in Fréchet spaces, we introduce Gaussian measures and we prove the Fernique Theorem. This is a powerful tool to get precise estimates on the Gaussian integrals. In most of the course, our framework is an infinite dimensional separable real Fréchet or Banach space that we denote by X. In the latter case, the norm is denoted by $\|\cdot\|$ or $\|\cdot\|_X$ when there is risk of confusion. The separability assumption is not essential in most of the theory, but *Radon* Gaussian measures are concentrated on separable subspaces, see [3, Theorem 3.6.1], hence we assume saparability from the beginning. The open (resp. closed) ball with centre $x \in X$ and radius $r > 0$ will be denoted by $B(x,r)$ (resp. $\overline{B}(x,r)$). We denote by X^*, with norm $\|\cdot\|_{X^*}$, the topological dual of X consisting of all linear continuous functions $f : X \to \mathbb{R}$. Sometimes, we shall discuss the case of a separable Hilbert space in order to highlight some special features. The only example that is not a separable Banach space is $\mathbb{R}^I$ (see Lecture 3), but we prefer to describe this as a particular case rather than to present a more general abstract theory.

2.1 σ-algebras in infinite dimensional spaces and characteristic functions

In order to present further properties of measures in X, and in particular approximation of measures and functions, it is useful to start with a discussion on the relevant underlying σ-algebras. Besides the Borel σ-algebra, to take advantage of the finite dimensional reductions we shall frequently encounter in the sequel, we introduce the σ-algebra $\mathscr{E}(X)$ generated by the *cylindrical* sets, i.e, the sets of the form

$$C = \Big\{x \in X : \ (f_1(x), \dots, f_n(x)) \in C_0\Big\},$$

where $f_1, \dots, f_n \in X^*$ and $C_0 \in \mathscr{B}(\mathbb{R}^n)$, called a *base* of C. According to Definition 1.1.10, $\mathscr{E}(X) = \mathscr{E}(X, X^*)$. Notice that the cylindrical sets constitute an *algebra*. The following important result holds. We do not present its most general version, but we do not even confine ourselves to Banach spaces because we shall apply it to $\mathbb{R}^\infty$, which is a Fréchet space. We recall that a Fréchet space is a complete metrisable locally convex topological vector space, i.e., a vector space endowed with a sequence of seminorms that generate a metrisable topology such that the space is complete.

Theorem 2.1.1. *If X is a separable Fréchet space, then $\mathscr{E}(X) = \mathscr{B}(X)$. Moreover, there is a countable family $F \subset X^*$ separating the points in X (i.e., such that for every pair of points $x \neq y \in X$ there is $f \in F$ such that $f(x) \neq f(y)$) such that $\mathscr{E}(X) = \mathscr{E}(X, F)$.*

Proof. Let (x_n) be a sequence dense in X, and denote by (p_k) a family of seminorms which defines the topology of X. By the Hahn-Banach theorem for every n and k there is $\ell_{n,k} \in X^*$ such that $p_k(x_n) = \ell_{n,k}(x_n)$ and $\sup\{\ell_{n,k}(x) :\ x \in X,\, p_k(x) \leq 1\} = 1$, whence $\ell_{n,k}(x) \leq p_k(x)$ for every $n, k \in \mathbb{N}$ and $x \in X$. As a consequence, for every $x \in X$ and $k \in \mathbb{N}$ we have $p_k(x) = \sup_n\{\ell_{n,k}(x)\}$. Indeed, if (x_{n_h}) is a subsequence of (x_n) such that $x_{n_h} \to x$ as $h \to \infty$, we have $\lim_h \ell_{n_h,k}(x_{n_h}) = p_k(x)$. Therefore, for every $r > 0$

$$\overline{B}_k(x, r) := \{y \in X :\ p_k(y - x) \leq r\} = \bigcap_{n \in \mathbb{N}} \{y \in X :\ \ell_{n,k}(y - x) \leq r\} \in \mathscr{E}(X).$$

As X is separable, there is a countable base of its topology. For instance, we may take the sets $B_k(x_n, r)$ with rational r, see [10, I.6.2 and I.6.12]. These sets are in $\mathscr{E}(X)$, as

$$B_k(x_n, r) = \bigcap_{m \in \mathbb{N}} \{x \in X :\ \ell_{m,k}(x - x_n) < r\}.$$

Hence, the inclusion $\mathscr{B}(X) \subset \mathscr{E}(X)$ holds. The converse inclusion is trivial.
To prove the last statement, just take $F = \{\ell_{n,k},\ n, k \in \mathbb{N}\}$. It is obviously a countable family; let us show that it separates points. If $x \neq y$, there is $k \in \mathbb{N}$ such that $p_k(x - y) = \sup_n \ell_{n,k}(x - y) > 0$ and therefore there is $\bar{n} \in \mathbb{N}$ such that $\ell_{\bar{n},k}(x - y) > 0$. $\square$

In the discussion of the properties of Gaussian measures in infinite dimensional spaces, as in $\mathbb{R}^d$, the characteristic functions, defined by

$$\hat{\mu}(f) := \int_X \exp\{if(x)\}\, \mu(dx), \qquad f \in X^*, \tag{2.1.1}$$

play an important role. The properties of characteristic functions seen in Lecture 1 can be extended to the present context. We discuss in detail only the extension of property (iii) (the injectivity), which is the most important for our purposes. In the following proposition, we use the coincidence criterion for measures agreeing on an algebra of generators of the σ-algebra, see e.g. [9, Theorem 3.1.10].

Proposition 2.1.2. *Let X be a separable Fréchet space, and let $\mu_1,\ \mu_2$ be two probability measures on $(X, \mathscr{B}(X))$. If $\widehat{\mu}_1 = \widehat{\mu}_2$ then $\mu_1 = \mu_2$.*

Proof. It is enough to show that if $\hat{\mu} = 0$ then $\mu = 0$ and in particular, by Theorem 2.1.1, that $\mu(C) = 0$ when C is a cylinder with base $C_0 \in \mathscr{B}(\mathbb{R}^d)$. Let be $\hat{\mu} = 0$, consider $F = \operatorname{span}\{f_1, \ldots, f_d\} \subset X^*$ and define $\mu_F = \mu \circ P_F^{-1}$, where $P_F : X \to \mathbb{R}^d$ is given by $P_F(x) = (f_1(x), \ldots, f_d(x))$. Then for any $\xi \in \mathbb{R}^d$

$$\widehat{\mu}_F(\xi) = \int_F \exp\{i\xi \cdot y\}\, \mu_F(dy) = \int_X \exp\{i\xi \cdot P_F(x)\}\, \mu(dx) = \int_X \exp\{iP_F^*\xi(x)\}\, \mu(dx) = 0,$$

where $P_F^* : \mathbb{R}^d \to X^*$ is the adjoint map

$$P_F^*(\xi) = \sum_{i=1}^{d} \xi_i f_i.$$

It follows that $\mu_F = 0$ and therefore the restriction of μ to the σ-algebra $\mathscr{E}(X, F)$ is the null measure. □

2.2 Gaussian measures in infinite dimensional spaces

Measure theory in infinite dimensional spaces is far from being a trivial issue, because there is no equivalent of the Lebesgue measure, i.e., there is no nontrivial measure invariant by translations.

Proposition 2.2.1. *Let X be an infinite dimensional separable Hilbert space. If $\mu : \mathscr{B}(X) \to [0, \infty]$ is a σ-additive set function such that:*

(i) $\mu(x + B) = \mu(B)$ for every $x \in X$, $B \in \mathscr{B}(X)$,

(ii) $\mu(B(0, r)) > 0$ for every $r > 0$,

then $\mu(A) = \infty$ for every open set A.

Proof. Assume that μ satisfies (i) and (ii), and let $\{e_n : n \in \mathbb{N}\}$ be an orthonormal basis in X. For any $n \in \mathbb{N}$ consider the balls B_n with centre $2re_n$ and radius $r > 0$; they are pairwise disjoint and by assumption they have the same measure, say $\mu(B_n) = m > 0$ for all $n \in \mathbb{N}$. Then,

$$\bigcup_{n=1}^{\infty} B_n \subset B(0, 3r) \quad \Longrightarrow \quad \mu(B(0, 3r)) \geq \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} m = \infty,$$

hence $\mu(A) = \infty$ for every open set A. □

Definition 2.2.2 (Gaussian measures on X)**.** *Let X be a separable Fréchet space. A probability measure γ on $(X, \mathscr{B}(X))$ is said to be Gaussian if $\gamma \circ f^{-1}$ is a Gaussian measure in $\mathbb{R}$ for every $f \in X^*$. The measure γ is called* centred *(or* symmetric*) if all the measures $\gamma \circ f^{-1}$ are centred and it is called* nondegenerate *if for any $f \neq 0$ the measure $\gamma \circ f^{-1}$ is nondegenerate.*

Our first task is to give characterisations of Gaussian measures in infinite dimensions in terms of characteristic functions analogous to those seen in $\mathbb{R}^d$, Proposition 1.2.4.

Notice that if $f \in X^*$ then $f \in L^p(X,\gamma)$ for every $p \geq 1$: indeed, the integral

$$\int_X |f(x)|^p \,\gamma(dx) = \int_{\mathbb{R}} |t|^p \,(\gamma \circ f^{-1})(dt)$$

is finite because $\gamma \circ f^{-1}$ is Gaussian in $\mathbb{R}$. Therefore, we can give the following definition.

Definition 2.2.3. *We define the* mean a_γ *and the* covariance B_γ *of* γ *by*

$$a_\gamma(f) := \int_X f(x)\,\gamma(dx), \tag{2.2.1}$$

$$B_\gamma(f,g) := \int_X [f(x) - a_\gamma(f)]\,[g(x) - a_\gamma(g)]\,\gamma(dx), \tag{2.2.2}$$

$f, g \in X^*$.

Observe that $f \mapsto a_\gamma(f)$ is linear and $(f,g) \mapsto B_\gamma(f,g)$ is bilinear in X^*. Moreover, $B_\gamma(f,f) = \|f - a_\gamma(f)\|^2_{L^2(X,\gamma)} \geq 0$ for every $f \in X^*$.

Theorem 2.2.4. *A Borel probability measure γ on X is Gaussian if and only if its characteristic function is given by*

$$\hat{\gamma}(f) = \exp\Big\{ia(f) - \frac{1}{2}B(f,f)\Big\}, \qquad f \in X^*, \tag{2.2.3}$$

where a is a linear functional on X^* and B is a nonnegative symmetric bilinear form on X^*.

Proof. Assume that γ is Gaussian. Let us show that $\hat{\gamma}$ is given by (2.2.3) with $a = a_\gamma$ and $B = B_\gamma$. Indeed, we have:

$$\widehat{\gamma}(f) = \int_{\mathbb{R}} \exp\{i\xi\}(\gamma \circ f^{-1})(d\xi) = \exp\Big\{im - \frac{1}{2}\sigma^2\Big\},$$

where m and σ^2 are the mean and the covariance of $\gamma \circ f^{-1}$, given by

$$m = \int_{\mathbb{R}} \xi(\gamma \circ f^{-1})(d\xi) = \int_X f(x)\gamma(dx) = a_\gamma(f),$$

and

$$\sigma^2 = \int_{\mathbb{R}} (\xi - m)^2(\gamma \circ f^{-1})(d\xi) = \int_X (f(x) - a_\gamma(f))^2\gamma(dx) = B_\gamma(f,f).$$

Conversely, let γ be a Borel probability measure on X and assume that (2.2.3) holds. Since a is linear and B is bilinear, we can compute the Fourier transform of $\gamma \circ f^{-1}$, for $f \in X^*$, as follows:

$$\begin{aligned}\widehat{\gamma \circ f^{-1}}(\tau) &= \int_{\mathbb{R}} \exp\{i\tau t\}\,(\gamma \circ f^{-1})(dt) = \int_X \exp\{i\tau f(x)\}\,\gamma(dx) \\ &= \exp\Big\{i\tau a(f) - \frac{1}{2}\tau^2 B(f,f)\Big\}.\end{aligned}$$

According to Remark 1.2.2, $\gamma \circ f^{-1} = \mathscr{N}(a(f), B(f,f))$ is Gaussian and we are done. □

Remark 2.2.5. We point out that at the moment we have proved that a_γ is linear on X^* and B_γ is bilinear on $X^* \times X^*$, but they are not necessarily continuous. We shall see that if X is a Banach space then a_γ and B_γ are in fact continuous.

As in the finite dimensional case, we say that γ is *centred* if $a_\gamma = 0$; in this case, the bilinear form B_γ is nothing but the restriction of the inner product in $L^2(X, \gamma)$ to X^*,

$$B_\gamma(f, g) = \int_X f(x)g(x)\,\gamma(dx), \qquad B_\gamma(f, f) = \|f\|^2_{L^2(X,\gamma)}. \tag{2.2.4}$$

In the sequel we shall frequently consider centred Gaussian measures; this requirement is equivalent to the following symmetry property.

Proposition 2.2.6. *Let γ be a Gaussian measure on a Fréchet space X and define the measure μ by*

$$\mu(B) := \gamma(-B), \qquad \forall B \in \mathscr{B}(X).$$

Then, γ is centred if and only if $\gamma = \mu$.

Proof. We know that $\widehat{\gamma}(f) = \exp\{ia_\gamma(f) - \frac{1}{2}\|f - a_\gamma(f)\|^2_{L^2(X,\gamma)}\}$. On the other hand, since $\mu = \gamma \circ R^{-1}$ with $R : X \to X$ given by $R(x) = -x$,

$$\widehat{\mu}(f) = \int_X e^{if(x)}\mu(dx) = \int_X e^{-if(x)}\gamma(dx) = \exp\Big\{-ia_\gamma(f) - \frac{1}{2}\|f - a_\gamma(f)\|^2_{L^2(X,\gamma)}\Big\}.$$

Then $\widehat{\mu} = \widehat{\gamma}$ if and only if $a_\gamma(f) = 0$ for any $f \in X^*$, whence the statement follows by Proposition 2.1.2. $\square$

Let us draw some interesting (and useful) consequences from the above result.

Proposition 2.2.7. *Let X be a separable Freéchet space and let γ be a Gaussian measure on X.*

(i) If μ is a Gaussian measure on a Fréchet space Y, then $\gamma \otimes \mu$ is a Gaussian measure on $X \times Y$.

(ii) If μ is another Gaussian measure on X, then the convolution measure $\gamma * \mu$, defined as the image measure in X of $\gamma \otimes \mu$ on $X \times X$ under the map $(x, y) \mapsto x + y$ is a Gaussian measure and is given by

$$\gamma * \mu(B) = \int_X \mu(B - x)\gamma(dx) = \int_X \gamma(B - x)\mu(dx). \tag{2.2.5}$$

(iii) If γ is centred, then for every $\theta \in \mathbb{R}$ the image measure $(\gamma \otimes \gamma) \circ R_\theta^{-1}$ in $X \times X$ under the map $R_\theta : X \times X \to X \times X$, $R_\theta(x, y) := (x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta)$ is again $\gamma \otimes \gamma$.

(iv) If γ is centred, then for every $\theta \in \mathbb{R}$ the image measures $(\gamma \otimes \gamma) \circ \phi_i^{-1}$, $i = 1, 2$ in X under the maps $\phi_i : X \times X \to X$,

$$\phi_1(x, y) := x \cos\theta + y \sin\theta, \quad \phi_2(x, y) := -x \sin\theta + y \cos\theta$$

are again γ.

Proof. All these results follow quite easily by computing the relevant characteristic functions. We start with (i) taking into account that for $f \in X^*$ and $g \in Y^*$ we have

$$B_\gamma(f, f) = \|f - a_\gamma(f)\|^2_{L^2(X,\gamma)}, \qquad B_\mu(g, g) = \|g - a_\mu(g)\|^2_{L^2(Y,\mu)}.$$

For every $\ell \in (X \times Y)^*$, we define f and g by

$$\ell(x, y) = \ell(x, 0) + \ell(0, y) =: f(x) + g(y).$$

Then

$$\begin{aligned}
\widehat{\gamma \otimes \mu}(\ell) &= \int_{X\times Y} \exp\{i\ell(x, y)\}(\gamma \otimes \mu)(d(x, y)) \\
&= \int_X \exp\{if(x)\}\gamma(dx) \int_Y \exp\{ig(y)\}\mu(dy) \\
&= \exp\left\{ia_\gamma(f) - \frac{1}{2}\|f - a_\gamma(f)\|^2_{L^2(X,\gamma)} + a_\mu(g) - \frac{1}{2}\|g - a_\mu(g)\|^2_{L^2(Y,\mu)}\right\} \\
&= \exp\left\{i(a_\gamma(f) + a_\mu(g)) - \frac{1}{2}\left(\|f - a_\gamma(f)\|^2_{L^2(X,\gamma)} + \|g - a_\mu(g)\|^2_{L^2(Y,\mu)}\right)\right\}.
\end{aligned}$$

On the other hand, since

$$\begin{aligned}
&\int_{X\times Y} (f(x) - a_\gamma(f))(g(y) - a_\mu(g))(\gamma \otimes \mu)(d(x, y)) = \\
&\qquad\qquad = \int_X (f(x) - a_\gamma(f))\gamma(dx) \int_Y (g(y) - a_\mu(g))\,\mu(dy) = 0,
\end{aligned}$$

we have

$$\begin{aligned}
&\|f - a_\gamma(f)\|^2_{L^2(X,\gamma)} + \|g - a_\mu(g)\|^2_{L^2(Y,\mu)} = \\
&= \int_X (f(x) - a_\gamma(f))^2\gamma(dx) + \int_Y (g(y) - a_\mu(g))^2\mu(dy) \\
&= \int_X (\ell(x, 0) - a_\gamma(f))^2\gamma(dx) + \int_Y (\ell(0, y) - a_\mu(g))^2\mu(dy) \\
&= \int_{X\times Y} (\ell(x, 0) - a_\gamma(f))^2(\gamma \otimes \mu)(d(x, y)) + \int_{X\times Y} (\ell(0, y) - a_\mu(g))^2(\gamma \otimes \mu)(d(x, y)) \\
&= \int_{X\times Y} (\ell(x, 0) - a_\gamma(f))^2 + (\ell(0, y) - a_\mu(g))^2(\gamma \otimes \mu)(d(x, y)) \\
&= \int_{X\times Y} (\ell(x, 0) + \ell(0, y) - (a_\gamma(f) + a_\mu(g)))^2(\gamma \otimes \mu)(d(x, y)) \\
&= \int_{X\times Y} (\ell(x, y) - (a_\gamma(f) + a_\mu(g)))^2(\gamma \otimes \mu)(d(x, y)).
\end{aligned}$$

So we have

$$B_{\gamma\otimes\mu}(\ell,\ell) = \|\ell(x,y) - (a_\gamma(f) + a_\mu(g)\|_{L^2(X\times Y,\gamma\otimes\mu)} = B_\gamma(f,f) + B_\mu(g,g)$$

if we decompose ℓ by $\ell(x,y) = f(x) + g(y)$, $f(x) = \ell(x,0)$, $g(y) = \ell(0,y)$ as before.

The proof of statement (ii) is similar; indeed if $h : X \times X \to X$ is given by $h(x,y) = x + y$, then

$$\begin{aligned}
(\gamma \widehat{\otimes \mu) \circ} h^{-1}(\ell) =& \int_X \exp\{i\ell(x)\} \left((\gamma\otimes\mu)\circ h^{-1}\right)(dx) \\
=& \int_{X\times X} \exp\{i\ell(h(x,y))\}(\gamma\otimes\mu)(d(x,y)) \\
=& \int_X \exp\{i\ell(x)\}\gamma(dx) \int_X \exp\{i\ell(y)\}\mu(dy) \\
=& \exp\left\{ia_\gamma(\ell) - \frac{1}{2}\left(\|\ell\|^2_{L^2(X,\gamma)} + ia_\mu(\ell)\|\ell\|^2_{L^2(X,\mu)}\right)\right\}.
\end{aligned}$$

for every $\ell \in X^*$. Using the notation of Remark 1.1.16, for every $B \in \mathscr{B}(X)$ we have

$$\begin{aligned}
\{(x,y) \in X\times X : h(x,y) \in B\}_x = B - x, \\
\{(x,y) \in X\times X : h(x,y) \in B\}^y = B - y.
\end{aligned}$$

Applying the Fubini Theorem to the characteristic functions of $h^{-1}(B)$ we deduce that the convolution measure is given by (2.2.5),

$$\begin{aligned}
\gamma * \mu(B) = (\gamma\otimes\mu)(h^{-1}(B)) =& \int_{X\times X} 1\!\!1_{h^{-1}(B)}(\gamma\otimes\mu)(d(x,y)) \\
=& \int_X \gamma(dx) \int_{\{y\in X: x+y\in B\}} \mu(dy) = \int_X \mu(B-x)\gamma(dx).
\end{aligned}$$

To show (iii), set $\mu := (\gamma\otimes\gamma)\circ R_\theta^{-1}$; taking into account that for any $\ell \in (X\times X)^*$ we have

$$\begin{aligned}
\ell(R_\theta(x,y)) =& \ell(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta) \\
=& \underbrace{\ell(x,0)\cos\theta - \ell(0,x)\sin\theta}_{=:f_\theta(x)} + \underbrace{\ell(y,0)\sin\theta + \ell(0,y)\cos\theta}_{=:g_\theta(y)}.
\end{aligned}$$

We find

$$\begin{aligned}
\hat{\mu}(\ell) =& \int_{X\times X} \exp\{i\ell(R_\theta(x,y))\}(\gamma\otimes\gamma)(d(x,y)) \\
=& \int_X \exp\{if_\theta(x)\}\gamma(dx) \int_X \exp\{ig_\theta(y)\}\gamma(dy) \\
=& \exp\left\{-\frac{1}{2}\left(B_\gamma(f_\theta,f_\theta) + B_\gamma(g_\theta,g_\theta)\right)\right\}.
\end{aligned}$$

Since B_γ is bilinear,

$$B_\gamma(f_\theta, f_\theta) + B_\gamma(g_\theta, g_\theta) = B_\gamma(f_0, f_0) + B_\gamma(g_0, g_0),$$

where $f_0(x) = \ell(x, 0)$, $g_0(y) = \ell(0, y)$ and

$$B_\gamma(f_0, f_0) + B_\gamma(g_0, g_0) = B_{\gamma\otimes\gamma}(\ell, \ell)$$

by the proof of statement (i).

To prove (iv), we notice that the laws of $\gamma \otimes \gamma$ under the projection maps $p_1, p_2 : X \times X \to X$,

$$p_1(x, y) := x, \qquad p_2(x, y) := y$$

are γ by definition of product measure. Since $\phi_i = p_i \circ R_\theta$, $i = 1, 2$ we deduce

$$(\gamma \otimes \gamma) \circ \phi_i^{-1} = (\gamma \otimes \gamma) \circ (p_i \circ R_\theta)^{-1} = ((\gamma \otimes \gamma) \circ R_\theta^{-1}) \circ p_i^{-1} = (\gamma \otimes \gamma) \circ p_i^{-1} = \gamma.$$

$\square$

2.3 The Fernique Theorem

In this section we prove the Fernique Theorem; we start by proving it in the case of centred Gaussian measures and then we extend the result to any Gaussian measure.

Theorem 2.3.1 (Fernique)**.** *Let γ be a centred Gaussian measure on a separable Banach space X. Then there exists $\alpha > 0$ such that*

$$\int_X \exp\{\alpha \|x\|^2\}\, \gamma(dx) < \infty.$$

Proof. If γ is a Dirac measure the result is trivial, therefore we may assume that this is not the case. The idea of the proof is to show that the measures of suitable annuli decay fast enough to compensate the growth of the exponential function in the integral. Let us fix $t > \tau > 0$ and let us estimate $\gamma(\{\|x\| \le \tau\})\gamma(\{\|x\| > t\})$. Using property (iii) of Proposition 2.2.7 with $\theta = -\frac{\pi}{4}$, we obtain

$$\begin{aligned}
&\gamma(\{x \in X : \|x\| \le \tau\})\gamma(\{x \in X : \|x\| > t\}) \\
&= (\gamma \otimes \gamma)\left(\{(x, y) \in X \times X : \|x\| \le \tau\} \cap \{(x, y) \in X \times X : \|y\| > t\}\right) \\
&= (\gamma \otimes \gamma)\left(\left\{(x, y) \in X \times X : \frac{\|x - y\|}{\sqrt{2}} \le \tau\right\} \cap \left\{(x, y) \in X \times X : \frac{\|x + y\|}{\sqrt{2}} > t\right\}\right).
\end{aligned}$$

The triangle inequality yields $\|x\|, \|y\| \ge \frac{\|x+y\|}{2} - \frac{\|x-y\|}{2}$, which implies the inclusion

$$\begin{aligned}
&\left\{(x, y) \in X \times X : \frac{\|x - y\|}{\sqrt{2}} \le \tau\right\} \cap \left\{(x, y) \in X \times X : \frac{\|x + y\|}{\sqrt{2}} > t\right\} \\
&\subset \left\{(x, y) \in X \times X : \|x\| > \frac{t - \tau}{\sqrt{2}}\right\} \cap \left\{(x, y) \in X \times X : \|y\| > \frac{t - \tau}{\sqrt{2}}\right\}.
\end{aligned}$$

As a consequence, we have the estimate

$$\gamma(\{\|x\| \le \tau\})\gamma(\{\|x\| > t\}) \le \gamma\left(X \setminus \overline{B}\Big(0, \frac{t-\tau}{\sqrt{2}}\Big)\right)^2. \tag{2.3.1}$$

We leave as an exercise, Exercise 2.1, the fact that if γ is not a Dirac measure, then $\gamma(\overline{B}(0,\tau)) < 1$ for any $\tau > 0$. Let us fix $\tau > 0$ such that $c := \gamma\big(\overline{B}(0,\tau)\big) \in (1/2, 1)$ and set

$$\begin{aligned}
&\alpha := \frac{1}{24\tau^2}\log\Big(\frac{c}{1-c}\Big),\\
&t_0 := \tau, \quad t_n := \tau + \sqrt{2}t_{n-1} = \tau(1+\sqrt{2})\big(\sqrt{2}^{n+1} - 1\big), \qquad n \ge 1.
\end{aligned}$$

Applying estimate (2.3.1) with $t = t_n$ and recalling that $\frac{t_n - \tau}{\sqrt{2}} = t_{n-1}$, we obtain

$$\gamma(X \setminus \bar{B}(0, t_n)) \le \frac{\gamma(X \setminus \bar{B}(0, t_{n-1}))^2}{\gamma(\bar{B}(0,\tau))} = \Big(\frac{\gamma(X \setminus \bar{B}(0, t_{n-1}))}{c}\Big)^2 c$$

and iterating

$$\gamma(X \setminus \bar{B}(0, t_n)) \le c\left(\frac{1-c}{c}\right)^{2^n}.$$

Therefore

$$\begin{aligned}
\int_X \exp\{\alpha\|x\|^2\}\,\gamma(dx) = &\int_{\overline{B}(0,\tau)} \exp\{\alpha\|x\|^2\}\,\gamma(dx) + \\
&+ \sum_{n=0}^{\infty} \int_{\overline{B}(0,t_{n+1})\setminus\overline{B}(0,t_n)} \exp\{\alpha\|x\|^2\}\,\gamma(dx)\\
\le &c\exp\{\alpha\tau^2\} + \sum_{n=0}^{\infty} \exp\{\alpha t_{n+1}^2\}\,\gamma(X \setminus \overline{B}(0, t_n)).
\end{aligned}$$

Since $(\sqrt{2}^{n+2} - 1)^2 \le 2^{n+2}$ for every $n \in \mathbb{N}$,

$$\begin{aligned}
\int_X \exp\{\alpha\|x\|^2\}\,\gamma(dx) &\le c\Big(\exp\{\alpha\tau^2\} + \sum_{n=0}^{\infty} \exp\{4\alpha\tau^2(1+\sqrt{2})^2 2^n\}\Big(\frac{1-c}{c}\Big)^{2^n}\Big)\\
&= c\Big(\exp\{\alpha\tau^2\} + \sum_{n=0}^{\infty} \exp\Big\{2^n\Big(\log\frac{1-c}{c} + 4\alpha\tau^2(1+\sqrt{2})^2\Big)\Big\}\Big)\\
&= c\Big(\exp\{\alpha\tau^2\} + \sum_{n=0}^{\infty} \exp\Big\{2^n\Big(\frac{1}{2} - \frac{\sqrt{2}}{3}\Big)\log\Big(\frac{1-c}{c}\Big)\Big\}\Big).
\end{aligned}$$

The last series is convergent because $c > 1/2$ and hence $\log\left(\frac{1-c}{c}\right) < 0$. $\square$

The validity of the Fernique Theorem can be extended to any Gaussian measure, not necessarily centred.

Corollary 2.3.2. *Let γ be a Gaussian measure on a separable Banach space X. Then there exists $\alpha > 0$ such that*

$$\int_X \exp\{\alpha\|x\|^2\}\,\gamma(dx) < \infty.$$

Proof. Let us set $\mu(B) = \gamma(-B)$ for any $B \in \mathscr{B}(X)$. According to (2.2.5), the measure $\gamma_1 = \gamma * \mu$ is given by $(\gamma\otimes\mu)\circ h^{-1}$, $h(x,y) = x+y$, and it is a centred Gaussian measure. Therefore, there exists $\alpha_1 > 0$ such that

$$\begin{aligned}\infty > \int_X \exp\{\alpha_1\|x\|^2\}\,\gamma_1(dx) &= \int_{X\times X} \exp\{\alpha_1\|x+y\|^2\}\,(\gamma\otimes\mu)(d(x,y))\\ = \int_{X\times X} \exp\{\alpha_1\|x-y\|^2\}\,(\gamma\otimes\gamma)(d(x,y)) &= \int_X \gamma(dy)\int_X \exp\{\alpha_1\|x-y\|^2\}\,\gamma(dx).\end{aligned}$$

Then, for γ–a.e. $y \in X$,

$$\int_X \exp\{\alpha_1\|x-y\|^2\}\,\gamma(dx) < \infty.$$

Using the inequality $2ab \le \varepsilon a^2 + \frac{1}{\varepsilon}b^2$, which holds for any a, b and $\varepsilon > 0$, we get

$$\|x\|^2 \le \|x-y\|^2 + \|y\|^2 + 2\|x-y\|\|y\| \le (1+\varepsilon)\|x-y\|^2 + \Big(1+\frac{1}{\varepsilon}\Big)\|y\|^2.$$

For any $\alpha \in (0,\alpha_1)$, by setting $\varepsilon = \frac{\alpha_1}{\alpha} - 1$, we obtain

$$\int_X \exp\{\alpha\|x\|^2\}\,\gamma(dx) \le \exp\left\{\frac{\alpha\alpha_1}{\alpha_1-\alpha}\|y\|^2\right\}\int_X \exp\{\alpha_1\|x-y\|^2\}\,\gamma(dx).$$

Then for any $\alpha < \alpha_1$,

$$\int_X \exp\{\alpha\|x\|^2\}\,\gamma(dx) < \infty.$$

□

As a first application of the Fernique theorem, we notice that for every $1 \le p < \infty$ we have

$$\int_X \|x\|^p\,\gamma(dx) < \infty \tag{2.3.2}$$

since $\|x\|^p \le c_{\alpha,p}\exp\{\alpha\|x\|^2\}$ for all $x \in X$ and for some constant $c_{\alpha,p}$ depending on α and p only. We already know, through the definition of Gaussian measure, that the functions $f \in X^*$ belong to all $L^p(X,\gamma)$ spaces, for $1 \le p < \infty$. The Fernique Theorem tells us much more, since it gives a rather precise description of the allowed growth of the functions in $L^p(X,\gamma)$. Moreover, estimate (2.3.2) has important consequences on the functions a_γ and B_γ.

Proposition 2.3.3. *If γ is a Gaussian measure on a separable Banach space X, then $a_\gamma : X^* \to \mathbb{R}$ and $B_\gamma : X^*\times X^* \to \mathbb{R}$ are continuous. In addition, there exists $a \in X$ representing a_γ, i.e., such that*

$$a_\gamma(f) = f(a), \qquad \forall f \in X^*.$$

Proof. Let us define

$$c_1 := \int_X \|x\|\,\gamma(dx), \qquad c_2 := \int_X \|x\|^2\,\gamma(dx). \tag{2.3.3}$$

Then, if $f, g \in X^*$

$$\begin{aligned}
|a_\gamma(f)| &\le \|f\|_{X^*} \int_X \|x\|\,\gamma(dx) = c_1\|f\|_{X^*},\\
|B_\gamma(f,g)| &\le \int_X |f(x) - a_\gamma(f)||g(x) - a_\gamma(g)|\,\gamma(dx)\\
&\le \|f\|_{X^*}\|g\|_{X^*} \int_X (\|x\| + c_1)^2\,\gamma(dx) = (c_2 + 3c_1^2)\|f\|_{X^*}\|g\|_{X^*}.
\end{aligned}$$

To show that a_γ can be represented by an element $a \in X$, by the general duality theory (see e.g. [1, VIII, Théorème 8]), it is enough to show that the map $f \mapsto a_\gamma(f)$ is weakly* continuous on X^*, i.e., continuous with respect to the duality $\sigma(X^*, X)$. Moreover, since X is separable, by [4, Theorem 3.28] weak* continuity is equivalent to continuity along weak* convergent sequences.

Let (f_j) be a sequence weakly* convergent to f, i.e.,

$$f_j(x) \to f(x), \qquad \forall x \in X.$$

By the Uniform Boundedness Principle,

$$\sup_{j\in\mathbb{N}} \|f_j\|_{X^*} < \infty,$$

and by the Lebesgue Dominated Convergence Theorem, we deduce

$$\lim_{j\to\infty} a_\gamma(f_j) = \lim_{j\to\infty} \int_X f_j(x)\,\gamma(dx) = \int_X f(x)\,\gamma(dx) = a_\gamma(f).$$

□

If X is a Banach space, it is easily seen that the space X^* is contained in $L^2(X,\gamma)$ and the inclusion map $j : X^* \to L^2(X,\gamma)$,

$$j(f) = f - a_\gamma(f), \qquad f \in X^* \tag{2.3.4}$$

is continuous because $\|j(f)\|_{L^2(X,\gamma)} \le (c_2^{1/2} + c_1)\|f\|_{X^*}$, where c_1 and c_2 are defined in (2.3.3). If X is a Fréchet space, then the range of the function j defined in (2.3.4) is still contained in $L^2(X,\gamma)$. We define the *reproducing kernel*[(1)] as the closure of the range of the map j.

[(1)] This terminology comes from the general theory of Reproducing Kernel Hilbert Spaces due to N. Aronszajn, see *Theory of Reproducing Kernels,* Trans. Amer. Math. Soc. **68** (1950), 337-404 and [29, §III.9].

Definition 2.3.4 (Reproducing kernel)**.** *The* reproducing kernel *is defined by*

$$X_\gamma^* := \text{ the closure of } j(X^*) \text{ in } L^2(X,\gamma), \tag{2.3.5}$$

i.e. X_γ^* *consists of all limits in* $L^2(X,\gamma)$ *of sequences of functions* $j(f_h) = f_h - a_\gamma(f_h)$ *with* $(f_h) \subset X^*$.

For the moment we have defined the functions a_γ, $\hat{\gamma}$ in X^* and the function B_γ in $X^* \times X^*$. Of course, using formulas (2.2.1), (2.2.2) and (2.1.1) a_γ could be defined in $L^1(X,\gamma)$, B_γ could be defined in $L^2(X,\gamma) \times L^2(X,\gamma)$, and $\hat{\gamma}$ could be defined in the space of all measurable functions f, but we are not interested to study the extensions of a_γ, B_γ and $\hat{\gamma}$ to such domains. Our attention is restricted here to X_γ^*.

The extension of a_γ to X_γ^* is trivial, since the mean value of every element of X_γ^* is zero. The extension of B_γ to $X_\gamma^* \times X_\gamma^*$ is obviously continuous ($X_\gamma^* \times X_\gamma^*$ is endowed with the $L^2(X,\gamma) \times L^2(X,\gamma)$ norm), and since $a_\gamma \equiv 0$ on X_γ^*,

$$B_\gamma(f,g) = \int_X f(x)g(x)\gamma(dx) = \langle f,g\rangle_{L^2(X,\gamma)}, \quad f,\ g \in X_\gamma^*.$$

Concerning the extension of $\hat{\gamma}$, still defined by

$$\hat{\gamma}(f) = \int_X e^{if(x)}\gamma(dx),$$

we have the following proposition.

Proposition 2.3.5. *Let* γ *be a Gaussian measure on a separable Banach space. Then,*

$$\widehat{\gamma}(f) = \exp\Big\{-\frac{1}{2}\|f\|^2_{L^2(X,\gamma)}\Big\}, \qquad \forall f \in X_\gamma^*.$$

Proof. Let $f \in X_\gamma^*$ and let $g_h = j(f_h)$, $f_h \in X^*$, be a sequence of functions converging to f in $L^2(X,\gamma)$. Then, $a_\gamma(g_h) = 0$ for all $h \in \mathbb{N}$. Using the fact that the map $t \mapsto e^{it}$ is 1-Lipschitz, we have

$$\begin{aligned}\Big|\int_X \exp\{ig_h(x)\} - \exp\{if(x)\}\,\gamma(dx)\Big| &\le \int_X |\exp\{ig_h(x)\} - \exp\{if(x)\}|\,\gamma(dx)\\ &\le \int_X |g_h(x) - f(x)|\,\gamma(dx)\\ &\le \Big(\int_X |g_h(x) - f(x)|^2\,\gamma(dx)\Big)^{1/2} \to 0.\end{aligned}$$

Therefore,

$$\hat{\gamma}(f) = \lim_{h\to\infty}\hat{\gamma}(g_h) = \lim_{h\to\infty}\exp\Big\{-\frac{1}{2}B_\gamma(g_h,g_h)\Big\} = \exp\Big\{-\frac{1}{2}\|f\|_{L^2(X,\gamma)}\Big\}.$$

$\square$

Notice that if γ is nondegenerate then two different elements of X^* define two different elements of X^*_γ, but if γ is degenerate two different elements of X^* may define elements coinciding γ-a.e.

We define the operator $R_\gamma : X^*_\gamma \to (X^*)'$ by

$$R_\gamma f(g) := \int_X f(x)[g(x) - a_\gamma(g)]\,\gamma(dx), \qquad f \in X^*_\gamma,\ g \in X^*. \tag{2.3.6}$$

Observe that

$$R_\gamma f(g) = \langle f, g - a_\gamma(g)\rangle_{L^2(X,\gamma)}. \tag{2.3.7}$$

It is important to notice that indeed R_γ maps X^*_γ into X, if X is a separable Banach space. This is true also for Fréchet (and, even more genarally, locally convex) spaces, see [3, Theorem 3.2.3], but the proof is more difficult and we do not need the result in its full generality.

Proposition 2.3.6. *Let X be a separable banach space. Then, the range of R_γ is contained in X, i.e., for every $f \in X^*_\gamma$ there is $y \in X$ such that $R_\gamma f(g) = g(y)$ for all $g \in X^*$.*

Proof. As in the proof of Proposition 2.3.3, we show that for every $f \in X^*_\gamma$ the map $g \mapsto R_\gamma f(g)$ is weakly* continuous on X^*, i.e., continuous with respect to the duality $\sigma(X^*, X)$. By the general duality theory (see e.g. [1, VIII, Théorème 8]) we deduce that $R_\gamma f \in X$. Recall that, since X is separable, weak* continuity is equivalent to continuity along weak* convergent sequences. Let then $(g_k) \subset X^*$ be weakly* convergent to g, i.e., $g_k(x) \to g(x)$ for every $x \in X$. Then, by the Uniform Boundedness Principle the sequence (g_k) is bounded in X^* and by the Dominated Convergence Theorem and Proposition 2.3.3 $a_\gamma(g_k) \to a_\gamma(g)$ and

$$R_\gamma f(g_k) = \int_X f(x)[g_k(x) - a_\gamma(g_k)]\,\gamma(dx) \quad\longrightarrow\quad \int_X f(x)[g(x) - a_\gamma(g)]\,\gamma(dx) = R_\gamma f(g).$$

□

Remark 2.3.7. Thanks to Proposition 2.3.6, we can identify $R_\gamma f$ with the element $y \in X$ representing it, i.e. we shall write

$$R_\gamma f(g) = g(R_\gamma f), \qquad \forall g \in X^*.$$

2.4 Exercises

Exercise 2.1. Prove that if γ is a Gaussian measure and γ is not a Dirac measure, then for any $r > 0$ and $x \in X$,

$$\gamma(B(x, r)) < 1.$$

Exercise 2.2. Let X be an infinite dimensional Banach space. Prove that there is no nontrivial measure μ on X invariant under translations and such that $\mu(B) > 0$ for any ball B. *Hint:* modify the construction described in the Hilbert case using a sequence of elements in the unit ball having mutual distance $1/2$.

Exercise 2.3. Prove that a centred Gaussian measure on a Banach space is degenerate iff there exists $X^* \ni f \neq 0$ such that $\hat{\gamma}(f) = 1$ and hence iff there exists a proper closed subspace $V \subset X$ with $\gamma(V) = 1$.

Exercise 2.4. Let γ be centred. Prove that for any choice $f_1, \dots, f_d \in X^*$, setting

$$P(x) = (f_1(x), \dots, f_d(x)),$$

$\gamma \circ P^{-1}$ is the Gaussian measure $\mathscr{N}(0, Q)$, with $Q_{i,j} = \langle f_i, f_j \rangle_{L^2(X,\gamma)}$. If $L : \mathbb{R}^d \to \mathbb{R}^n$ is another linear map, compute the covariance matrix of $\gamma \circ (L \circ P)^{-1}$.

Exercise 2.5. Let $\gamma = \mathscr{N}(a, Q)$ be a nondegenerate Gaussian probability measure on $\mathbb{R}^d$. Show that

$$\int_{\mathbb{R}^d} |x|^4 \, \gamma(dx) = (\mathrm{Tr}\, Q)^2 + 2\mathrm{Tr}(Q^2).$$

Hint: Consider the function $F(\varepsilon) = \int_{\mathbb{R}^d} e^{\varepsilon |x|^2} \gamma(dx)$ and compute $F''(0)$.

Exercise 2.6. Let γ be a centred Gaussian measure on a separable Banach space X. Compute the integrals

$$\int_X e^{f(x)} \, \gamma(dx), \qquad \int_X \big(f(x)\big)^k \, \gamma(dx), \; k \in \mathbb{N}$$

for every $f \in X^*$.

Lecture 3

The Cameron–Martin space

In this Lecture X is a separable Fréchet space, and we present the Cameron-Martin space. It consists of the elements $h \in X$ such that the measure $\gamma_h(B) := \gamma(B - h)$ is absolutely continuous with respect to γ. As we shall see, the Cameron-Martin space is fundamental when dealing with the differential structure in X mainly in connection with integration by parts formulas.

3.1 The Cameron–Martin space

We start with the definition of the Cameron–Martin space.

Definition 3.1.1 (Cameron-Martin space)**.** *For every $h \in X$ set*

$$|h|_H := \sup\Big\{f(h) : \ f \in X^*, \ \|j(f)\|_{L^2(X,\gamma)} \leq 1\Big\}, \tag{3.1.1}$$

where $j : X^* \to L^2(X, \gamma)$ is the inclusion defined in (2.3.4)*. The* Cameron-Martin space *is defined by*

$$H := \Big\{h \in X : \ |h|_H < \infty\Big\}. \tag{3.1.2}$$

If $X$is a Banach space, calling c the norm of $j : X^* \to L^2(X, \gamma)$, we have

$$\|h\|_X = \sup\{f(h) : \ \|f\|_{X^*} \leq 1\} \leq \sup\{f(h) : \ \|j(f)\|_{L^2(X,\gamma)} \leq c\} = c|h|_H, \tag{3.1.3}$$

and then H is continuously embedded in X. We shall see that this embedding is even compact and that the norms $\|\cdot\|_X$ and $|\cdot|_H$ are not equivalent in H, in general.

The Cameron-Martin space inherits a natural Hilbert space structure from the space X^*_γ through the $L^2(X, \gamma)$ Hilbert structure.

Proposition 3.1.2. *An element $h \in X$ belongs to H if and only if there is $\hat{h} \in X^*_\gamma$ such that $h = R_\gamma \hat{h}$. In this case,*

$$|h|_H = \|\hat{h}\|_{L^2(X,\gamma)}. \tag{3.1.4}$$

Therefore $R_\gamma : X^*_\gamma \to H$ is an isometry and H is a Hilbert space with the inner product

$$[h,k]_H := \langle \hat{h}, \hat{k}\rangle_{L^2(X,\gamma)}$$

whenever $h = R_\gamma \hat{h}$, $k = R_\gamma \hat{k}$.

Proof. If $|h|_H < \infty$, we define the linear map $L : j(X^*) \to \mathbb{R}$ setting

$$L(j(f)) := f(h), \qquad \forall f \in X^*.$$

Such map is well-defined since the estimate

$$|f(h)| \le \|j(f)\|_{L^2(X,\gamma)} |h|_H, \tag{3.1.5}$$

that comes from (3.1.1), implies that if $j(f_1) = j(f_2)$, then $f_1(h) = f_2(h)$. The map L is also continuous with respect to the L^2 topology again by estimate (3.1.5). Then L can be continuously extended to X^*_γ; by the Riesz representation theorem there is a unique $\hat{h} \in X^*_\gamma$ such that the extension (still denoted by L) is given by

$$L(\phi) = \int_X \phi(x)\hat{h}(x)\,\gamma(dx), \qquad \forall \phi \in X^*_\gamma.$$

In particular, for any $f \in X^*$,

$$f(h) = L(j(f)) = \int_X j(f)(x)\hat{h}(x)\,\gamma(dx) = f(R_\gamma \hat{h}), \tag{3.1.6}$$

therefore $R_\gamma \hat{h} = h$ and

$$|h|_H = \sup\Big\{ f(h) : \ f \in X^*, \ \|j(f)\|_{L^2(X,\gamma)} \le 1 \Big\} = \|\hat{h}\|_{L^2(X,\gamma)}.$$

Conversely, if $h = R_\gamma \hat{h}$, then for every $f \in X^*$ we have

$$f(h) = \int_X j(f)(x)\hat{h}(x)\,\gamma(dx) \le \|\hat{h}\|_{L^2(X,\gamma)} \|j(f)\|_{L^2(X,\gamma)}, \tag{3.1.7}$$

by (3.1.6), whence $|h|_H < \infty$. $\square$

The space $L^2(X,\gamma)$ (hence its subspace X^*_γ as well) is separable, because X is separable, see e.g. [4, Theorem 4.13]. Therefore, H, being isometric to a separable space, is separable.

Remark 3.1.3. The map $R_\gamma : X^*_\gamma \to X$ can be defined directly using the Bochner integral through the formula

$$R_\gamma f := \int_X (x-a) f(x)\,\gamma(dx),$$

where a is the mean of γ. We do not assume the knowledge of Bochner integral. We shall say something about it in one of the following lectures.

Before going on, let us describe the finite dimensional case $X = \mathbb{R}^d$. If $\gamma = \mathscr{N}(a, Q)$ then for $f \in \mathbb{R}^d$ we have

$$\|j(f)\|^2_{L^2(\mathbb{R}^d,\gamma)} = \int_{\mathbb{R}^d} (x-a)\cdot f\, \mathscr{N}(a,Q)(dx) = (Qf)\cdot f$$

and therefore $|h|_H$ is finite if and only if $h \in Q(\mathbb{R}^d)$ and, as a consequence, $H = Q(\mathbb{R}^d)$ is the range of Q. According to the notation introduced in Proposition 3.1.2, if γ is nondegenerate, namely Q is invertible, $h = R_\gamma \hat{h}$ iff $\hat{h}(x) = \langle Q^{-1}h, x\rangle_{\mathbb{R}^d}$. Moreover, if γ is nondegenerate the measures γ_h defined by $\gamma_h(B) = \gamma(B-h)$ are all equivalent to γ in the sense of Section 1.1 and an elementary computation shows that, writing $\gamma_h = \varrho_h\gamma$, we have

$$\varrho_h(x) := \exp\Big\{(Q^{-1}h)\cdot x - \frac{1}{2}|h|^2\Big\} = \exp\Big\{\hat{h}(x) - \frac{1}{2}|h|^2\Big\}.$$

In the infinite dimensional case the situation is completely different. We start with a preliminary result.

Lemma 3.1.4. *For any $g \in X^*_\gamma$, the measure*

$$\mu_g = \exp\Big\{g - \frac{1}{2}\|g\|^2_{L^2(X,\gamma)}\Big\}\gamma$$

is a Gaussian measure with characteristic function

$$\widehat{\mu}_g(f) = \exp\Big\{if(R_\gamma g) + ia_\gamma(f) - \frac{1}{2}\|j(f)\|^2_{L^2(X,\gamma)}\Big\}. \tag{3.1.8}$$

Proof. First of all, we notice that the image of γ under the measurable function $g : X \to \mathbb{R}$ is still a Gaussian measure given by $\mathscr{N}(0, \|g\|^2_{L^2(X,\gamma)})$ thanks to Proposition 2.3.5. Indeed,

$$\hat{\gamma}(tg) = exp\Big\{-\frac{1}{2}\|g\|^2_2 t^2\Big\} \qquad \text{for all } t \in \mathbb{R},$$

and then $\hat{\gamma}(tg)$ can also be computed by

$$\int_X \exp\{itg(x)\}\gamma(dx) = \int_{\mathbb{R}} \exp\{it\tau\}(\gamma\circ g^{-1})\,d\tau = \hat{(}\gamma\circ g^{-1})(t).$$

Therefore,

$$\int_X \exp\{|g(x)|\}\,\gamma(dx) = \int_{\mathbb{R}} e^{|t|}\,\mathscr{N}(0, \|g\|^2_{L^2(X,\gamma)})(dt) < \infty,$$

hence $\exp\{|g|\} \in L^1(X,\gamma)$ and μ_g is a finite measure. In addition, μ_g is a probability measure since

$$\begin{aligned}\mu_g(X) =& \int_X \exp\Big\{g(x) - \frac{1}{2}\|g\|^2_{L^2(X,\gamma)}\Big\}\,\gamma(dx)\\ =& \exp\Big\{-\frac{1}{2}\|g\|^2_{L^2(X,\gamma)}\Big\}\int_{\mathbb{R}} e^t \mathscr{N}(0, \|g\|^2_{L^2(X,\gamma)})(dt) = 1.\end{aligned}$$

In order to prove that (3.1.8) holds, we observe that for every $t \in \mathbb{R}$ we have

$$
\begin{aligned}
&\exp\Big\{-\frac{1}{2}\|g\|^2_{L^2(X,\gamma)}\Big\}\int_X \exp\{i(f(x)-tg(x))\}\,\gamma(dx)\\
&=\exp\Big\{-\frac{1}{2}\|g\|^2_{L^2(X,\gamma)}\Big\}\widehat{\gamma}(f-tg)\\
&=\exp\Big\{-\frac{1}{2}\|g\|^2_{L^2(X,\gamma)}\Big\}\exp\Big\{ia_\gamma(f-tg)-\frac{1}{2}\|j(f-tg)\|^2_{L^2(X,\gamma)}\Big\}\\
&=\exp\Big\{tf(R_\gamma g)-\frac{1+t^2}{2}\|g\|^2_{L^2(X,\gamma)}+ia_\gamma(f)-\frac{1}{2}\|j(f)\|^2_{L^2(X,\gamma)}\Big\}.
\end{aligned}
$$

As $\int_X \exp\{\alpha|g(x)|\}\gamma(dx) < \infty$ for all $\alpha > 0$ (which can be proved exactly as at the beginning of the proof of Lemma 3.1.4, the functions

$$
\begin{aligned}
&z \mapsto \exp\Big\{-\frac{1}{2}\|g\|^2_{L^2(X,\gamma)}\Big\}\int_X \exp\{i(f(x)-zg(x))\}\,\gamma(dx)\\
&z \mapsto \exp\Big\{zf(R_\gamma g)-\frac{1+z^2}{2}\|g\|^2_{L^2(X,\gamma)}+ia_\gamma(f)-\frac{1}{2}\|j(f)\|^2_{L^2(X,\gamma)}\Big\}
\end{aligned}
$$

are entire holomorphic and coincide for $z \in \mathbb{R}$, hence they coincide in $\mathbb{C}$. In particular, taking $z = i$ we obtain

$$
\widehat{\mu}_g(f) = \exp\Big\{ia_\gamma(f)-\frac{1}{2}\|j(f)\|^2_{L^2(X,\gamma)}+iR_\gamma g(f)\Big\}.
$$

□

Theorem 3.1.5 (Cameron-Martin Theorem)**.** *For $h \in X$, define the measure $\gamma_h(B) := \gamma(B-h)$. If $h \in H$ the measure γ_h is equivalent to γ and $\gamma_h = \varrho_h\gamma$, with*

$$
\varrho_h(x) := \exp\Big\{\hat{h}(x)-\frac{1}{2}|h|^2_H\Big\}, \tag{3.1.9}
$$

where $\hat{h} = R_\gamma^{-1}h$. If $h \notin H$ then $\gamma_h \perp \gamma$. Hence, $\gamma_h \approx \gamma$ if and only if $h \in H$.

Proof. For $h \in H$, let us compute the characteristic function of γ_h. For any $f \in X^*$ we have

$$
\begin{aligned}
\hat{\gamma}_h(f) &= \int_X \exp\{if(x)\}\,\gamma_h(dx) = \int_X \exp\{if(x+h)\}\,\gamma(dx)\\
&= \exp\Big\{if(R_\gamma\hat{h})+ia_\gamma(f)-\frac{1}{2}\|j(f)\|^2_{L^2(X,\gamma)}\Big\}, \qquad f \in X^*.
\end{aligned}
$$

Taking into account Lemma 3.1.4 and Proposition 2.1.2, we obtain $\gamma_h = \varrho_h\gamma$, where the density ϱ_h is given by (3.1.9).

Now, let us see that if $h \notin H$ then $\gamma_h \perp \gamma$. To this aim, let us first consider the 1-dimensional case. If γ is a Dirac measure in $\mathbb{R}$, then $\gamma_h \perp \gamma$ for any $h \neq 0$ and $|\gamma-\gamma_h|(\mathbb{R}) = 2$. Otherwise, if $\gamma = \mathscr{N}(a,\sigma^2)$ is a nondegenerate Gaussian measure in $\mathbb{R}$, then $\gamma_h \ll \gamma$ with $\frac{d\gamma_h}{d\gamma}(t) = \exp\{-\frac{h^2}{2\sigma^2} + \frac{h(t-a)}{\sigma^2}\}$. We can apply Theorem 1.1.12 (Hellinger) with $\lambda = \gamma$, whence by Exercise 3.2

$$H(\gamma, \gamma_h) = \exp\Big\{-\frac{h^2}{8\sigma^2}\Big\}, \tag{3.1.10}$$

and then (1.1.7) implies

$$|\gamma - \gamma_h|(\mathbb{R}) \geq 2\left(1 - \exp\Big\{-\frac{1}{8\sigma^2}h^2\Big\}\right). \tag{3.1.11}$$

In any case, (3.1.11) holds true.

Let us go back to X. For every $f \in X^*$, using just the definition, it is immediate to verify that $\gamma_h \circ f^{-1} = (\gamma \circ f^{-1})_{f(h)}$ and

$$|\gamma \circ f^{-1} - (\gamma \circ f^{-1})_{f(h)}|(\mathbb{R}) \leq |\gamma - \gamma_h|(X). \tag{3.1.12}$$

If $h \notin H$, there exists a sequence $(f_n) \subset X^*$ with $\|j(f_n)\|_{L^2(X,\gamma)} = 1$ and $f_n(h) \geq n$. By (3.1.11) and (3.1.12) we obtain

$$\begin{aligned} |\gamma - \gamma_h|(X) &\geq |(\gamma \circ f_n^{-1}) - (\gamma \circ f_n^{-1})_{f_n(h)}|(\mathbb{R}) \geq 2\left(1 - \exp\Big\{-\frac{1}{8}f_n(h)^2\Big\}\right) \\ &\geq 2\left(1 - \exp\Big\{-\frac{1}{8}n^2\Big\}\right). \end{aligned}$$

This implies that $|\gamma - \gamma_h|(X) = 2$, hence by Corollary 1.1.13, $\gamma_h \perp \gamma$. □

From now on, we denote by $B^H(0,r)$ the open ball of centre 0 and radius r in H and by $\overline{B}^H(0,r)$ its closure in H. In the proof of Theorem 3.1.9 we need the following result.

Proposition 3.1.6. *If $A \in \mathscr{B}(X)$ is such that $\gamma(A) > 0$, then there is $r > 0$ such that $B^H(0,r) \subset A - A$.*

Proof. Let us introduce the function $H \ni h \mapsto \phi(h) := \gamma\big((A+h) \cap A\big)$, i.e.,

$$\phi(h) = \int_X 1\!\!1_A(x-h) 1\!\!1_A(x)\, \gamma(dx).$$

We claim that

$$\liminf_{|h|_H \to 0} \phi(h) \geq \gamma(A).$$

Assume first that A is open. For $\varepsilon > 0$ define

$$A_\varepsilon := \{x \in A : \ \mathrm{dist}\,(x, A^c) > \varepsilon\}.$$

Then $A_\varepsilon \subset (A+h) \cap A$ for all $h \in X$ with $\|h\|_X < \varepsilon$, and therefore

$$\gamma(A_\varepsilon) \leq \liminf_{\|h\|_X \to 0} \phi(h).$$

Since $\gamma(A_\varepsilon) \to \gamma(A)$ as $\varepsilon \to 0$, one obtains $\gamma(A) \le \liminf_{\|h\|_X \to 0} \phi(h)$.

To prove the claim for every measurable set A, notice that $\lim_{|h|_H \to 0} \gamma(A+h) = \gamma(A)$. Indeed, since the image measure of γ under $\hat{h}$ is $\mathscr{N}(0, |h|_H^2)$, we have

$$\begin{aligned}|\gamma(A+h) - \gamma(A)| &\le \int_X \left|\exp\Big\{-\hat{h}(x) - \frac{1}{2}|h|_H^2\Big\} - 1\right| 1\!\!1_A(x)\,\gamma(dx) \\ &\le \int_{\mathbb{R}} \left|\exp\Big\{-t|h|_H - \frac{1}{2}|h|_H^2\Big\} - 1\right| \gamma_1(dt),\end{aligned}$$

and the right hand side vanishes as $|h|_H \to 0$ by the Dominated Convergence Theorem.

Let now A be a measurable set. We have seen in the proof of Proposition 1.1.6 that for any $\varepsilon > 0$ there exists an open set $A_\varepsilon \supset A$ such that $\gamma(A_\varepsilon \setminus A) < \varepsilon$. Therefore, for $h \in H$ we have

$$(A+h) \cap A \supseteq \big[(A_\varepsilon + h) \cap A_\varepsilon\big] \setminus \big[((A_\varepsilon \setminus A) + h) \cup (A_\varepsilon \setminus A)\big];$$

hence

$$\gamma((A+h) \cap A) \ge \gamma((A_\varepsilon + h) \cap A_\varepsilon) - \gamma((A_\varepsilon \setminus A) + h) - \gamma(A_\varepsilon \setminus A).$$

By the first part of the proof,

$$\liminf_{|h|_H \to 0} \gamma((A_\varepsilon + h) \cap A_\varepsilon) \ge \gamma(A_\epsilon)$$

and then

$$\liminf_{|h|_H \to 0} \gamma((A+h) \cap A) \ge \gamma(A_\epsilon) - 2\gamma(A_\varepsilon \setminus A) \ge \gamma(A) - 2\epsilon > 0$$

if $\varepsilon < \gamma(A)/2$. Then there is $r > 0$ such that $\phi(h) > 0$ for $|h|_H < r$ and therefore for any $|h|_H < r$, $(A+h) \cap A \ne \emptyset$, so that $B^H(0,r) \subset A - A$. □

We give the following technical result that we shall need for instance in the proof of Theorem 3.1.9; it will be rephrased with a probabilistic language in the sequel.

Lemma 3.1.7. *Let $f, g \in X^*$ and set $T : X \to \mathbb{R}^2$, $T(x) := (f(x), g(x))$. Then*

$$\gamma \circ T^{-1} = (\gamma \circ f^{-1}) \otimes (\gamma \circ g^{-1})$$

iff $j(f)$ and $j(g)$ are orthogonal in $L^2(X, \gamma)$.

Proof. We just compute the characteristic function. For every $\xi \in \mathbb{R}^2$ we have

$$\begin{aligned}\widehat{\gamma \circ T^{-1}}(\xi) &= \int_X \exp\{i\xi(T(x))\}\gamma(dx) = \int_X \exp\{i(\xi_1 f + \xi_2 g)(x)\}\gamma(dx) \\ &= \exp\left\{i\xi_1 a_\gamma(f) + i\xi_2 a_\gamma(g) - \frac{1}{2}\|j(\xi_1 f + \xi_2 g)\|^2_{L^2(X,\gamma)}\right\}.\end{aligned}$$

On the other hand, if $\mu = (\gamma \circ f^{-1}) \otimes (\gamma \circ g^{-1})$, then

$$\begin{aligned}\widehat{\mu}(\xi) &= (\widehat{\gamma \circ f^{-1}})(\xi_1)(\widehat{\gamma \circ g^{-1}})(\xi_2) \\ &= \exp\left\{ i\xi_1 a_\gamma(f) + i\xi_2 a_\gamma(g) - \frac{\xi_1^2}{2}\|j(f)\|^2_{L^2(X,\gamma)} - \frac{\xi_2^2}{2}\|j(g)\|^2_{L^2(X,\gamma)} \right\},\end{aligned}$$

whence the conclusion, since

$$\|j(\xi_1 f + \xi_2 g)\|^2_{L^2(X,\gamma)} = \xi_1^2\|j(f)\|^2_{L^2(X,\gamma)} + \xi_2^2\|j(g)\|^2_{L^2(X,\gamma)}$$

if and only if $\langle f, g\rangle_{L^2(X,\gamma)} = 0$. □

Let us show that it is always possible to consider orthonormal bases in X^*_γ made by elements of $j(X^*)$; this fact can be useful in some proofs.

Lemma 3.1.8. *There exists an orthonormal basis of X^*_γ contained in $j(X^*)$.*

Proof. Let $F = \{f_k : \ k \in \mathbb{N}\}$ be a dense sequence in $j(X^*)$, which is a pre-Hilbert space when endowed with the $L^2(X,\gamma)$ inner product. Starting from F, by the Gram–Schmidt procedure (see e.g. [18, Theorem V.2.1]), we can build an orthonormal basis of $j(X^*)$, say $\{e_k : \ k \in \mathbb{N}\}$. Then, $\{e_k : \ k \in \mathbb{N}\}$ is an orthonormal basis of X^*_γ as well. □

Theorem 3.1.9. *Let γ be a Gaussian measure in a separable Banach space X, and let H be its Cameron–Martin space. The following statements hold.*

(i) The unit ball $B^H(0,1)$ of H is relatively compact in X and hence the embedding $H \hookrightarrow X$ is compact.

(ii) If γ is centred then H is the intersection of all the Borel full measure subspaces of X.

(iii) If γ is centred and X^*_γ is infinite dimensional then $\gamma(H) = 0$.

Proof. (i) It is sufficient to prove that $B^H(0,r)$ is relatively compact in X for some $r > 0$. Fix any compact set $K \subset X$ with $\gamma(K) > 0$ (such a set exists by Proposition 1.1.6); by Lemma 3.1.6 there is $r > 0$ such that the ball $B^H(0,r)$ is contained in the compact set $K - K$, which implies that $\overline{B}^H(0,r)$ is contained in $K - K$ and the proof is complete.
(ii) Let V be a subspace of X with $\gamma(V) = 1$ and fix $h \in H$; by Theorem 3.1.5,

$$\begin{aligned}\gamma(V-h) =& \gamma_h(V) = \int_V \exp\Big\{\hat{h}(x) - \frac{1}{2}|h|^2_H\Big\}\,\gamma(dx) \\ &= \int_X \exp\Big\{\hat{h}(x) - \frac{1}{2}|h|^2_H\Big\}\,\gamma(dx) = 1.\end{aligned}$$

This implies that $h \in V$, since otherwise $V \cap (V - h) = \emptyset$ and we would have

$$1 = \gamma(X) \geq \gamma(V) + \gamma(V-h) = 2,$$

a contradiction. Therefore, $H \subset V$ for all subspaces V of full measure.

To prove that the intersection of all subspaces of X with full measure is contained in H, fixed any $h \notin H$, we construct a full measure subspace V such that $h \notin V$. If $h \notin H$, then $|h|_H = \infty$ and there is a sequence $(f_n) \subset X^*$ with $\|j(f_n)\|_{L^2(X,\gamma)} = 1$ and $f_n(h) \geq n$. Since

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \int_X |j(f_n)(x)|\, \gamma(dx) \leq \sum_{n=1}^{\infty} \frac{1}{n^2} \|j(f_n)\|_{L^2(X,\gamma)} < \infty$$

the subspace

$$V := \Big\{ x \in X : \ \text{ the series } \sum_{n=1}^{\infty} \frac{1}{n^2} j(f_n)(x) \ \text{ is convergent } \Big\} \tag{3.1.13}$$

(which is a Borel set, see Exercise 3.1) has full measure, and $h \notin V$.

(iii) Let us assume that X^*_γ is infinite dimensional. Then, there exists an orthonormal basis $\{f_n : \ n \in \mathbb{N}\} \subset X^*$ (see lemma 3.1.8) of X^*_γ; in particular for any $n \in \mathbb{N}$, $\gamma \circ f_n^{-1} = \mathscr{N}(0,1)$. For every $M > 0$ and $n \in \mathbb{N}$ we have

$$\gamma(\{x \in X : \ |f_n(x)| \leq M\}) = \mathscr{N}(0,1)(-M,M) =: a_M < 1;$$

as a consequence, since the functions f_n are mutually orthogonal, by Lemma 3.1.7 we have

$$\gamma(\{x \in X : \ |f_k(x)| \leq M \ \text{ for } k = 1, \ldots, n\}) = a_M^n \to 0 \quad \text{ as } n \to \infty$$

and then

$$\gamma\Big(\Big\{x \in X : \ \sup_{n \in \mathbb{N}} |f_n(x)| \leq M\Big\}\Big) = \gamma\Big(\bigcap_{n \in \mathbb{N}} \{x \in X : \ |f_k(x)| \leq M, k = 1, \ldots, n\}\Big) = 0.$$

Since $\{f_n : \ n \in \mathbb{N}\} \subset X^*$ is an orthonormal basis of X^*_γ, for any $h \in H$ we have

$$|h|_H^2 = \|\hat{h}\|^2_{L^2(X,\gamma)} = \sum_{n=1}^{\infty} \langle f_n, \hat{h} \rangle^2_{L^2(X,\gamma)} = \sum_{n=1}^{\infty} f_n(h)^2.$$

Therefore

$$H = \Big\{ x \in X : \ \sum_{n=1}^{\infty} f_n(x)^2 < \infty \Big\} \subset \bigcup_{M > 0} \Big\{ x \in X : \ \sup_{n \in \mathbb{N}} |f_n(x)| \leq M \Big\}$$

and it has measure 0. $\square$

We close this lecture with a couple of properties of the reproducing kernel and of the Cameron–Martin space. Then we see that the norm of the space X is somehow irrelevant in the theory, in the sense that the Cameron–Martin space remains unchanged if we replace the norm of X by a weaker norm.

Proposition 3.1.10. *Let γ be a Gaussian measure on a Banach space X. Let us assume that X is continuously embedded in another Banach space Y, i.e., there exists a continuous injection $i : X \to Y$. Then the image measure $\gamma_Y := \gamma \circ i^{-1}$ in Y is Gaussian and the Cameron–Martin space H associated with the measure γ is isomorphic to the Cameron–Martin space H_Y associated with the measure γ_Y in Y.*

Proof. Let $f \in Y^*$; then $f \circ i \in X^*$ by the continuity of the injection i. Moreover

$$a_\gamma(f \circ i) = \int_X f(i(x))\,\gamma(dx) = \int_Y f(y)\,\gamma_Y(dy) = a_{\gamma_Y}(f).$$

Denoting by $j_Y : Y^* \to Y^*_{\gamma_Y}$ the inclusion of Y^* into $L^2(Y, \gamma_Y)$, we have $j(f \circ i) = j_Y(f) \circ i$ and

$$\|j(f \circ i)\|^2_{L^2(X,\gamma)} = \int_X j(f \circ i)(x)^2\,\gamma(dx) = \int_Y j_Y(f)(y)^2\,\gamma_Y(dy) = \|j_Y(f)\|^2_{L^2(Y,\gamma_Y)}.$$

We prove now that $i : H \to H_Y$ is an isometry. First of all, $i(h) \in H_Y$ for any $h \in H$ since for any $f \in Y^*$

$$|f(i(h))| = |(f \circ i)(h)| \le \|j(f \circ i)\|_{L^2(X,\gamma)}|h|_H$$

and then

$$|i(h)|_{H_Y} = \sup\{f(i(h)) : f \in Y^*,\ \|j_Y(f)\|_{L^2(Y,\gamma_Y)} \le 1\} \le |h|_H < \infty.$$

Hence $i(H) \subset H_Y$ and

$$|i(h)|_{H_Y} \le |h|_H. \tag{3.1.14}$$

We prove now the inclusion $H_Y \subset i(H)$; since $i(X)$ has full measure in Y, we have $H_Y \subset i(X)$ by statement (ii) of Theorem 3.1.9. Then, for any $h_Y \in H_Y$, there exists a unique $h \in X$ with $i(h) = h_Y$; since

$$\gamma_Y(B - h_Y) = \gamma(i^{-1}(B) - h),$$

$h_Y \in H_Y$ if and only if $h \in H$. In this case

$$\begin{aligned}\gamma_Y(B - h_Y) &= \int_B \exp\Big\{\hat h_Y(y) - \frac{1}{2}|h_Y|^2_{H_Y}\Big\}\,\gamma_Y(dy)\\ &= \int_{i^{-1}(B)} \exp\Big\{\hat h_Y(i(x)) - \frac{1}{2}|h_Y|^2_{H_Y}\Big\}\,\gamma(dx)\end{aligned}$$

is equal to

$$\gamma(i^{-1}(B) - h) = \int_{i^{-1}(B)} \exp\Big\{\hat h(x) - \frac{1}{2}|h|^2_H\Big\}\,\gamma(dx).$$

This implies

$$\hat h_Y(i(x)) - \frac{1}{2}|h_Y|^2_{H_Y} = \hat h(x) - \frac{1}{2}|h|^2_H \tag{3.1.15}$$

for γ–a.e. $x \in X$. By (3.1.14) we obtain $\hat h_Y(i(x)) - \hat h(x) \le 0$ for γ–a.e. $x \in X$, and then, since

$$\int_X (\hat h_Y(i(x)) - \hat h(x))\,\gamma(dx) = \int_Y \hat h_Y(y)\,\gamma_Y(dy) - \int_X \hat h(x)\,\gamma(dx) = 0,$$

we conclude that $\hat h_Y(i(x)) = \hat h(x)$ for γ–a.e. $x \in X$ and then by (3.1.15) $|h_Y|_{H_Y} = |i(h)|_{H_Y} = |h|_H$. □

3.2 Exercises 3

Exercise 3.1. Prove that the space V in (3.1.13) is a Borel set.

Exercise 3.2. Show that (3.1.10) holds.

Exercise 3.3. Let γ be the measure on $\mathbb{R}^2$ defined by

$$\gamma(B) = \gamma_1(\{x \in \mathbb{R} : (x,0) \in B\}), \qquad B \in \mathscr{B}(\mathbb{R}^2).$$

Prove that the Cameron–Martin space H is given by $\mathbb{R} \times \{0\}$.

Exercise 3.4. Let γ, μ be equivalent Gaussian measures in X, and denote by H_γ, H_μ the associated Cameron–Martin spaces. Prove that for every $x \in X$, $x \in H_\gamma$ iff $x \in H_\mu$, and if in addition γ and μ are centred, then $X^*_\gamma = X^*_\mu$. Prove that if γ, μ are centred Gaussian measures in X such that $\gamma \perp \mu$, then $\gamma_x \perp \mu_y$, for all x, $y \in X$.

Exercise 3.5. Prove that the Cameron–Martin space is invariant by translation, i.e. for any $x \in X$, the measure

$$\gamma_x(B) = \gamma(B - x), \qquad \forall B \in \mathscr{B}(X)$$

has the same Cameron–Martin space as γ even if $\gamma_x \perp \gamma$.

Lecture 4

Examples

In this Lecture we present two basic examples that provide in some sense the extreme cases. The first one is $\mathbb{R}^\infty$, which is not even a Banach space, but, being a countable product of real lines, admits a canonical product Gaussian measure which generalises Proposition 1.2.4. The second example is a Hilbert space, whose richer structure allows a more detailed, though simplified, description of the framework. After presenting the relevant Gaussian measures in these spaces, we describe the Reproducing Kernel X^*_γ and the Cameron–Martin space H.

4.1 The product $\mathbb{R}^\infty$

The space $\mathbb{R}^I$ of all the real functions defined in the set I is the only example of non Banach space that is relevant in our lectures. Its importance comes both from some "universal" properties it enjoys and (even more) from the fact that it appears naturally when dealing with stochastic processes, in particular with the Brownian motion. Here, we restrict our attention to the countable case, i.e., $\mathbb{R}^\infty := \mathbb{R}^\mathbb{N}$, the space of all the real sequences. It is obviously a vector space, that we endow with a metrisable locally convex topology coming from the family of semi-norms $p_k(x) := |x_k|$, where $x = (x_k)_{k\in\mathbb{N}} \in \mathbb{R}^\infty$. A distance is defined as follows,

$$d(x,y) := \sum_{k=1}^\infty \frac{1}{2^k}\frac{|x_k - y_k|}{1 + |x_k - y_k|}.$$

It is easily seen that under this metric every Cauchy sequence is convergent, hence $\mathbb{R}^\infty$ turns out to be a *Fréchet space*. Moreover, the subspace $\mathbb{R}^\infty_c$ of finite sequences, namely the sequences (x_k) that vanish eventually, is isomorphic to the topological dual of $\mathbb{R}^\infty$, under the obvious isomorphism (see Exercise 4.1(i))

$$\mathbb{R}^\infty_c \ni (\xi_k) \mapsto f, \qquad f(x) = \sum_{k=1}^\infty \xi_k x_k \text{ (finite sum)} .$$

The elements of $\mathbb{R}^\infty_c$ with rational entries are a countable dense set, hence $\mathbb{R}^\infty$ is separable. The *cylindrical σ-algebra* $\mathscr{E}(\mathbb{R}^\infty)$ is generated by the sets of the form

$$\Big\{x \in \mathbb{R}^\infty : \ (x_1, \dots, x_k) \in B, B \in \mathscr{B}(\mathbb{R}^k)\Big\}.$$

According to Definition 1.1.10, we consider the cylindrical σ–algebra $\mathscr{E}(\mathbb{R}^\infty, F)$ with $F := \{\delta_j, \ j \in \mathbb{N}\}$, generated by the evaluations $\delta_j(x) := x_j$. By Theorem 2.1.1, it coincides with the Borel σ-algebra. According to Remark 1.1.18, we endow $\mathbb{R}^\infty$ with the product measure

$$\gamma := \bigotimes_{k\in\mathbb{N}} \gamma_1 \tag{4.1.1}$$

where γ_1 is the standard Gaussian measure. As in the Banach space case, we say that a probability measure μ in $\mathbb{R}^\infty$ is Gaussian if for every $\xi \in \mathbb{R}^\infty_c$ the measure $\mu \circ \xi^{-1}$ is Gaussian on $\mathbb{R}$. Moreover, it is easily seen that Theorem 2.2.4 holds in $\mathbb{R}^\infty$ as well, see [3, Theorem 2.2.4]. Finally, γ is obviously a Gaussian measure.

Theorem 4.1.1. *The countable product measure γ on $\mathbb{R}^\infty$ is a centred Gaussian measure. Its characteristic function is*

$$\hat\gamma(\xi) = \exp\Big\{-\frac{1}{2}\sum_{k=1}^\infty |\xi_k|^2\Big\} = \exp\Big\{-\frac{1}{2}\|\xi\|^2_{\ell^2}\Big\}, \qquad \xi \in \mathbb{R}^\infty_c, \tag{4.1.2}$$

the Reproducing Kernel is

$$X^*_\gamma = \Big\{f \in L^2(\mathbb{R}^\infty, \gamma) : \ f(x) = \sum_{k=1}^\infty \xi_k x_k, \ (\xi_k) \in \ell^2\Big\}$$

and the Cameron-Martin space H is ℓ^2.

Proof. We compute the characteristic function of γ. For $f(x) = \sum_k \xi_k x_k$, with $x \in \mathbb{R}^\infty$ and $\xi \in \mathbb{R}^\infty_c$ we have

$$\begin{aligned}
\hat\gamma(f) &= \int_{\mathbb{R}^\infty} \exp\{if(x)\}\,\gamma(dx) = \int_{\mathbb{R}^\infty} \exp\Big\{i\sum_{k=1}^\infty \xi_k x_k\Big\} \bigotimes_{k=1}^\infty \gamma_1(dx) \\
&= \prod_{k=1}^\infty \int_{\mathbb{R}} \exp\{ix_k\xi_k\}\,\gamma_1(dx_k) = \prod_{k=1}^\infty \exp\Big\{-\frac{1}{2}|\xi_k|^2\Big\} = \exp\Big\{-\frac{1}{2}\|\xi\|^2_{\ell^2}\Big\}.
\end{aligned} \tag{4.1.3}$$

According to Theorem 2.2.4, γ is a Gaussian measure with mean $a_\gamma = 0$ and covariance $B_\gamma(\xi, \xi) = \|\xi\|^2_{\ell^2}$.

Let us come to the Cameron-Martin space. Since the mean of γ is 0, for every $f = (\xi_k) \in \mathbb{R}^\infty_c$ we have

$$\begin{aligned}\|f\|^2_{L^2(X,\gamma)} &= \int_X \Big|\sum_k \xi_k x_k\Big|^2 \gamma(dx) = \int_X \Bigg(\sum_k \xi_k^2 x_k^2 + \sum_{j\neq k} \xi_j\xi_k x_j x_k\Bigg)\, \gamma(dx) \\ &= \sum_k \xi_k^2 \int_{\mathbb{R}} x_k^2\, \gamma_1(dx_k) + \sum_{j\neq k} \xi_j\xi_k \int_{\mathbb{R}} x_j\, \gamma(dx_j) \int_{\mathbb{R}} x_k\, \gamma(dx_k) \\ &= \sum_k \xi_k^2 = \|\xi\|^2_{\ell^2}\end{aligned}$$

(we recall that all sums are finite) and this shows that X^*_γ, being the closure of $j(X^*) = X^*$, consists of all the functions $f(x) = \sum_k \xi_k x_k$, with $(\xi_k) \in \ell^2$, see Exercise 4.1(ii). On the other hand, for any sequence $h = (h_k) \in \mathbb{R}^\infty$ we have

$$\begin{aligned}|h|_H &= \sup\Big\{ f(h) :\ f \in X^*,\ \|j(f)\|_{L^2(X,\gamma)} \le 1\Big\} \\ &= \sup\Big\{\sum_k \xi_k h_k :\ \xi \in \mathbb{R}^\infty_c,\ \sum_k |\xi_k|^2 \le 1\Big\} = \|h\|_{\ell^2},\end{aligned}$$

and then H coincides with ℓ^2. $\square$

Remark 4.1.2. It is possible to consider more general Gaussian measures on $\mathbb{R}^\infty$, i.e.,

$$\mu = \bigotimes_{k=1}^\infty \mathscr{N}(a_k, \lambda_k). \tag{4.1.4}$$

In this case,

$$\hat{\mu}(\xi) = \exp\Big\{ i \sum_{k=1}^\infty \xi_k a_k - \frac{1}{2}\sum_{k=1}^\infty \lambda_k \xi_k^2\Big\}, \qquad \xi = (\xi_k) \in \mathbb{R}^\infty_c, \tag{4.1.5}$$

where as before all the sums contain a finite number of nonzero terms. Let us show that if $(a_k) \in \ell^2$ and $\sum_k \lambda_k < \infty$ then μ is concentrated on ℓ^2, i.e., $\mu(\ell^2) = 1$. Indeed

$$\int_{\mathbb{R}^\infty} \sum_{k=1}^\infty |x_k|^2\, \mu(dx) = \sum_{k=1}^\infty \int_{\mathbb{R}} |x_k|^2\, \mathscr{N}(a_k, \lambda_k)(dx_k) = \sum_{k=1}^\infty \lambda_k + \sum_{k=1}^\infty |a_k|^2 < \infty.$$

Then $\|x\|_{\ell^2} < \infty$ μ-a.e in $\mathbb{R}^\infty$, hence $\mu(\ell^2) = 1$.

4.2 The Hilbert space case

Let X be an infinite dimensional separable Hilbert space, with norm $\|\cdot\|_X$ and inner product $\langle\cdot,\cdot\rangle_X$. As usual, we identify X^* with X via the Riesz representation.

We say that an operator $L \in \mathcal{L}(X)$ is nonnegative if $\langle Lx, x\rangle_X \ge 0$ for all $x \in X$. We also recall that an operator $L \in \mathcal{L}(X)$ is compact if (and only if) L is the limit in the operator norm of a sequence of finite rank operators. Then, we have the following result.

Proposition 4.2.1. *Let L be a self-adjoint nonnegative bounded linear operator on X. If there is an orthonormal basis $\{f_k : \ k \in \mathbb{N}\}$ of X such that*

$$T := \sum_{k=1}^{\infty} \langle Lf_k, f_k \rangle_X < \infty \tag{4.2.1}$$

then L is compact.

Proof. Without loss of generality, assume that $\|L\|_{\mathcal{L}(X)} = 1$. For every $n \in \mathbb{N}$, let X_n be the linear span of $\{f_1, \ldots, f_n\}$. Let P_n be the orthogonal projection onto X_n and define the finite rank operator $L_n = P_n L P_n$. It is nonnegative and self-adjoint, $\|L_n\|_{\mathcal{L}(X)} \leq 1$, and the equality

$$\langle L_n f_k, f_k \rangle_X = \langle Lf_k, f_k \rangle_X$$

holds for every $k = 1, \ldots, n$. Denoting by $L_n^{1/2}$ the square root of L_n, we have $\|L_n^{1/2} f_k\|_X^2 = \langle Lf_k, f_k \rangle_X$ for $k = 1, \ldots, n$, whence

$$\sum_{k=1}^{n} \|L_n^{1/2} f_k\|_X^2 \leq T \quad \forall n \in \mathbb{N}.$$

From $\|L_n^{1/2}\|_{\mathcal{L}(X)} \leq 1$ it follows that $\|L_n f_k\|_X^2 \leq \|L_n^{1/2}\|_X^2 \|L_n^{1/2} f_k\|_X^2 \leq \|L_n^{1/2} f_k\|_X^2$ for $k = 1 \ldots, n$ and

$$\sum_{k=1}^{n} \|L_n f_k\|_X^2 \leq \sum_{k=1}^{n} \|L_n^{1/2} f_k\|_X^2 \leq \mathrm{tr}(L) \quad \forall n \in \mathbb{N}.$$

Notice that for $k = 1, \ldots, m$ and $n \geq m$ we have $L_n f_k = P_n L f_k$. therefore, for every $m \in \mathbb{N}$ and $\lim_{n\to\infty} L_n f_k = L f_k$. Therefore, for every $m \in \mathbb{N}$ we have

$$\sum_{k=1}^{m} \|Lf_k\|_X^2 = \lim_{n\to\infty} \sum_{k=1}^{m} \|L_n f_k\|_X^2 \leq T$$

and then letting $m \to \infty$

$$\sum_{k=1}^{\infty} \|Lf_k\|_X^2 \leq T. \tag{4.2.2}$$

Using (4.2.2) we prove that L is compact. Let

$$x_n = \sum_{k=1}^{\infty} \langle x_n, f_k \rangle_X f_k \quad \longrightarrow \quad 0 \qquad \text{weakly}.$$

Then, (x_n) is bounded, say $\|x_n\|_X \leq M$ for any $n \in \mathbb{N}$. Moreover,

$$Lx_n = \sum_{k=1}^{\infty} \langle x_n, f_k \rangle_X L f_k$$

whence for every $N \in \mathbb{N}$ we have

$$\begin{aligned}\|Lx_n\|_X &\le \sum_{k=1}^{\infty} |\langle x_n, f_k\rangle_X| \|Lf_k\|_X = \sum_{k=1}^{N} |\langle x_n, f_k\rangle_X| \|Lf_k\|_X + \sum_{k=N+1}^{\infty} |\langle x_n, f_k\rangle_X| \|Lf_k\|_X \\ &\le \sum_{k=1}^{N} |\langle x_n, f_k\rangle_X| + M\Big(\sum_{k=N+1}^{\infty} \|Lf_k\|_X^2\Big)^{1/2}\end{aligned}$$

and for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that the second term is $< \varepsilon$ because the series in (4.2.2) is convergent. Once this N has been fixed, there is $\nu > 0$ such that the first term is $< \varepsilon$ for $n > \nu$ by the weak convergence of the sequence (x_n) to 0. So, (Lx_n) converges in X and L is compact. □

Let us recall that if L is a compact operator on X, the spectrum of L is at most countable and if the spectrum is infinite it consists of a sequence of eigenvalues (λ_k) that can cluster only at 0. If L is compact and self-adjoint, there is an orthonormal basis of X consisting of eigenvectors, see e.g. [4, Theorem 6.11]. Moreover, L has the representation

$$Lx = \sum_{k=1}^{\infty} \lambda_k \langle x, e_k\rangle_X, \qquad x \in X, \tag{4.2.3}$$

where $\{e_k : k \in \mathbb{N}\}$ is an orthonormal basis of eigenvectors and $Le_k = \lambda_k e_k$ for any $k \in \mathbb{N}$. If in addition L is nonnegative, then its eigenvalues are nonnegative.

Lemma 4.2.2. *Every operator that can be written in the form*

$$Lx = \sum_{k=1}^{\infty} \alpha_k \langle x, f_k\rangle_X f_k,$$

for some orthonormal basis $\{f_k : k \in \mathbb{N}\}$ *with* $(\alpha_k) \subset \mathbb{R}$, $\lim_{k\to\infty} \alpha_k = 0$, *is compact.*

Proof. Let us show that L is the limit in the operator norm of the sequence of finite rank operators

$$L_n x = \sum_{k=1}^{n} \alpha_k \langle x, f_k\rangle_X f_k.$$

Indeed,

$$\|Lx - L_n x\|_X = \Big\|\sum_{k=n+1}^{\infty} \alpha_k \langle x, f_k\rangle_X f_k\Big\|_X \le \sup_{k>n} |\alpha_k| \|x\|_X,$$

whence $\|L - L_n\|_{\mathcal{L}(X)} \le \sup_{k>n} |\alpha_k| \to 0$ because $\alpha_k \to 0$ as $k \to \infty$. □

It follows that if L is a nonnegative self-adjoint compact operator and $\{e_k : k \in \mathbb{N}\}$ is an orthonormal basis of eigenvectors with $Le_k = \lambda_k e_k$. We may define the *square root of* L by

$$L^{1/2} x = \sum_{k=1}^{\infty} \lambda_k^{1/2} \langle x, e_k\rangle_X e_k.$$

The operator $L^{1/2}$ is obviously self-adjoint, and it is also compact by Lemma 4.2.2.

Let us show that if $L \in \mathcal{L}(X)$ verifies the hypotheses of Proposition 4.2.1 then the sum in (4.2.1) is independent of the basis. Indeed, if $\{f_k : \ k \in \mathbb{N}\}$ satisfies (4.2.1) and $\{e_n : \ n \in \mathbb{N}\}$ a the basis of eigenvectors, we have

$$\begin{aligned}\sum_{k=1}^{\infty}\langle Lf_k, f_k\rangle_X &= \sum_{k=1}^{\infty}\Big\langle L\Big(\sum_{n=1}^{\infty}\langle f_k, e_n\rangle_X e_n\Big), \sum_{m=1}^{\infty}\langle f_k, e_m\rangle_X e_m\Big\rangle_X \\ &= \sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\Big\langle\langle f_k, e_n\rangle_X Le_n, \langle f_k, e_m\rangle_X e_m\Big\rangle_X \\ &= \sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\lambda_n\Big\langle\langle f_k, e_n\rangle_X e_n, \langle f_k, e_m\rangle_X e_m\Big\rangle_X \\ &= \sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\lambda_n\langle f_k, e_n\rangle_X^2 = \sum_{n=1}^{\infty}\lambda_n.\end{aligned}$$

The same computation with *any* orthonormal basis $\{g_k : \ k \in \mathbb{N}\}$ shows that the sum is independent of the basis and is finite for any basis if it is finite for one. So, we may define the *trace-class* operators.

Definition 4.2.3 (Trace-class operators)**.** *A nonnegative self-adjoint operator $L \in \mathcal{L}(X)$ is of* trace-class *or* nuclear *if there is an orthonormal basis $\{e_k : \ k \in \mathbb{N}\}$ of X such that*

$$\sum_{k=1}^{\infty}\langle Le_k, e_k\rangle_X < \infty$$

and the trace of L *is*

$$\mathrm{tr}(L) := \sum_{k=1}^{\infty}\langle Le_k, e_k\rangle_X \tag{4.2.4}$$

for any orthonormal basis $\{e_k : \ k \in \mathbb{N}\}$ of X.

For a complete treatment of the present matter we refer e.g. to [10, §VI.5], [11, §§XI.6, XI.9].

Let γ be a Gaussian measure in X. According to Theorem 2.2.4 and (2.2.1), (2.2.2) we have

$$\hat{\gamma}(f) = \exp\Big\{ia_\gamma(f) - \frac{1}{2}B_\gamma(f,f)\Big\}, \quad f \in X^*,$$

where the linear mapping $a_\gamma : X^* \to \mathbb{R}$ and the bilinear symmetric mapping $B_\gamma : X^* \times X^* \to \mathbb{R}$ are continuous by Proposition 2.3.3. Then, there are $a \in X$ and a self-adjoint $Q \in \mathcal{L}(X)$ such that $a_\gamma(f) = \langle f, a\rangle_X$ and $B_\gamma(f,g) = \langle Qf, g\rangle_X$ for every $f, \ g \in X^* = X$ (see Exercise 4.2). So,

$$\langle Qf, g\rangle_X = \int_X \langle f, x-a\rangle_X\langle g, x-a\rangle_X\gamma(dx), \quad f, \ g \in X, \tag{4.2.5}$$

and

$$\hat{\gamma}(f) = \exp\Big\{i\langle f, a\rangle_X - \frac{1}{2}\langle Qf, f\rangle_X\Big\}, \quad f \in X. \tag{4.2.6}$$

We denote by $\mathscr{N}(a, Q)$ the Gaussian measure γ whose Fourier transform is given by (4.2.6). As in finite dimension, a is called the mean and Q is called the covariance of γ.

The following theorem is analogous to Theorem 2.2.4, but there is an important difference. In Theorem 2.2.4 a measure is given and we give a criterion to see if it is Gaussian. Instead, in Theorem 4.2.4 we characterise all Gaussian measures in X.

Theorem 4.2.4. *If γ is a Gaussian measure on X then its characteristic function is given by* (4.2.6)*, where $a \in X$ and Q is a self-adjoint nonnegative trace-class operator. Conversely, for every $a \in X$ and for every nonnegative self-adjoint trace-class operator Q, the function $\hat{\gamma}$ in* (4.2.6) *is the characteristic function of a Gaussian measure with mean a and covariance operator Q.*

Proof. Let γ be a Gaussian measure and let $\hat{\gamma}$ be its characteristic function, given by (4.2.6). The vector a is the mean of γ by definition, and the symmetry of Q follows from the fact that the bilinear form B_γ is symmetric. By (4.2.5), Q is nonnegative, and for every orthonormal basis $\{e_k : k \in \mathbb{N}\}$ we have

$$\sum_{k=1}^{\infty}\langle Qe_k, e_k\rangle_X = \sum_{k=1}^{\infty}\int_X \langle x - a, e_k\rangle_X^2\,\gamma(dx) = \int_X \|x - a\|_X^2\,\gamma(dx)$$

which is finite by Corollary 2.3.2. Therefore, Q is a trace-class operator.

Conversely, let Q be a self-adjoint nonnegative trace-class operator. Then Q is given by

$$Qx = \sum_{k=1}^{\infty}\lambda_k\langle x, e_k\rangle_X e_k,$$

where $\lambda_k \geq 0$ for all $k \in \mathbb{N}$, $\sum_k \lambda_k < \infty$ and $\{e_k : \ k \in \mathbb{N}\}$ is an orthonormal basis of eigenvectors such that $Qe_k = \lambda_k e_k$ for any $k \in \mathbb{N}$. Let us consider the measure μ on $\mathbb{R}^\infty$ defined by (4.1.4) and its characteristic function,

$$\hat{\mu}(\xi) = \exp\Big\{i\sum_{k=1}^{\infty}\xi_k a_k - \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k|\xi_k|^2\Big\}, \qquad \xi \in \mathbb{R}^\infty_c$$

(recall that the series contains only a finite number of nonzero elements). Let $u : \ell^2 \to X$ be defined by $u(y) = \sum_{k=1}^{\infty} y_k e_k$ (and extended arbitrarily in the μ-negligible set $\mathbb{R}^\infty \setminus \ell^2$, see Remark 4.1.2). Let us show that $\gamma := \mu \circ u^{-1}$ and let us prove that $\gamma = \mathscr{N}(a, Q)$ by computing its characteristic function. For $x \in \mathbb{R}^\infty_c$, setting $z = u(y)$ we have

$$\begin{aligned}
(\widehat{\mu \circ u^{-1}})(x) &= \int_X \exp\{i\langle z, \sum_{k=1}^\infty x_k e_k\rangle_X\}\,(\mu \circ u^{-1})(dz)\\
&= \int_{\mathbb{R}^\infty} \exp\Big\{i\sum_{k=1}^\infty y_k x_k\Big\}\,\mu(dy)\\
&= \int_{\ell^2} \exp\Big\{i\sum_{k=1}^\infty x_k y_k\Big\}\,\bigotimes_{k=1}^\infty \mathscr{N}(a_k,\lambda_k)(dy)\\
&= \exp\Big\{i\sum_{k=1}^\infty x_k a_k - \frac{1}{2}\sum_{k=1}^\infty \lambda_k x_k^2\Big\}\\
&= \exp\Big\{i\langle x, a\rangle_X - \frac{1}{2}\langle Qx, x\rangle_X\Big\}.
\end{aligned}$$

By Theorem 2.2.4, $(\widehat{\mu \circ u^{-1}})$ is the characteristic function of a unique Gaussian measure with mean a and covariance Q. □

Remark 4.2.5. Since in infinite dimensions the identity is not a trace-class operator, the function $x \mapsto \exp\{-\frac{1}{2}\|x\|_X^2\}$ cannot be the characteristic function of any Gaussian measure on X.

As a consequence of Theorem 4.2.4 we compute the best constant in Theorem 2.3.1 (Fernique).

Proposition 4.2.6. *Let $\gamma = \mathscr{N}(a,Q)$ be a Gaussian measure on X and let (λ_k) be the sequence of the eigenvalues of Q. If γ is not a Dirac measure, the integral*

$$\int_X \exp\{\alpha\|x\|_X^2\}\,\gamma(dx)$$

is finite if and only if

$$\alpha < \inf\left\{\frac{1}{2\lambda_k} :\ \lambda_k > 0\right\}. \tag{4.2.7}$$

Proof. Let $\{e_k :\ k \in \mathbb{N}\}$ be an orthonormal basis of eigenvectors of Q, and $Qe_k = \lambda_k e_k$ for any $k \in \mathbb{N}$. For every $\alpha > 0$, we compute

$$\begin{aligned}
\int_X \exp\{\alpha\|x\|_X^2\}\,\gamma(dx) &= \int_{\mathbb{R}^\infty} \exp\Big\{\alpha\sum_{k=1}^\infty x_k^2\Big\}\,\bigotimes_{k=1}^\infty \mathscr{N}(a_k,\lambda_k)(dx)\\
&= \prod_{k=1}^\infty \int_{\mathbb{R}} \exp\{\alpha x_k^2\}\,\mathscr{N}(a_k,\lambda_k)(dx_k)\\
&= \prod_{k:\lambda_k=0} \exp\{\alpha a_k^2\} \prod_{k:\lambda_k>0} \frac{1}{\sqrt{2\pi\lambda_k}} \int_{\mathbb{R}} \exp\{\alpha x_k^2\}\exp\Big\{-\frac{1}{2\lambda_k}(x_k-a_k)^2\Big\}\,dx_k.
\end{aligned}$$

If $\alpha \geq (2\lambda_k)^{-1}$ for some $k \in \mathbb{N}$ then the integral with respect to dx_k is infinite, and the function $x \mapsto \exp\{\alpha\|x\|^2\}$ is not in $L^1(X,\gamma)$. If $\alpha < \inf\left\{\frac{1}{2\lambda_k} :\ \lambda_k > 0\right\}$ then each integral is finite and we have

$$\frac{1}{\sqrt{2\pi\lambda_k}}\int_{\mathbb{R}} \exp\{\alpha x_k^2\}\exp\Big\{-\frac{1}{2\lambda_k}(x_k-a_k)^2\Big\}\,dx_k = \exp\Big\{\frac{\alpha a_k^2}{1-2\alpha\lambda_k}\Big\}\frac{1}{\sqrt{1-2\alpha\lambda_k}} \tag{4.2.8}$$

for every $k \in \mathbb{N}$. Therefore

$$\begin{aligned}&\int_X \exp\{\alpha\|x\|_X^2\}\,\gamma(dx)\\ &= \exp\Big\{\alpha\sum_{k:\lambda_k=0} a_k^2\Big\}\exp\Big\{\alpha\sum_{k:\lambda_k>0}\frac{a_k^2}{1-2\alpha\lambda_k}\Big\}\exp\Big\{\sum_{k:\lambda_k>0}\log\Big(\frac{1}{\sqrt{1-2\alpha\lambda_k}}\Big)\Big\}.\end{aligned}$$

The convergence of the series follows from $\sum_k \lambda_k < \infty$. $\square$

Let us characterise X_γ^* and the Cameron-Martin space H. By definition, X_γ^* is the closure of $j(X^*)$ in $L^2(X,\gamma)$. For the rest of the lecture, if $\gamma = \mathscr{N}(a,Q)$ we fix an orthonormal basis $\{e_k : k \in \mathbb{N}\}$ of eigenvectors of Q such that $Qe_k = \lambda_k e_k$ for any $k \in \mathbb{N}$ and for every $x \in X$, $k \in \mathbb{N}$, we set $x_k := \langle x, e_k\rangle_X$.

Theorem 4.2.7. *Let $\gamma = \mathscr{N}(a,Q)$ be a nondegenerate Gaussian measure in X. The space X_γ^* is*

$$X_\gamma^* = \Big\{f : X \to \mathbb{R} :\ f(x) = \sum_{k=1}^\infty (x_k - a_k)z_k\lambda_k^{-1/2},\ z \in X\Big\} \tag{4.2.9}$$

and the Cameron-Martin space is the range of $Q^{1/2}$, i.e.,

$$H = \Big\{x \in X :\ \sum_{k=1}^\infty x_k^2\lambda_k^{-1} < \infty\Big\}. \tag{4.2.10}$$

For $h = Q^{1/2}z \in H$, we have

$$\hat{h}(x) = \sum_{k=1}^\infty (x_k - a_k)z_k\lambda_k^{-1/2}. \tag{4.2.11}$$

and

$$[h,k]_H = \langle Q^{-1/2}h, Q^{-1/2}k\rangle_X \qquad \forall\ h,k \in H. \tag{4.2.12}$$

Proof. Let $z \in X$. The sequence

$$f_n(x) := \sum_{k=1}^n (x_k - a_k)z_k\lambda_k^{-1/2}$$

converges in $L^2(X,\gamma)$, since for $m>n$,

$$\|f_m - f_n\|^2_{L^2(X,\gamma)} = \sum_{k=n+1}^{m} \lambda_k^{-1} z_k^2 \int_{\mathbb{R}} (x_k-a_k)^2 \mathscr{N}(a_k,\lambda_k)(dx_k) = \sum_{k=n+1}^{m} z_k^2,$$

and the limit function $f(x) = \sum_{k=1}^\infty (x_k-a_k) z_k \lambda_k^{-1/2}$ satisfies

$$\|f\|^2_{L^2(X,\gamma)} = \sum_{k=1}^{\infty} \lambda_k^{-1} z_k^2 \int_{\mathbb{R}} (x_k-a_k)^2 \mathscr{N}(a_k,\lambda_k)(dx_k) = \|z\|_X^2. \tag{4.2.13}$$

Moreover, for every $n\in\mathbb{N}$, $f_n = j(g_n)$, with $g_n\in X^*$,

$$g_n(x) = \sum_{k=1}^{n} \lambda_k^{-1/2} z_k x_k.$$

So, denoting by V the set in the right hand side of (4.2.9), V is contained in the closure of $j(X^*)$ in $L^2(X,\gamma)$, which is precisely X^*_γ.

Let us show that $X^*_\gamma \subset V$. Let $f\in X^*_\gamma$ and let $(w^{(n)})\subset X$ be a sequence such that $f_n(x) = \langle x-a, w^{(n)}\rangle_X$ converges to f in $L^2(X,\gamma)$. Setting $z^{(n)} := Q^{1/2}w^{(n)}$, we have $f_n(x) = \sum_{k=1}^\infty (x_k-a_k) z_k^{(n)} \lambda_k^{-1/2}$, and by (4.2.13),

$$\|z^{(n)} - z^{(m)}\|_X = \|f_n - f_m\|_{L^2(X,\gamma)}, \quad n,\ m\in\mathbb{N},$$

so that $(z^{(n)})$ is a Cauchy sequence, and it converges to some $z\in X$. Then, still by (4.2.13),

$$\begin{aligned}
&\int_X \Big(f(x) - \sum_{k=1}^{\infty} (x_k-a_k) z_k \lambda_k^{-1/2}\Big)^2 \gamma(dx)\\
&= \lim_{n\to\infty} \int_X \left(\sum_{k=1}^{\infty} (x_k-a_k)(z_k^{(n)} - z_k)\lambda_k^{-1/2}\right)^2 \gamma(dx)\\
&= \lim_{n\to\infty} \|z^{(n)} - z\|_X^2 = 0,
\end{aligned}$$

so that $f\in V$.

Let us come to the Cameron-Martin space. We know that H is the range of $R_\gamma : X^*_\gamma \to X$, and that $R_\gamma f = h$ iff $\langle f, j(g)\rangle_{L^2(X,\gamma)} = g(h)$ for every $g\in X^*$.

Given any $f\in X^*_\gamma$, $f(x) = \sum_{k=1}^\infty (x_k-a_k) z_k \lambda_k^{-1/2}$ for some $z\in X$, and $g\in X^*$, $g(x) = \sum_{k=1}^\infty g_k x_k$, we have

$$\int_X f(x) j(g)(x)\gamma(dx) = \int_X \sum_{k=1}^{\infty} (x_k-a_k)^2 z_k \lambda_k^{-1/2} g_k \, \gamma(dx) = \sum_{k=1}^{\infty} z_k \lambda_k^{1/2} g_k.$$

This is equal to $g(h)$ for $h = \sum_{k=1}^\infty z_k\lambda_k^{1/2} e_k$, namely $h = Q^{1/2}z$. So, by definition $R_\gamma f = Q^{1/2}z$, hence $H = Q^{1/2}(X)$ and for $h = Q^{1/2}z$ we have $\hat{h}(x) = \sum_{k=1}^\infty (x_k-a_k)\lambda_k^{-1/2} z_k$, nmely (4.2.11) holds. It implies that for every $h,k\in H$ we have $[h,k]_H = \langle \hat{h},\hat{k}\rangle_{L^2(X,\gamma)} = \langle Q^{-1/2}h, Q^{-1/2}k\rangle_X$. $\square$

4.3 Exercises

Exercise 4.1. (i) Prove that the dual space of $\mathbb{R}^\infty$ is $\mathbb{R}^\infty_c$.

(ii) Prove that the embedding $\iota : (\ell^2, \|\cdot\|_{\ell^2}) \hookrightarrow (\mathbb{R}^\mathbb{N}, d)$ is continuous and is the unique continuous extension of $\iota : (\mathbb{R}^\mathbb{N}, \|\cdot\|_{\ell^2}) \hookrightarrow (\mathbb{R}^\mathbb{N}, d)$.

Exercise 4.2. Let X be a real Hilbert space, and let $B : X \times X \to \mathbb{R}$ be bilinear, symmetric and continuous. Prove that there exists a unique self-adjoint operator $Q \in \mathcal{L}(X)$ such that $B(x, y) = \langle Qx, y\rangle_X$, for every x, $y \in X$.

Exercise 4.3. Let $L : \ell^2 \to \ell^2$ be the operator defined by $Lx = (x_2, x_1, x_4, x_3....)$, for $x = (x_1, ...x_n, ...) \in \ell^2$. Show that L is self-adjoint and $\langle Le_k, e_k\rangle_{\ell^2} = 0$ and that L is not compact.

Exercise 4.4. Check the computation of the integral in (4.2.8).

Exercise 4.5. Prove that if X is a separable Hilbert space, then $h \in R_\gamma(j(X^*))$ if and only if $h = Qx$, $x \in X$, and that in this case $|h|_H = \|Q^{1/2}x\|_X$.

Exercise 4.6. Modify the proof of Theorem 4.2.7 in order to consider also degenerate Gaussian measures.

Exercise 4.7. Show that ℓ^2 is a Borel measurable subset of $\mathbb{R}^\mathbb{N}$.

Lecture 5

The Brownian motion

In this and in the next Lecture we present a very important example: the classical Wiener space, which is to some extent the basic and main reference example of the theory. To do this, we introduce the Wiener measure and we define the Brownian motion. With these tools, in the next Lecture we shall define the stochastic integral and we shall use it to characterise the reproducing kernel X_γ^* when γ is the Wiener measure on $X = C([0,1])$, the Banach space of real valued continuous functions. For the material of this chapter we refer the reader for instance to the books [2, 9].

5.1 Some notions from Probability Theory

In this section we recall a few notions of probability theory. As in Definition 1.1.2, a probability is nothing but a positive measure $\mathbb{P}$ on a measurable (or *probability*) space $(\Omega, \mathscr{F})$ such that $\mathbb{P}(\Omega) = 1$.

Any measurable $\mathbb{R}^d$-valued function defined on a a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is called *random variable*. Usually, random variables are denoted by the last letters of the alphabet.

Using the image measure, we call $\mathbb{P} \circ X^{-1}$ the *law* of the $\mathbb{R}^d$-valued random variable $X : \Omega \to \mathbb{R}^d$. The law of a random variable is obviously a probability measure.

Given a real valued random variable $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, we denote by

$$\mathbb{E}[X] = \int_\Omega X d\mathbb{P}$$

the average or the *expectation* of X. We also define the variance of the real random variable X, in case $X \in L^2(\Omega, \mathscr{F}, \mathbb{P})$, as

$$\mathrm{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \int_\Omega (X - \mathbb{E}[X])^2 d\mathbb{P}.$$

Let us introduce the notion of *stochastic process.*

Definition 5.1.1. *A stochastic process* $(X_t)_{t\in I}$ *on a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *indexed on the interval* $[0,1]$ *is a function* $X : [0,1] \times \Omega \to \mathbb{R}$ *such that for any* $t \in [0,1]$ *the function* $X_t(\cdot) = X(t, \cdot)$ *is a random variable on* $(\Omega, \mathscr{F}, \mathbb{P})$.

We give now the notion of independence, both for sets and for functions. Notice that a measurable set is often called *event* in the present context.

Definition 5.1.2 (Independence). *Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space. Two sets or events* $A, B \in \mathscr{F}$ *are independent if*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

Two sub-σ*-algebras* $\mathscr{F}_1, \mathscr{F}_2$ *of* $\mathscr{F}$ *are independent if any set* $A \in \mathscr{F}_1$ *is independent of any set* $B \in \mathscr{F}_2$*, that is*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B), \qquad \forall A \in \mathscr{F}_1, \forall B \in \mathscr{F}_2.$$

Given a real random variable X *and* $\mathscr{F}'$ *sub-*σ*-algebra contained in* $\mathscr{F}$*, we say that* X *is independent of* $\mathscr{F}'$ *if the* σ*-algebras* $\sigma(X)$ *and* $\mathscr{F}'$ *are independent* (1)*. Two random variables* X *and* Y *are independent if* $\sigma(X)$ *and* $\sigma(Y)$ *are independent. Two stochastic processes* $(X_t)_{t\in I}$ *and* $(Y_t)_{t\in I}$ *are independent if* $\sigma(X_t)$ *and* $\sigma(Y_s)$ *are independent for any* $t, s \in I$.

One of the first properties of independence is expressed in the following

Proposition 5.1.3. *Let* X *and* Y *be two independent real random variables on* $(\Omega, \mathscr{F}, \mathbb{P})$*. If* $X, Y, X \cdot Y \in L^1(\Omega, \mathscr{F}, \mathbb{P})$*, then*

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y].$$

Proof. Splitting both X and Y in positive and negative part, it is not restrictive to assume that X and Y are nonnegative. Let us consider two sequences of simple functions $(s_i)_{i\in\mathbb{N}}, (s'_i)_{i\in\mathbb{N}} \subset \mathcal{S}_+$ such that s_i is $\sigma(X)$-measurable and s'_i is $\sigma(Y)$-measurable for any $i \in \mathbb{N}$, and such that $0 \le s_i \le X_i$, $0 \le s'_i \le Y_i$ and

$$\mathbb{E}[X] = \lim_{i\to\infty} \mathbb{E}[s_i], \qquad \mathbb{E}[Y] = \lim_{i\to\infty} \mathbb{E}[s'_i].$$

We have

$$s_i = \sum_{h=1}^{n_i} c_{i,h} \mathbb{1}_{A_{i,h}}, \quad s'_i = \sum_{h=1}^{m_i} c'_{i,k} \mathbb{1}_{A'_{i,k}}$$

with $A_{i,h} \in \sigma(X)$, $A'_{i.k} \in \sigma(Y)$. Then $(s_i \cdot s'_i)_{i\in\mathbb{N}}$ is a sequence of simple functions converging

(1) We recall that $\sigma(X)$ is the σ–algebra generated by the sets $\{\omega \in \Omega : X(\omega) < a\}$ with $a \in \mathbb{R}$.

to $X \cdot Y$ and then, by independence

$$
\begin{aligned}
\mathbb{E}[X \cdot Y] = \lim_{i\to\infty} \mathbb{E}[s_i \cdot s_i'] &= \lim_{i\to\infty} \sum_{h=1}^{n_i}\sum_{k=1}^{m_i} c_{i,h}c'_{i,k}\mathbb{E}[\mathbb{1}_{A_{i,h}} \cdot \mathbb{1}_{A'_{i,k}}] \\
&= \lim_{i\to\infty} \sum_{h=1}^{n_i}\sum_{k=1}^{m_i} c_{i,h}c'_{i,k}\mathbb{P}(A_{i,h} \cap A'_{i,k}) \\
&= \lim_{i\to\infty} \sum_{h=1}^{n_i}\sum_{k=1}^{m_i} c_{i,h}c'_{i,k}\mathbb{P}(A_{i,h}) \cdot \mathbb{P}(A'_{i,k}) \\
&= \lim_{i\to\infty} \sum_{h=1}^{n_i} c_{i,h}\mathbb{P}(A_{i,h}) \sum_{k=1}^{m_i} c'_{i,k}\mathbb{P}(A'_{i,k}) \\
&= \lim_{i\to\infty} \mathbb{E}[s_i] \cdot \mathbb{E}[s_i'] = \mathbb{E}[X] \cdot \mathbb{E}[Y].
\end{aligned}
$$

□

Noticing that $f(X)$ and $g(Y)$ are independent if X and Y are independent and $f, g : \mathbb{R} \to \mathbb{R}$ are Borel functions, the following corollary is immediate.

Corollary 5.1.4. *Let X and Y be two independent real random variables on $(\Omega, \mathscr{F}, \mathbb{P})$ and let $f, g : \mathbb{R} \to \mathbb{R}$ be two Borel functions. If $f(X), g(Y), f(X) \cdot g(Y) \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, then*

$$\mathbb{E}[f(X) \cdot g(Y)] = \mathbb{E}[f(X)] \cdot \mathbb{E}[g(Y)].$$

Remark 5.1.5. Using Corollary 5.1.4, it is possible to prove that two random variables X and Y are independent if and only if $\mathbb{P}\circ(X, Y)^{-1} = (\mathbb{P}\circ X^{-1})\otimes(\mathbb{P}\circ Y^{-1})$, see Exercise 5.1. As a consequence, Lemma 3.1.7 can be rephrased saying that for every $f,\ g \in X^*$, the elements $j(f), j(g)$ are orthogonal in X^*_γ iff they are independent.

5.2 The Wiener measure $\mathbb{P}^W$ and the Brownian motion

We start by considering the space $\mathbb{R}^{[0,1]}$, the set of all real valued functions defined on $[0, 1]$. We introduce the σ–algebra $\mathscr{F}$ generated by the sets

$$\{\omega \in \mathbb{R}^{[0,1]} : P_F(\omega) \in B\},$$

where $F = \{t_1, \dots, t_m\}$ is any finite set contained in $[0, 1]$, $B \in \mathscr{B}(\mathbb{R}^m)$ and $P_F : \mathbb{R}^{[0,1]} \to \mathbb{R}^m$ is defined by

$$P_F(\omega) = (\omega(t_1), \dots, \omega(t_m)).$$

We denote by $\mathscr{C}_F$ the σ–algebra $P_F^{-1}(\mathscr{B}(\mathbb{R}^m))$, and we define a measure μ_F on $\mathscr{C}_F$ by setting, in the case $0 < t_1 < \ldots < t_m$

$$\mu_F(A) = \frac{1}{(2\pi)^{\frac{m}{2}}\sqrt{t_1(t_2 - t_1)\cdot \ldots \cdot (t_m - t_{m-1})}} \int_{P_F(A)} e^{-\frac{x_1^2}{2t_1} + \ldots - \frac{(x_m - x_{m-1})^2}{2(t_m - t_{m-1})}}\, dx;$$

in the case $0 = t_1 < \ldots < t_m$

$$\mu_F(A) = \frac{1}{(2\pi)^{\frac{m-1}{2}}\sqrt{t_2 \cdot \ldots \cdot (t_m - t_{m-1})}} \int_{(P_F(A))_0} e^{-\frac{x_2^2}{2t_2} + \ldots - \frac{(x_m - x_{m-1})^2}{2(t_m - t_{m-1})}} dx'$$

where

$$(P_F(A))_0 = \{x' \in \mathbb{R}^{m-1} : (0, x') \in P_F(A)\}.$$

For $F = \{0\}$, we set $\mu_{\{0\}} = \delta_0$, the Dirac measure at 0. In this way we have defined a family of measures μ_F on the σ–algebras $\mathscr{C}_F$.

We shall use the following result to extend the family of measure μ_F to a unique probability measure on $(\mathbb{R}^{[0,1]}, \mathscr{F})$. It is known as the Daniell–Kolmogorov extension theorem, that we present only in the version we need for our purposes. Its proof relies on the following basic results.

Proposition 5.2.1. *Let μ be a nonnegative real valued finitely additive set function on an algebra $\mathscr{A}$. Then μ is countably additive on $\mathscr{A}$ if and only if it is continuous at $\emptyset$, i.e.*

$$\lim_{n\to\infty} \mu(A_n) = 0$$

for every decreasing sequence of sets $(A_n) \subset \mathscr{A}$ such that $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$.

Theorem 5.2.2. *Let μ be a nonnegative real valued countably additive set function on an algebra $\mathscr{A}$. Then μ may be extended to a unique finite measure on the σ-algebra generated by $\mathscr{A}$.*

The proof of Proposition 5.2.1 is left as an exercise, see Exercise 5.2. For Theorem 5.2.2 we refer to [9, Theorems 3.1.4, 3.1.10]).

Theorem 5.2.3 (Daniell–Kolmogorov extension)**.** *There exists a unique probability measure $\mathbb{P}^W$, called the* Wiener measure *on $(\mathbb{R}^{[0,1]}, \mathscr{F})$ such that for every finite $F \subset [0,1]$, $\mathbb{P}^W(A) = \mu_F(A)$ if $A \in \mathscr{C}_F$.*

Proof. We notice that if $F' = F \cup \{t_{m+1}\}$ with $t_m < t_{m+1} \le 1$, then for any $B \in \mathscr{B}(\mathbb{R}^m)$, $P_F^{-1}(B) = P_{F'}^{-1}(B \times \mathbb{R})$, so that

$$\mu_F(P_F^{-1}(B)) = \mu_{F'}(P_{F'}^{-1}(B \times \mathbb{R})).$$

This argument can be generalised to the case $F \subset G \subset [0,1]$, F and G finite sets with cardinality m and n respectively, to conclude that if $A = P_F^{-1}(B) = P_G^{-1}(B')$, $B \in \mathscr{B}(\mathbb{R}^m)$, $B' \in \mathscr{B}(\mathbb{R}^n)$, then $\mu_F(A) = \mu_G(A)$. So, for $A \in \mathscr{C}_F$, we can set

$$\mathbb{P}^W(A) := \mu_F(A).$$

The set function $\mathbb{P}^W$ is defined on the algebra

$$\mathscr{A} = \bigcup_{F\subset[0,1] \text{ finite}} \mathscr{C}_F;$$

it is finitely additive since if $A \in \mathscr{C}_F$ and $B \in \mathscr{C}_G$ are two disjoint sets, $A \cup B \in \mathscr{C}_{F\cup G}$ and

$$\begin{aligned}\mathbb{P}^W(A \cup B) &= \mu_{F\cup G}(A \cup B) = \mu_{F\cup G}(A) + \mu_{F\cup G}(B) = \mu_F(A) + \mu_G(B)\\ &= \mathbb{P}^W(A) + \mathbb{P}^W(B).\end{aligned}$$

Moreover, $\mathbb{P}^W(\mathbb{R}^{[0,1]}) = 1$. To extend $\mathbb{P}^W$ to the σ–algebra $\mathscr{F}$, we apply Proposition 5.2.1 and Theorem 5.2.2. Let us prove that $\mathbb{P}^W$ is continuous at $\emptyset$. Assume by contradiction that there are $\varepsilon > 0$ and a decreasing sequence $(A_n) \subset \mathscr{A}$ of sets whose intersection is empty, such that

$$\mathbb{P}^W(A_n) > \varepsilon, \qquad \forall n \in \mathbb{N}.$$

Without loss of generality, we may assume that $A_n = P_{F_n}^{-1}(B_n)$ with F_n containing n points and $B_n \in \mathscr{B}(\mathbb{R}^n)$, and that $F_n \subset F_{n+1}$. Denote by $\pi_n : \mathbb{R}^{n+1} \to \mathbb{R}^n$ the projection such that $\pi_n \circ P_{F_{n+1}} = P_{F_n}$. Since each measure $\mu_{F_n} \circ P_{F_n}^{-1}$ is a Radon measure in $\mathbb{R}^n$, for every $n \in \mathbb{N}$ there is a compact set $K_n \subset B_n$ such that $\mathbb{P}^W(A_n \setminus C_n) < \frac{\varepsilon}{2^n}$, where $C_n = P_{F_n}^{-1}(K_n)$. By replacing K_{n+1} by $\tilde{K}_{n+1} = K_{n+1} \cap \pi_n^{-1}(K_n)$, we get $\tilde{K}_{n+1} \subset \pi_n^{-1}(\tilde{K}_n)$. In order to see that the $\tilde{K}_n$ are nonempty, we bound from below their measure. Setting as before $\tilde{C}_n = P_{F_n}^{-1}(\tilde{K}_n)$, we have

$$\begin{aligned}\mu_{F_n} \circ P_{F_n}^{-1}(\tilde{K}_n) &= \mathbb{P}^W(\tilde{C}_n) = \mathbb{P}^W(A_n) - \mathbb{P}^W(A_n \setminus \tilde{C}_n)\\ &\geq \mathbb{P}^W(A_n) - \mathbb{P}^W\Big(\bigcup_{k=1}^n A_n \setminus C_k\Big) \geq \mathbb{P}^W(A_n) - \mathbb{P}^W\Big(\bigcup_{k=1}^n A_k \setminus C_k\Big)\\ &\geq \varepsilon - \sum_{k=1}^n \frac{\varepsilon}{2^k} > 0.\end{aligned}$$

Therefore, for any $n \in \mathbb{N}$ we can pick an element

$$x^{(n)} = (x_1^{(n)}, \ldots, x_n^{(n)}) \in \tilde{K}_n.$$

Since $\tilde{K}_n \subset \pi_{n-1}^{-1}(\tilde{K}_{n-1})$, the sequence $(x_1^{(n)})$ is contained in $\tilde{K}_1$, there is a subsequence $(x_1^{(k_n)}) \subset \tilde{K}_1$ converging to $y_1 \in \tilde{K}_1$. The sequence $(x_1^{(k_n)}, x_2^{(k_n)})$ is contained in $\tilde{K}_2$, then up to subsequences, there exists y_2 such that it converges to $(y_1, y_2) \in \tilde{K}_2$. Iterating the procedure, and taking the diagonal sequence, we obtain a sequence (y_n) such that

$$(y_1, \ldots, y_n) \in \tilde{K}_n, \qquad \forall n \in \mathbb{N}.$$

Then

$$P_{F_n}^{-1}(\{(y_1, \ldots, y_n)\}) \subset \tilde{C}_n \subset A_n, \qquad \forall n \in \mathbb{N},$$

hence

$$S := \{\omega \in \mathbb{R}^{[0,1]} : \ \omega(t_j) = y_j \ \forall j \in \mathbb{N}\} \subset \bigcap_{n=1}^\infty A_n$$

which is a contradiction, as $S \neq \emptyset$. Therefore, $\mathbb{P}^W$ is continuous at $\emptyset$. By Proposition 5.2.1, $\mathbb{P}^W$ is countably additive, and by Theorem 5.2.2 it has a unique extension (still denoted by $\mathbb{P}^W$) to the σ-algebra $\mathscr{F}$ generated by $\mathscr{A}$. $\square$

Once the Wiener measure has been defined, we give a formal definition of the Brownian motion.

Definition 5.2.4 (Standard Brownian motion)**.** *A real valued standard Brownian motion on* $[0,1]$ *is a stochastic process* $(B_t)_{t\in[0,1]}$ *on a probability space* $(\Omega,\mathscr{F},\mathbb{P})$ *such that:*

1. $B_0 = 0$ *almost surely, i.e.* $\mathbb{P}$*–a.e.;*
2. *for any* $t,s\in[0,1]$*,* $s<t$*, the law of both random variables* $B_t - B_s$ *and* B_{t-s} *is equal to* $\mathscr{N}(0,t-s)$*;*
3. *for any* $0\le t_0\le t_1\le\ldots\le t_n$ *the random variables* $B_{t_0}, B_{t_1}-B_{t_0},\ldots,B_{t_n}-B_{t_{n-1}}$ *are independent.*

An explicit construction of a Brownian motion is in the following proposition.

Proposition 5.2.5 (Construction and properties of Brownian motion)**.** *Given the probability space* $(\mathbb{R}^{[0,1]},\mathscr{F},\mathbb{P}^W)$*, the family of functions* $B_t:\mathbb{R}^{[0,1]}\to\mathbb{R}$ *defined by*

$$B_t(\omega)=\omega(t),\qquad t\in[0,1]$$

is a real valued standard Brownian motion on $[0,1]$*.*

Proof. The proof relies on the equalities

$$B_t(\omega)=\omega(t)=P_{\{t\}}(\omega).$$

First of all we notice that for any $t\in[0,1]$

$$B_t^{-1}(\mathscr{B}(\mathbb{R}))=\mathscr{C}_{\{t\}}.$$

Then B_t is $\mathscr{F}$-measurable and $(B_t)_{t\in[0,1]}$ is a stochastic process.
By the definition of the Wiener measure, we have

$$\mathbb{P}^W(B_0\in A)=\mu_{\{0\}}(P_{\{0\}}^{-1}(A))=\delta_0(A),\qquad \forall A\in\mathscr{B}(\mathbb{R}),$$

and then $B_0=0$, $\mathbb{P}^W$–almost surely. Let us now compute $\mathbb{P}^W\circ B_{t-s}^{-1}$, for $t>s$. For every Borel set $A\subset\mathbb{R}$,

$$\begin{aligned}\mathbb{P}^W(B_{t-s}\in A)&=\mathbb{P}^W(P_{\{t-s\}}^{-1}(A))=\mu_{\{t-s\}}(P_{\{t-s\}}^{-1}(A))\\&=\frac{1}{\sqrt{2\pi(t-s)}}\int_A e^{-\frac{x^2}{2(t-s)}}\,dx=\mathscr{N}(0,t-s)(A).\end{aligned}$$

On the other hand, if we define $h:\mathbb{R}^2\to\mathbb{R}$, $h(x,y):=y-x$, then

$$\{\omega\in\mathbb{R}^{[0,1]}:B_t(\omega)-B_s(\omega)\in A\}=\{\omega\in\mathbb{R}^{[0,1]}:\omega(t)-\omega(s)\in A\}$$

$$=\{\omega\in\mathbb{R}^{[0,1]}:h(P_{\{s,t\}}(\omega))\in A\}=P_{\{s,t\}}^{-1}(h^{-1}(A)).$$

Hence

$$\begin{aligned}\mathbb{P}^W(\{B_t - B_s \in A\}) =& \mu_{\{s,t\}}(P^{-1}_{\{s,t\}}(h^{-1}(A))) = \frac{1}{2\pi\sqrt{s(t-s)}} \int_{h^{-1}(A)} e^{-\frac{x^2}{2s} + \frac{(y-x)^2}{2(t-s)}} dxdy \\ =& \frac{1}{\sqrt{2\pi s}} \int_{\mathbb{R}} \left(\frac{1}{\sqrt{2\pi(t-s)}} \int_{(h^{-1}(A))_x} e^{-\frac{(y-x)^2}{2(t-s)}} dy \right) e^{-\frac{x^2}{2s}}\ dx\end{aligned}$$

where

$$(h^{-1}(A))_x = \{y \in \mathbb{R} : (x,y) \in h^{-1}(A)\} = \{y \in \mathbb{R} : y - x \in A\} = A + x.$$

As a consequence,

$$\begin{aligned}\frac{1}{\sqrt{2\pi(t-s)}} \int_{(h^{-1}(A))_x} e^{-\frac{(y-x)^2}{2(t-s)}} dy &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{A+x} e^{-\frac{(y-x)^2}{2(t-s)}} dy \\ &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{A} e^{-\frac{z^2}{2(t-s)}} dz \\ &= \mathscr{N}(0, t-s)(A),\end{aligned}$$

and therefore $\mathbb{P}^W(\{B_t - B_s \in A\}) = \mathscr{N}(0, t-s)(A)$.

In order to verify independence, we fix $0 < s < t$ and $A_1, A_2 \in \mathscr{B}(\mathbb{R})$. Then

$$\{B_s \in A_1\} = P^{-1}_{\{s\}}(A_1) = P^{-1}_{\{s,t\}}(A_1 \times \mathbb{R}),$$

and

$$\{B_t - B_s \in A_2\} = P^{-1}_{\{s,t\}}(h^{-1}(A_2)),$$

so we have

$$\begin{aligned}\mathbb{P}^W(\{B_s \in A_1\} \cap \{B_t - B_s \in A_2\}) &= \mathbb{P}^W(P^{-1}_{\{s,t\}}((A_1 \times \mathbb{R}) \cap h^{-1}(A_2)) \\ &= \mu_{\{s,t\}}(P^{-1}_{\{s,t\}}((A_1 \times \mathbb{R}) \cap h^{-1}(A_2)) \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{s(t-s)}} \int_{(A_1 \times \mathbb{R}) \cap h^{-1}(A_2)} e^{-\frac{x^2}{2s} - \frac{(y-x)^2}{2(t-s)}} dxdy \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{s(t-s)}} \int_{\mathbb{R}} e^{-\frac{x^2}{2s}} \left(\int_{((A_1 \times \mathbb{R}) \cap h^{-1}(A_2))_x} e^{-\frac{(y-x)^2}{2(t-s)}} dy \right) dx \\ &= \frac{1}{2\pi} \frac{1}{\sqrt{s(t-s)}} \int_{A_1} e^{-\frac{x^2}{2s}} \left(\int_{A_2 + x} e^{-\frac{(y-x)^2}{2(t-s)}} dy \right) dx \\ &= \frac{1}{\sqrt{2\pi s}} \int_{A_1} e^{-\frac{x^2}{2s}} dx \frac{1}{\sqrt{2\pi(t-s)}} \int_{A_2} e^{-\frac{z^2}{2(t-s)}} dz \\ &= \mathbb{P}^W(\{B_s \in A_1\}) \cdot \mathbb{P}^W(\{B_t - B_s \in A_2\}).\end{aligned}$$

$\square$

Now, we have a measure on $(\mathbb{R}^{[0,1]}, \mathscr{F})$, but we are looking for a measure on a separable Banach space. We now show how to define the measure $\mathbb{P}^W$ on $C([0,1])$; this is not immediate because $C([0,1])$ does not belong to $\mathscr{F}$. To avoid this problem, the main point is to prove that the Brownian motion (B_t) can be modified in a convenient way to obtain a process with continuous trajectories. We prove something more, namely that the trajectories are Hölder continuous for $\mathbb{P}^W$–a.e. $\omega \in \mathbb{R}^{[0,1]}$.

We need the following useful lemma. We recall that the limsup of a sequence of sets (A_n) is defined by

$$\limsup_{n\to\infty} A_n := \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k$$

and it is the set of points ω such that $\omega \in A_n$ for infinitely many $n \in \mathbb{N}$.

Lemma 5.2.6 (Borel-Cantelli)**.** *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $(A_n)_{n\in\mathbb{N}} \subset \mathscr{F}$ be a sequence of measurable sets. If*

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty,$$

then

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 0.$$

Proof. We define the sets

$$B_n := \bigcup_{k\geq n} A_k.$$

Then $B_{n+1} \subset B_n$ for every n, and setting

$$B := \bigcap_{n\in\mathbb{N}} B_n = \limsup_{n\to\infty} A_n,$$

by the continuity property of measures along monotone sequences (see Remark 1.1.3)

$$\mathbb{P}(B) = \lim_{n\to\infty} \mathbb{P}(B_n).$$

On the other hand,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(B_n) &= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} A_k\right) \\ &\leq \lim_{n\to\infty} \sum_{k\geq n} \mathbb{P}(A_k) = 0.\end{aligned}$$

□

Now we state and prove the Kolmogorov continuity theorem; we need the notion of version of a stochastic process. Given two stochastic processes $X_t, \widetilde{X}_t$, $t \in [0,1]$ on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, we say that $\widetilde{X}_t$ is a *version* of X_t if

$$\mathbb{P}(\{X_t \neq \widetilde{X}_t\}) = 0, \qquad \forall t \in [0,1].$$

We use also Chebychev's inequality, whose proof is left as Exercise 5.3. For any $\beta > 0$, for any measurable function f such that $|f|^\beta \in L^1(\Omega, \mu)$ we have

$$\mu(\{|f| \geq \lambda\}) \leq \frac{1}{\lambda^\beta} \int_\Omega |f|^\beta d\mu \qquad \forall \lambda > 0.$$

Moreover, we use the classsical Egoroff's theorem, see e.g. [9, Theorem 7.5.1].

Theorem 5.2.7. *Let μ be a positive finite measure on the measurable space $(X, \mathscr{F})$, and for $n \in \mathbb{N}$ let $f, f_n : X \to \mathbb{R}$ be measurable functions such that $f_n(x) \to f(x)$ for μ-a.e. $x \in X$. Then, for very $\varepsilon > 0$ there is a measurable set A with $\mu(X \setminus A) < \varepsilon$ such that $f_n \to f$ uniformly on A.*

Theorem 5.2.8 (Kolmogorov continuity Theorem)**.** *Let $(X_t)_{t\in[0,1]}$ be a stochastic process on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and assume that there exist $\alpha, \beta > 0$, such that*

$$\mathbb{E}[|X_t - X_s|^\beta] \leq C|t-s|^{1+\alpha}, \quad t,\ s \in [0,1].$$

Then there exist a set $A \in \mathscr{F}$ with $\mathbb{P}(A) = 1$ and a version $(\widetilde{X}_t)_{t\in[0,1]}$ such that the map $t \mapsto \widetilde{X}_t(\omega)$ is γ–Hölder continuous for any $\gamma < \frac{\alpha}{\beta}$ and for any $\omega \in A$.

Proof. Let us define

$$\mathscr{D}_n = \left\{ \frac{k}{2^n} : k = 0, \dots, 2^n \right\}, \qquad \mathscr{D} = \bigcup_{n\in\mathbb{N}} \mathscr{D}_n.$$

We compute the measures of the sets

$$A_n = \left\{ \max_{1\leq k\leq 2^n} \left| X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}} \right| \geq \frac{1}{2^{\gamma n}} \right\};$$

using Chebychev's inequality. We have

$$\begin{aligned}
\mathbb{P}(A_n) &= \mathbb{P}\left(\bigcup_{k=1}^{2^n} \left\{ \left| X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}} \right| \geq \frac{1}{2^{\gamma n}} \right\} \right) \\
&\leq \sum_{k=1}^{2^n} \mathbb{P}\left(\left\{ \left| X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}} \right| \geq \frac{1}{2^{\gamma n}} \right\} \right) \\
&\leq \sum_{k=1}^{2^n} 2^{\gamma n\beta} \mathbb{E}\left[\left| X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}} \right|^\beta \right] \\
&\leq C \sum_{k=1}^{2^n} 2^{\gamma n\beta} \left| \frac{k}{2^n} - \frac{k-1}{2^n} \right|^{1+\alpha} \\
&= C 2^{-n(\alpha - \gamma\beta)}.
\end{aligned}$$

As a consequence we obtain that the series

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) \le C \sum_{n\in\mathbb{N}} 2^{-n(\alpha-\gamma\beta)}$$

is convergent if $\gamma < \frac{\alpha}{\beta}$. In this case, by the Borel-Cantelli Lemma 5.2.6, the set

$$A = \Omega \setminus \limsup_{n\to\infty} A_n$$

has full measure, $\mathbb{P}(A) = 1$. By construction, for any $\omega \in A$ there exists $N(\omega)$ such that

$$\max_{1\le k\le 2^n} \left| X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega) \right| \le 2^{-\gamma n}, \qquad \forall n \ge N(\omega).$$

We claim that for every $\omega \in A$ the restriction of the function $t \to X_t(\omega)$ to $\mathscr{D}$ is γ–Hölder continuous, i.e.,

$$\exists\, C > 0 \quad \text{such that} \quad |X_t(\omega) - X_s(\omega)| \le C|t-s|^\gamma \tag{5.2.1}$$

for all $t, s \in \mathscr{D}$. Indeed, it is enough prove that (5.2.1) holds for $t, s \in \mathscr{D}$ with $|t-s| \le 2^{-N(\omega)}$.

Fixed $t, s \in \mathscr{D}$ such that $|t-s| \le 2^{-N(\omega)}$, there exists a unique $n \ge N(\omega)$ such that $2^{-n-1} < |t-s| \le 2^{-n}$. We consider the sequences $s_k \le s$, $t_k \le t$, $s_k, t_k \in \mathscr{D}$, defined by $s_0 = t_0 = 0$, and for $k \ge 1$

$$s_k = \frac{[2^k s]}{2^k}, \quad t_k = \frac{[2^k t]}{2^k},$$

where $[x]$ is the integer part of x. Such sequences are monotone increasing, and since $t, s \in \mathscr{D}$, they are eventually constant. Moreover,

$$s_{k+1} - s_k \le \frac{1}{2^{k+1}}, \;\; t_{k+1} - t_k \le \frac{1}{2^{k+1}}, \quad k \in \mathbb{N}.$$

Then,

$$X_t(\omega) - X_s(\omega) = X_{t_n}(\omega) - X_{s_n}(\omega) + \sum_{k\ge n}(X_{t_{k+1}}(\omega) - X_{t_k}(\omega)) - \sum_{k\ge n}(X_{s_{k+1}}(\omega) - X_{s_k}(\omega))$$

where the series are indeed finite sums. Hence

$$|X_t(\omega) - X_s(\omega)| \le 2^{-\gamma n} + 2\sum_{k\ge n} 2^{-\gamma(k+1)} = \frac{2^{-\gamma n}}{1-2^{-\gamma}} \le \frac{2^{-\gamma}}{1-2^{-\gamma}}|t-s|^\gamma.$$

So (5.2.1) holds with $C = \frac{2^{-\gamma}}{1-2^{-\gamma}}$, for $t, s \in \mathscr{D}$ with $|t-s| \le 2^{-N(\omega)}$. Covering $[0,1]$ by a finite number of intervals with length $2^{-N(\omega)}$, we obtain that (5.2.1) holds for every $t, s \in \mathscr{D}$ (possibly, with a larger constant C). In particular, the mapping $t \mapsto X_t(\omega)$ is uniformly continuous on the dense set $\mathscr{D}$; therefore it admits a unique continuous extension to the whole $[0,1]$ which is what we need to define $\widetilde{X}_t(\omega)$.

Let us define for $\omega \in A$

$$\widetilde{X}_t(\omega) = \lim_{\mathscr{D} \ni s \to t} X_s(\omega),$$

and for $\omega \notin A$

$$\widetilde{X}_t(\omega) = 0.$$

It is clear that $\mathbb{P}(\{X_t \neq \widetilde{X}_t\}) = 0$ if $t \in \mathscr{D}$. For an arbitrary $t \in [0,1]$, there exists a sequence (t_h) in $\mathscr{D}$ such that X_{t_h} converges to $\widetilde{X}_t$ $\mathbb{P}$–a.e. We use Egoroff's Theorem 5.2.7, and for any $\varepsilon > 0$ there exists $E_\varepsilon \in \mathscr{F}$ such that $\mathbb{P}(E_\varepsilon) < \varepsilon$ and X_{t_h} converges uniformly to $\widetilde{X}_t$ on $\Omega \setminus E_\varepsilon$. This implies convergence in measure, i.e. for any $\lambda > 0$

$$\lim_{h \to \infty} \mathbb{P}(\{|X_{t_h} - \widetilde{X}_t| > \lambda\}) = 0.$$

On the other hand, we know that

$$\mathbb{P}(\{|X_{t_h} - X_t| > \lambda\}) \leq \frac{1}{\lambda^\beta} \mathbb{E}[|X_{t_h} - X_t|^\beta] \leq \frac{C|t_h - t|^{1+\alpha}}{\lambda^\beta}$$

and then

$$\lim_{h \to \infty} \mathbb{P}(\{|X_{t_h} - X_t| > \lambda\}) = 0.$$

We deduce that $\widetilde{X}_t = X_t$ $\mathbb{P}$–a.e., see Exercise 5.6. Hence $(\widetilde{X}_t)$ is a version of (X_t). □

Our aim now is to define a Borel measure (the Wiener measure) on the Banach space $C([0,1])$ endowed as usual with the sup norm. To do this we use the Brownian motion $(B_t)_{t \in [0,1]}$ on $(\mathbb{R}^{[0,1]}, \mathscr{F})$.

Lemma 5.2.9. *Let $\mathbb{P}^W$ be the Wiener measure on $(\mathbb{R}^{[0,1]}, \mathscr{F})$ and let $(B_t)_{t \in [0,1]}$ be the Brownian motion defined in Proposition 5.2.5. Then, for any $k \in \mathbb{N}$*

$$\mathbb{E}[(B_t - B_s)^k] = \begin{cases} 0 & \textit{if } k \textit{ is odd} \\ \\ \frac{k!}{(\frac{k}{2})! 2^{\frac{k}{2}}} |t-s|^{\frac{k}{2}} & \textit{if } k \textit{ is even.} \end{cases}$$

Proof. Let us take $0 \leq s \leq t \leq 1$. Since the law of $B_t - B_s$ is $\mathscr{N}(0, t-s)$, we get

$$I_k := \mathbb{E}[(B_t - B_s)^k] = \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} x^k e^{-\frac{x^2}{2(t-s)}} dx.$$

As a consequence $I_k = 0$ if k is odd, whereas integrating by parts one obtains $I_0 = 1$, $I_2 = (t-s)$ and $I_{2h} = (2h-1)(t-s) I_{2h-2}$ for $h \geq 2$. □

Lemma 5.2.9 and Theorem 5.2.8 yield that, fixed any $\gamma < 1/2$, there exists a version $(\widetilde{B}_t)$ of (B_t) such that the trajectories $t \mapsto \widetilde{B}_t(\omega)$ are γ–Hölder continuous, in particular they are continuous. For any t, the random variables B_t and $\widetilde{B}_t$ have the same law, $\mathbb{P}^W \circ B_t^{-1} = \mathbb{P}^W \circ \widetilde{B}_t^{-1}$. The map $P : \mathbb{R}^{[0,1]} \to \mathbb{R}^n$,

$$P(\omega) = (\widetilde{B}_{t_1}(\omega), \ldots, \widetilde{B}_{t_n}(\omega)) \tag{5.2.2}$$

is measurable for any choice of $t_1, \dots, t_n \in [0,1]$, and the image measure of $\mathbb{P}^W$ under the map P is the same as the image measure of $\mathbb{P}^W$ under the map

$$\omega \mapsto (B_{t_1}(\omega), \dots, B_{t_n}(\omega)).$$

Finally, for $C \in \mathscr{C}_F$, $\mathbb{P}^W \circ P^{-1}(E) = \mu_F(E)$. We leave the verification of these properties as an exercise, see Exercise 5.5.

We recall some facts. The first one is the characterisation of the dual space $(C([0,1]))^*$. We denote by $\mathscr{M}([0,1])$ the space of all real finite measures on $[0,1]$; it is a real Banach space with the norm $\|\mu\| = |\mu|([0,1])$, see Exercise 5.4.

Theorem 5.2.10 (Riesz representation Theorem). *There is a linear isometry between the space $\mathscr{M}([0,1])$ of finite measures and $(C([0,1]))^*$, i.e. $L \in (C([0,1]))^*$ iff there exists $\mu \in \mathscr{M}([0,1])$ such that*

$$L(f) = \int_{[0.1]} f(t)\mu(dt), \qquad \forall f \in C([0,1]).$$

In addition $\|L\| = |\mu|([0,1])$.

We refer to [9, Theorem 7.4.1] for a proof.

On $C([0,1])$ we define the σ–algebra $\mathscr{C}'_F$ for $F = \{t_1, \dots, t_n\}$ as the family of sets

$$C = \{\omega \in C([0,1]) : (\omega(t_1), \dots, \omega(t_n)) \in B\},$$

where $B \in \mathscr{B}(\mathbb{R}^n)$. We also define the algebra

$$\mathscr{A}' = \bigcup_{F \subset [0,1],\ F \text{ finite}} \mathscr{C}'_F$$

and we denote by $\mathscr{F}'$ the σ–algebra generated by $\mathscr{A}'$. Using Theorem 5.2.10 and the fact that any Dirac measure δ_t is in $(C([0,1]))^*$, it is clear that $\mathscr{F}' \subset \mathscr{B}(C([0,1]))$; indeed, if $F = \{t_1, \dots, t_n\}$ and $B \in \mathscr{B}(\mathbb{R}^n)$, we have

$$\begin{aligned} C :=& \{\omega \in C([0,1]) : (\omega(t_1), \dots, \omega(t_n)) \in B\} \\ =& \{\omega \in C([0,1]) : (\delta_{t_1}(\omega), \dots, \delta_{t_n}(\omega)) \in B\} \in \mathscr{E}(C([0,1]), \{\delta_{t_1}, \dots, \delta_{t_n}\}). \end{aligned}$$

We have also the reverse inclusion, i.e. $\mathscr{B}(C([0,1])) \subset \mathscr{F}'$; the proof is similar to the proof of Theorem 2.1.1. Indeed, fix $\omega_0 \in C([0,1])$, $r > 0$ and let $\mathscr{D}$ be the set in the proof of the Kolmogorov continuity Theorem 5.2.8. Note that $\omega \in \overline{B}(\omega_0, r)$ if and only if $\|\omega - \omega_0\|_\infty \le r$, and by continuity this is equivalent to $|\omega(t) - \omega_0(t)| \le r$ for any $t \in \mathscr{D}$. Then

$$\overline{B}(\omega_0, r) = \bigcap_{n \in \mathbb{N}} \left\{\omega \in C([0,1]) : \omega\Big(\frac{k}{2^n}\Big) \in \Big[r - \omega_0\Big(\frac{k}{2^n}\Big), r + \omega_0\Big(\frac{k}{2^n}\Big)\Big],\ \forall k = 0, \dots, 2^n\right\}.$$

The set in the right hand side belongs to $\mathscr{F}'$. Since $C([0,1])$ is separable, as in the proof of Theorem 2.1.1 we deduce that $\mathscr{B}(C([0,1])) \subset \mathscr{F}'$.

Now we use the Kolmogorov continuity Theorem 5.2.8: there exists a set $A \in \mathscr{F}$ with $\mathbb{P}(A) = 1$ and a version $(\widetilde{B}_t)_{t\in[0,1]}$ such that the map $t \mapsto \widetilde{B}_t(\omega)$ is continuous for any $\omega \in A$. We define the restricted σ–algebra

$$\mathscr{F}_A = \{E \cap A : E \in \mathscr{F}\}$$

and the restriction $\mathbb{P}^W_A$ of $\mathbb{P}^W$ to $\mathscr{F}_A$.

Proposition 5.2.11. *The map* $\widetilde{B} : (A, \mathscr{F}_A) \to (C([0,1]), \mathscr{B}(C([0,1])))$ *defined by*

$$\widetilde{B}(\omega)(t) := \widetilde{B}_t(\omega)$$

is measurable. The image measure $\mathbb{P}^W_A \circ \widetilde{B}^{-1}$*, called the* Wiener measure on $C([0,1])$*, has the property that for any* $C' \in \mathscr{C}'_F$*,* $C' = C \cap C([0,1])$ *with* $C \in \mathscr{C}_F$*, and*

$$\mathbb{P}^W_A \circ \widetilde{B}^{-1}(C') = \mu_F(C). \tag{5.2.3}$$

Proof. We know that $\mathscr{B}(C([0,1])) = \mathscr{F}'$, so it is sufficient to prove that for every finite set $F = \{t_1, \dots, t_n\} \subset [0,1]$, we have

$$\widetilde{B}^{-1}(C') \in \mathscr{F}_A, \quad \forall C' \in \mathscr{C}'_F.$$

Let $E \in \mathscr{B}(\mathbb{R}^n)$ and

$$C' = \{\omega \in C([0,1]) : (\omega(t_1), \dots, \omega(t_n)) \in E\} = C \cap C([0,1]),$$

with $C \in \mathscr{C}_F$ given by

$$C = \{\omega \in \mathbb{R}^{[0,1]} : (\omega(t_1), \dots, \omega(t_n)) \in E\}$$

Since for any $t \in [0,1]$ we have $\mathbb{P}^W(\{B_t \neq \widetilde{B}_t\}) = 0$, we can find $A_{t_1}, \dots, A_{t_n} \in \mathscr{F}$ such that $\mathbb{P}^W(A_{t_i}) = 1$ for any $i = 1, \dots, n$ and $\widetilde{B}_{t_i} = B_{t_i}$ in A_{t_i}. Then

$$\begin{aligned}
\widetilde{B}^{-1}(C) &= \{\omega \in A : (\widetilde{B}_{t_1}(\omega), \dots, \widetilde{B}_{t_n}(\omega)) \in E\} \\
&= A \cap \bigcap_{i=1}^n A_{t_i} \cap \{\omega \in \mathbb{R}^{[0,1]} : (\omega(t_1), \dots, \omega(t_n)) \in E\} \\
&= A \cap \bigcap_{i=1}^n A_{t_i} \cap C.
\end{aligned}$$

Since $\bigcap_{i=1}^n A_{t_i} \cap C \in \mathscr{F}$, we deduce $\widetilde{B}^{-1}(C') \in \mathscr{F}_A$ and then $\widetilde{B}$ is measurable. The last assertion follows from the fact that

$$(\mathbb{P}^W_A \circ \widetilde{B}^{-1})(C') = \mathbb{P}^W\Big(A \cap \bigcap_{i=1}^n A_{t_i} \cap C\Big) = \mathbb{P}^W(C) = \mu_F(C).$$

□

5.3 Exercises

Exercise 5.1. Use Corollary 5.1.4 to prove that two random variables X and Y on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ are independent if and only if $\mathbb{P} \circ (X,Y)^{-1} = (\mathbb{P} \circ X^{-1}) \otimes (\mathbb{P} \circ Y^{-1})$.

Exercise 5.2. Prove Proposition 5.2.1.

Exercise 5.3. Prove *Chebychev's inequality*: if $\beta > 0$ and f is a measurable function such that $|f|^\beta \in L^1(\Omega, \mu)$, we have for any $\lambda > 0$

$$\mu(\{|f| \geq \lambda\}) \leq \frac{1}{\lambda^\beta} \int_\Omega |f|^\beta d\mu.$$

Exercise 5.4. Prove that the vector space $\mathscr{M}([0,1])$ with norm given by the total variation is a Banach space.

Exercise 5.5. Prove that the map $P : \mathbb{R}^{[0,1]} \to \mathbb{R}^n$, defined by (5.2.2), is measurable. Prove that $\mathbb{P}^W \circ P^{-1} = \mathbb{P}^W \circ T^{-1}$ with

$$T(\omega) = (\omega(t_1), \ldots, \omega(t_n)) = (B_{t_1}(\omega), \ldots, B_{t_n}(\omega)).$$

Prove in addition that if $C \in \mathscr{C}_F$, then

$$(\mathbb{P}^W \circ P^{-1})(C) = \mu_F(C).$$

Exercise 5.6. Let (X_n) be a sequence of random variables on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ that converge in measure to X and to Y. Prove that $X = Y$, $\mathbb{P}$–a.e.

Lecture 6

The classical Wiener space

In this Lecture we present the *classical Wiener space*, which is the archetype of the structure we are describing. Indeed, any triple (X, γ, H) (where X is a separable Banach space, γ is a Gaussian measure and H is the Cameron-Martin space) is called an *abstract Wiener space*. In the classical Wiener space the Banach space is that of continuous paths, $X = C([0,1])$, and all the objects involved can be described explicitly. The Gaussian measure is the Wiener measure γ^W defined in Lecture 5, the covariance operator is the integral operator with kernel $\min\{x, y\}$ on $[0,1]^2$ and both the Cameron-Martin space H and X^*_γ are spaces of functions defined on $[0,1]$.

6.1 The classical Wiener space

We start by considering the measure space $(X, \mathscr{B}(X), \gamma^W)$ where $X = C([0,1])$, $\mathscr{B}(X)$ is the Borel σ–algebra on X and $\gamma^W = \mathbb{P}^W \circ \widetilde{B}^{-1}$ is the measure defined in Proposition 5.2.11.

We give the following approximation result for measures in terms of Dirac measures. For every real measure $\mu \in \mathscr{M}([0,1])$ and $n \in \mathbb{N}$ we set

$$\mu_n = \mu(\{1\})\delta_1 + \sum_{i=0}^{2^n-1} \mu\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right)\delta_{\frac{i+1}{2^n}}. \tag{6.1.1}$$

Lemma 6.1.1. *The following statements hold:*

(i) if $\mu, \nu \in \mathscr{M}([0,1])$ are two finite measures and (μ_n), (ν_n) are two sequences of measure weakly convergent to μ and ν respectively, then $\mu_n \otimes \nu_n$ weakly converges to $\mu \otimes \nu$;

(ii) for every $\mu \in \mathscr{M}([0,1])$ the sequence (μ_n) defined in (6.1.1) *converges weakly to μ;*

Proof. (i) The statement is trivial for functions of the type $\varphi(x, y) = \varphi_1(x)\varphi_2(y)$ with $\varphi_1, \varphi_2 \in C([0,1])$. Since the linear span of such functions is dense in $C([0,1]^2)$ by the

Stone–Weierstrass Theorem, the conclusion follows.
(ii) Let us fix $f \in C([0,1])$. Then

$$\int_{[0,1]} f(x)\mu_n(dx) - \int_{[0,1]} f(x)\mu(dx) = \sum_{i=0}^{2^n-1} \int_{\left[\frac{i}{2^n},\frac{i+1}{2^n}\right)} \Big(f\Big(\frac{i+1}{2^n}\Big) - f(x) \Big) \mu(dx).$$

By the uniform continuity of f, for every $\varepsilon > 0$ there is $n_0 > 0$ such that for $n > n_0$ we have $|f(\frac{i+1}{2^n}) - f(x)| < \varepsilon$ for every $x \in [i2^{-n}, (i+1)2^{-n}]$ and for every $i = 0, \ldots, 2^n - 1$, whence for $n > n_0$

$$\Big|\sum_{i=0}^{2^n-1} \int_{\left[\frac{i}{2^n},\frac{i+1}{2^n}\right)} \Big(f\Big(\frac{i+1}{2^n}\Big) - f(x) \Big) \mu(dx)\Big| < \varepsilon|\mu|([0,1])$$

and $\mu_n(f) \to \mu(f)$. □

Proposition 6.1.2. *The characteristic function of the Wiener measure γ^W is*

$$\widehat{\gamma^W}(\mu) = \exp\Big\{-\frac{1}{2}\int_{[0,1]^2} \min\{x,y\}(\mu\otimes\mu)(d(x,y))\Big\}, \quad \mu \in \mathscr{M}([0,1]).$$

So, γ^W is a Gaussian measure with mean zero and covariance operator

$$B_{\gamma^W}(\mu,\nu) = \int_{[0,1]^2} \min\{t,s\}(\mu\otimes\nu)(d(t,s)), \qquad \mu,\nu \in \mathscr{M}([0,1]). \tag{6.1.2}$$

Proof. We start by considering a linear combination of two Dirac measures

$$\mu = \alpha\delta_s + \beta\delta_t$$

with $\alpha, \beta \in \mathbb{R}, s < t \in [0,1]$. Then

$$\widehat{\gamma^W}(\mu) = \int_X \exp\{i\alpha\delta_s(f) + i\beta\delta_t(f)\}\gamma^W(df) = \int_X \exp\{i\alpha f(s) + i\beta f(t)\}\gamma^W(df).$$

Since $\gamma^W = \mathbb{P}^W_A \circ \widetilde{B}^{-1}$ with $\mathbb{P}^W(A) = 1$ and $\widetilde{B}$ is a version of the Brownian motion which is continuous on A, noticing that $(\widetilde{B}_t)_{t\in[0,1]}$ and $(B_t)_{t\in[0,1]}$ have the same image measure, we obtain

$$\begin{aligned}
\widehat{\gamma^W}(\mu) &= \int_A \exp\{i\alpha\widetilde{B}(\omega)(s) + i\beta\widetilde{B}(\omega)(t)\}\mathbb{P}^W(d\omega) = \int_A \exp\{i\alpha\widetilde{B}_s(\omega) + i\beta\widetilde{B}_t(\omega)\}\mathbb{P}^W(d\omega) \\
&= \int_{\mathbb{R}^{[0,1]}} \exp\{i\alpha\widetilde{B}_s(\omega) + i\beta\widetilde{B}_t(\omega)\}\mathbb{P}^W(d\omega) = +\int_{\mathbb{R}^{[0,1]}} \exp\{i\alpha B_s(\omega) + i\beta B_t(\omega)\}\mathbb{P}^W(d\omega) \\
&= \int_{\mathbb{R}^{[0,1]}} \exp\{i(\alpha+\beta)B_s(\omega) + i\beta(B_t(\omega) - B_s(\omega)\}\mathbb{P}^W(d\omega).
\end{aligned}$$

Since B_s and $B_t - B_s$ are independent, we may write

$$\begin{aligned}
\widehat{\gamma^W}(\mu) &= \int_{\mathbb{R}^{[0,1]}} \exp\{i(\alpha+\beta)B_s(\omega) + i\beta(B_t(\omega) - B_s(\omega)\}\mathbb{P}^W(d\omega) \\
&= \int_{\mathbb{R}^{[0,1]}} \exp\{i(\alpha+\beta)B_s(\omega)\}\mathbb{P}^W(d\omega) \cdot \int_{\mathbb{R}^{[0,1]}} \exp\{i\beta(B_t(\omega) - B_s(\omega)\}\mathbb{P}^W(d\omega) \\
&= \int_{\mathbb{R}} \exp\{i(\alpha+\beta)x\}\mathscr{N}(0,s)(dx) \cdot \int_{\mathbb{R}} \exp\{i\beta y\}\mathscr{N}(0,t-s)(dy) \\
&= \exp\left\{-\frac{1}{2}(\alpha+\beta)^2 s\right\} \exp\left\{-\frac{1}{2}\beta^2(t-s)\right\} = \exp\left\{-\frac{1}{2}\Big((\alpha^2+2\alpha\beta)s + \beta^2 t\Big)\right\}.
\end{aligned}$$

We now compute the integral

$$\int_{[0,1]^2} \min\{x,y\}(\mu\otimes\mu)(d(x,y)) = \int_{[0,1]} \varphi(x)\mu(dx) = \alpha\varphi(s) + \beta\varphi(t),$$

where $\varphi(x) = \int_{[0,1]} \min\{x,y\}\mu(dy)$. We have

$$\begin{aligned}
\varphi(s) &= \int_{[0,1]} \min\{s,y\}\mu(dy) = \int_{[0,s]} y\mu(dy) + s\int_{(s,1]} \mu(dy) = \alpha s + \beta s \\
\varphi(t) &= \int_{[0,1]} \min\{t,y\}\mu(dy) = \int_{[0,t]} y\mu(dy) + t\int_{(t,1]} \mu(dy) = \alpha s + \beta t,
\end{aligned}$$

whence

$$\int_{[0,1]^2} \min\{x,y\}(\mu\otimes\mu)(d(x,y)) = (\alpha^2 + 2\alpha\beta)s + \beta^2 t.$$

So, the assertion of the theorem holds if μ is a linear combination of two Dirac measures.

Let us show by induction that the same assertion holds true if μ is a linear combination of a finite number of Dirac measures.

For $0 \le t_1 < t_2 < \ldots < t_n \le 1$ we define the matrix

$$Q_{t_1,\ldots,t_n} = \begin{pmatrix} t_1 & t_1 & \ldots & t_1 & t_1 \\ t_1 & t_2 & \ldots & t_2 & t_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_1 & t_2 & \ldots & t_{n-1} & t_{n-1} \\ t_1 & t_2 & \ldots & t_{n-1} & t_n \end{pmatrix}.$$

For $\alpha \in \mathbb{R}^n$, $\alpha = (\alpha_1, \ldots, \alpha_n)$

$$\mu := \sum_{j=1}^n \alpha_j \delta_{t_j},$$

one has

$$\int_{[0,1]^2} \min\{x,y\}(\mu\otimes\mu)(d(x,y)) = Q_{t_1,\ldots,t_n}\alpha\cdot\alpha.$$

We are now going to show by induction that

$$\widehat{\gamma^W}(\mu) = \exp\Big\{-\frac{1}{2}Q_{t_1,\ldots,t_n}\alpha\cdot\alpha\Big\}. \tag{6.1.3}$$

This is immediate for $n = 1$. Assume that it is true for n. Let $0 \le t_1 < t_2 < \ldots < t_n < t_{n+1} \le 1$,

$$\mu := \sum_{j=1}^{n+1} \alpha_j \delta_{t_j},$$

Then, by the induction hypothesis and using the independence of $B_{t_{n+1}} - B_{t_n}$ with respect to $\{B_s; 0 \le s \le t_n\}$, we obtain

$$\begin{aligned}
\widehat{\gamma^W}(\mu) &= \int_{\mathbb{R}^{[0,1]}} \exp\Big\{i\sum_{j=1}^{n+1}\alpha_j B_{t_j}(\omega)\Big\}\mathbb{P}^W(d\omega)\\
&= \int_{\mathbb{R}^{[0,1]}} \exp\Big\{i\sum_{j=1}^{n-1}\alpha_j B_{t_j}(\omega) + (\alpha_n+\alpha_{n+1})B_{t_n}(\omega)\Big\}\mathbb{P}^W(d\omega)\cdot\\
&\quad\cdot\int_{\mathbb{R}^{[0,1]}} \exp\Big\{\alpha_{n+1}(B_{t_{n+1}}(\omega) - B_{t_n}(\omega))\Big\}\mathbb{P}^W(d\omega)\\
&= \exp\Big\{-\frac{1}{2}Q_{t_1,\ldots,t_n}\tilde{\alpha}\cdot\tilde{\alpha} - \frac{1}{2}\alpha_{n+1}^2(t_{n+1}-t_n)\Big\}\\
&= \exp\Big\{-\frac{1}{2}Q_{t_1,\ldots,t_{n+1}}\alpha\cdot\alpha,\Big\}
\end{aligned}$$

where we have set $\tilde{\alpha} = (\alpha_1, \ldots, \alpha_{n-1}, \alpha_n + \alpha_{n+1})$. In the general case we conclude by using Lemma 6.1.1. Indeed, if μ_n is the approximation of μ defined in (6.1.1), then for any $f \in X$,

$$\lim_{n\to+\infty}\mu_n(f) = \lim_{n\to+\infty}\int_{[0,1]} f(x)\mu_n(dx) = \int_{[0,1]} f(x)\mu(dx) = \mu(f).$$

Hence $\exp\{i\mu_n(f)\}$ converges to $\exp\{i\mu(f)\}$ for any $f \in X$, so by the Lebesgue Dominated Convergence Theorem

$$\begin{aligned}
\widehat{\gamma^W}(\mu) &= \int_X \exp\{i\mu(f)\}\gamma^W(df) = \lim_{n\to+\infty}\int_X \exp\{i\mu_n(f)\}\gamma^W(df)\\
&= \lim_{n\to+\infty}\exp\left\{-\frac{1}{2}\int_{[0,1]^2}\min\{x,y\}(\mu_n\otimes\mu_n)(d(x,y))\right\}\\
&= \exp\left\{-\frac{1}{2}\int_{[0,1]^2}\min\{x,y\}(\mu\otimes\mu)(d(x,y))\right\}.
\end{aligned}$$

Then we conclude applying Theorem 2.2.4. $\square$

We notice that the space

$$C_0([0,1]) := \{f \in C([0,1]) : f(0) = 0\} = \delta_0^{-1}(\{0\})$$

is a closed subspace of $C([0,1])$. Since

$$\widehat{\gamma^W}(\delta_0) = \exp\left\{-\frac{1}{2}\int_{[0,1]^2} \min\{x,y\}(\delta_0 \otimes \delta_0)(d(x,y))\right\} = 1,$$

$\gamma^W \circ \delta_0^{-1} = \mathscr{N}(0,0) = \delta_0$, and so

$$\gamma^W(C_0([0,1])) = (\gamma^W \circ \delta_0^{-1})(\{0\}) = 1.$$

Then γ^W is degenerate and it is concentrated on $C_0([0,1])$.

6.2 The Cameron–Martin space

In order to determine the Cameron–Martin space of $(C([0,1]), \gamma^W)$, we use the embedding $\iota : C([0,1]) \to L^2(0,1)$, $\iota(f) = f$, which is a continuous injection since

$$\|\iota(f)\|_{L^2(0,1)} \le \|f\|_\infty.$$

If we consider the image measure $\widetilde{\gamma}^W := \gamma^W \circ \iota^{-1}$ on $L^2(0,1)$, the Cameron–Martin spaces on $C([0,1])$ and on $L^2(0,1)$ are the same in the sense of Proposition 3.1.10. The fact that $L^2(0,1)$ is a Hilbert space allows us to use the results of Section 4.2.

By using the identification $(L^2(0,1))^* = L^2(0,1)$, the characteristic function of the Gaussian measure $\widetilde{\gamma}^W$ is

$$\widehat{\widetilde{\gamma}^W}(g) = \int_{L^2(0,1)} \exp\Big\{i\langle f, g\rangle_{L^2(0,1)}\Big\}\, \tilde{\gamma}^W(df) = \int_{C([0,1])} \exp\{i\langle \iota(f), g\rangle_{L^2(0,1)}\}\, \gamma^W(df).$$

Let us compute $\langle \iota(f), g\rangle_{L^2(0,1)}$. If we denote by $\iota^* : L^2(0,1) \to \mathscr{M}([0,1])$ the adjoint of ι, then

$$\langle \iota(f), g\rangle_{L^2(0,1)} = \iota^*(g)(f). \tag{6.2.1}$$

Since $\iota(f)(x) = f(x)$, (6.2.1) yields

$$\int_0^1 f(x)g(x)dx = \int_{[0,1]} f(x)\iota^*(g)(dx), \qquad \forall f \in C([0,1]).$$

Hence $\iota^*(g) = g\lambda_1$, where λ_1 is the Lebesgue measure on $[0,1]$. Therefore, according to Proposition 6.1.2,

$$\widehat{\widetilde{\gamma}^W}(g) = \widehat{\gamma^W}(\iota^* g) = \exp\left\{-\frac{1}{2}\int_{[0,1]^2} \min\{x,y\} g(x)g(y)\, d(x,y)\right\}$$

so that $\widetilde{\gamma}^W$ is a Gaussian measure with covariance

$$B_{\widetilde{\gamma}^W}(f,g) = \int_{[0,1]^2} \min\{x,y\} f(x)g(y)d(x,y) = \int_0^1 Qf(y)g(y)dy,$$

where

$$Qf(y) = \int_0^1 \min\{x,y\} f(x)dx$$

is the covariance operator $Q : L^2(0,1) \to L^2(0,1)$ introduced in Section 4.2.

Theorem 6.2.1. *The Cameron–Martin space H of $\widetilde{\gamma}^W$ on $(L^2(0,1), \mathscr{B}(L^2(0,1)))$ is*

$$H^1_0([0,1]) := \{f \in L^2(0,1) : f' \in L^2(0,1) \text{ and } f(0) = 0\}.$$

Proof. As the Cameron-Martin space is the range of $Q^{1/2}$, see Theorem 4.2.7, we find the eigenvalues and eigenvectors of Q, i.e., we look for all $\lambda \in \mathbb{R}$ and $f \in L^2(0,1)$ such that $Qf = \lambda f$. Equality $Qf = \lambda f$ is equivalent to

$$\lambda f(x) = \int_0^1 \min\{x,y\} f(y)dy = \int_0^x y f(y)dy + x \int_x^1 f(y)dy \tag{6.2.2}$$

for a.e. $x \in [0,1]$. If (6.2.2) holds, f is weakly differentiable and

$$\lambda f'(x) = \int_x^1 f(y)dy.$$

For $\lambda = 0$, we get immediately $f \equiv 0$. For $\lambda \neq 0$ we get that f' is weakly differentiable and

$$\lambda f''(x) = -f(x) \qquad \text{a.e.}$$

Moreover, the continuous version of f vanishes at 0 and the continuous version of f' vanishes at 1. We have proved that if f is an eigenvector of Q with eigenvalue λ, then f is the solution of the following problem on $(0,1)$:

$$\begin{cases} \lambda f'' + f = 0, \\ f(0) = 0, \\ f'(1) = 0. \end{cases} \tag{6.2.3}$$

On the other hand, if f is the solution of problem (6.2.3), integrating between x and 1

$$\lambda f'(x) = \int_x^1 f(y)\, dy,$$

whence, integrating again between 0 and x

$$\begin{aligned}
\lambda f(x) &= \int_0^x \int_t^1 f(y)\, dy\, dt = \int_0^1 \mathbb{1}_{(0,x]}(t) \int_0^1 \mathbb{1}_{(t,1]}(y) f(y)\, dy\, dt \\
&= \int_0^1 f(y) \int_0^1 \mathbb{1}_{(0,x]}(t)\mathbb{1}_{(t,1]}(y) dt\, dy = \int_0^1 f(y) \int_0^1 \mathbb{1}_{(t,1]}(x)\mathbb{1}_{(t,1]}(y) dt\, dy \\
&= \int_0^1 \min\{x,y\} f(y)\, dy = Qf(x).
\end{aligned}$$

We leave as an exercise, see Exercise 6.2, to prove that if λ is an eigenvalue, then there exists $k \in \mathbb{N}$ such that $\lambda = \lambda_k$, where

$$\lambda_k = \frac{1}{\pi^2 \left(k + \frac{1}{2}\right)^2}, \qquad k \in \mathbb{N} \tag{6.2.4}$$

and $Qe_k = \lambda_k e_k$, $\|e_k\|_{L^2(0,1)} = 1$ if and only if

$$e_k(x) = \sqrt{2} \sin\Big(\frac{x}{\sqrt{\lambda_k}}\Big) = \sqrt{2} \sin\Big(\frac{2k+1}{2}\pi x\Big). \tag{6.2.5}$$

Let us now take $f \in L^2(0,1)$ and write

$$f = \sum_{k=1}^{\infty} f_k e_k, \qquad f_k = \langle f, e_k \rangle_{L^2(0,1)} \tag{6.2.6}$$

with e_k given by (6.2.5). Applying (4.2.10), we see that $f \in H$ if and only if

$$\sum_{k=1}^{\infty} f_k^2 \lambda_k^{-1} = \pi^2 \sum_{k=1}^{\infty} f_k^2 \left(\frac{2k+1}{2}\right)^2 < \infty.$$

This condition allows us to define the function

$$g(x) = \sqrt{2} \sum_{k=1}^{\infty} \frac{f_k}{\sqrt{\lambda_k}} \cos\left(\frac{x}{\sqrt{\lambda_k}}\right) = \pi\sqrt{2} \sum_{k=1}^{\infty} f_k \frac{2k+1}{2} \cos\left(x\pi\Big(\frac{2k+1}{2}\Big)\right)$$

and to obtain that $g \in L^2(0,1)$ is the weak derivative of f. Indeed, for any $\varphi \in C_c^\infty((0,1))$,

$$\begin{aligned}\int_0^1 f(x)\varphi'(x)dx &= \sum_{k=1}^{\infty} f_k\sqrt{2} \int_0^1 \sin\left(\frac{s}{\sqrt{\lambda_k}}\right) \varphi'(x)dx \\ &= -\sum_{k=1}^{\infty} \frac{f_k\sqrt{2}}{\sqrt{\lambda_k}} \int_0^1 \cos\left(\frac{x}{\sqrt{\lambda_k}}\right) \varphi(x)dx = -\int_0^1 g(x)\varphi(x)dx.\end{aligned}$$

In conclusion $f \in H^1(0,1)$ and, by (6.2.6), its continuous version vanishes at 0, whence $f \in H_0^1([0,1])$.

Finally, from the equality

$$|f|_H = \|Q^{-1/2} f\|_X,$$

we immediately get $|f|_H = \|f'\|_{L^2(0,1)}$. □

Remark 6.2.2. We have used the notation $H_0^1([0,1])$ to characterise the Cameron–Martin space; we point out that this space is not the closure of $C_c^\infty(0,1)$ in $H^1(0,1)$.

6.3 The reproducing kernel

In this section we determine the reproducing kernel, both for $X = C([0,1])$ and for $X = L^2(0,1)$. To do this, we need to introduce an important tool coming from probability, the stochastic Itô integral.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $(W_t)_{t\in[0,1]}$ be a Brownian motion, i.e., a stochastic process on $(\Omega, \mathscr{F}, \mathbb{P})$ satisfying the conditions of Definition 5.2.4.

If f is a simple function, i.e.,

$$f(t) = \sum_{i=0}^{n-1} c_i \mathbb{1}_{[t_i,t_{i+1})}(t)$$

with $c_i \in \mathbb{R}$, $0 = t_0 < \ldots < t_n = 1$, we define the random variable on Ω

$$\int_0^1 f(t)dW_t(\omega) := \sum_{i=0}^{n-1} c_i\big(W_{t_{i+1}}(\omega) - W_{t_i}(\omega)\big). \tag{6.3.1}$$

We claim that

$$\mathbb{E}\left[\left(\int_0^1 f(t)dW_t\right)^2\right] = \int_0^1 |f(x)|^2 dx. \tag{6.3.2}$$

Indeed, we have

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int_0^1 f(t)dW_t\Big)^2\Big] &= \sum_{i,j=0}^{n-1} c_i c_j \mathbb{E}[(W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j})] \\
&= \sum_{i=0}^{n-1} c_i^2 \mathbb{E}[(W_{t_{i+1}} - W_{t_i})^2] + 2\sum_{i=1}^{n-1}\sum_{j<i} c_i c_j \mathbb{E}[(W_{t_{j+1}} - W_{t_j})(W_{t_{i+1}} - W_{t_i})] \\
&= \sum_{i=0}^{n-1} c_i^2 \mathbb{E}[|W_{t_{i+1}-t_i}|^2] + 2\sum_{i=1}^{n-1}\sum_{j<i} c_i c_j \mathbb{E}[W_{t_{j+1}} - W_{t_j}]\cdot\mathbb{E}[W_{t_{i+1}} - W_{t_i}] \\
&= \sum_{i=0}^{n-1} c_i^2 (t_{i+1} - t_i) = \int_0^1 |f(x)|^2 dx,
\end{aligned}$$

where we have used the fact that $W_{t_{i+1}} - W_{t_i}$ is independent of $W_{t_{j+1}} - W_{t_j}$ if $j < i$ and the fact that $W_{t_{i+i}} - W_{t_i}$ has the same image measure as $W_{t_{i+1}-t_i}$, given by $\mathscr{N}(0, t_{i+1} - t_i)$.

For the next Theorem, we refer to [3].

Theorem 6.3.1 (Itô Integral). *There exists a unique continuous map $I_\Omega : L^2(0,1) \to L^2(\Omega, \mathbb{P})$ such that*

$$\mathbb{E}[|I_\Omega(f)|^2] = \int_0^1 |f(x)|^2 dx, \qquad \forall f \in L^2(0,1) \tag{6.3.3}$$

and such that

$$I_\Omega(f) = \int_0^1 f(t)dW_t$$

if f is a simple function.

Proof. Let $\mathscr{S}([0,1])$ be the linear subspace of $L^2(0,1)$ consisting of the simple functions. The map $I_\Omega : \mathscr{S}([0,1]) \to L^2(\Omega, \mathbb{P})$,

$$I_\Omega(f)(\omega) := \int_0^1 f(t) dW_t(\omega)$$

is a linear operator, defined on a dense subset of $L^2(0,1)$. Since it is continuous in the $L^2(0,1)$ topology by (6.3.2), it has a unique continuous extension to $L^2(0,1)$. □

The map defined in Theorem 6.3.1 is called the *Itô integral of f* with respect to the Brownian motion; such procedure can be performed also using different stochastic processes and more general classes of functions f. Identity (6.3.3) is called *Itô isometry* and the Itô integral is denoted by

$$I_\Omega(f) = \int_0^1 f(t) dW_t, \qquad f \in L^2(0,1).$$

In order to apply Theorem 6.3.1, we need to define Brownian motions on the probability spaces $(C([0,1]), \mathscr{B}(C([0,1])), \gamma^W)$ and $(L^2(0,1), \mathscr{B}(L^2(0,1)), \widetilde{\gamma}^W)$. We leave as an exercise (see Exercise 6.3) the verification that the evaluation map $W_t : C([0,1]) \to \mathbb{R}$ defined by

$$W_t(f) := f(t), \tag{6.3.4}$$

is indeed a standard Brownian motion on $C([0,1])$.

In the next Lemma we define the Brownian motion on $L^2(0,1)$ using the embedding $\iota : C([0,1]) \to L^2(0,1)$.

Lemma 6.3.2. *The space $C([0,1])$ belongs to $\mathscr{B}(L^2(0,1))$. As a consequence the maps $\widetilde{W_t} : L^2(0,1) \to \mathbb{R}$ defined by*

$$\widetilde{W_t}(f) = \begin{cases} W_t(\omega) & \text{if } f = \iota(\omega) \\ \\ 0 & \text{otherwise} \end{cases} \tag{6.3.5}$$

define a stochastic process on $L^2(0,1)$ and such process is a Brownian motion.

Proof. In order to see that $C([0,1])$ is a Borel set in $L^2(0,1)$, for a fixed $p \in [1,\infty]$ and for any $m, n \in \mathbb{N}$ and $r_1, r_2, s_1, s_2 \in [0,1] \cap \mathbb{Q}$ with $r_1 < r_2, s_1 < s_2$, $|r_1 - r_2| \le \frac{1}{n}, |s_1 - s_2| \le \frac{1}{n}$, $|\frac{r_1+r_2}{2} - \frac{s_1+s_2}{2}| \le \frac{1}{n}$ define the set

$$B_{p,n,m,r_1,r_2,s_1,s_2} := \Big\{ f \in L^p(0,1) : \Big| \frac{1}{r_2 - r_1} \int_{r_1}^{r_2} f(t) dt - \frac{1}{s_2 - s_1} \int_{s_1}^{s_2} f(t) dt \Big| \le \frac{1}{m} \Big\}.$$

It is enough to notice that for any $1 \le p \le \infty$ the set $B_{p,n,m,r_1,r_2,s_1,s_2}$ is Borel in $L^2(0,1)$ and the space $C([0,1])$ is a countable union of Borel sets,

$$C([0,1]) = \bigcap_{m \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} \bigcap_{\substack{r_1, r_2, s_1, s_2 \in [0,1] \cap \mathbb{Q} \\ r_1 < r_2, s_1 < s_2 \\ |r_1 - r_2| \le \frac{1}{n}, |s_1 - s_2| \le \frac{1}{n} \\ \left| \frac{r_1+r_2}{2} - \frac{s_1+s_2}{2} \right| \le \frac{1}{n}}} B_{p,n,m,r_1,r_2,s_1,s_2}.$$

Hence, the function defined in (6.3.5) is measurable. The fact that $\tilde{W}_t$ is a Brownian motion readily follows by the fact that W_t is a Brownian motion and $\tilde{\gamma}^W = \gamma^W \circ \iota^{-1}$. $\square$

We pass now to the characterisation of the reproducing kernels $X^*_{\gamma^W}$ and $X^*_{\tilde{\gamma}^W}$.

Proposition 6.3.3. *Let us consider the Gaussian measures γ^W and $\widetilde{\gamma}^W$ on $C([0,1])$ and on $L^2(0,1)$, respectively. Then*

$$X^*_{\gamma^W} = I_{C([0,1])}(L^2(0,1))$$

and

$$X^*_{\widetilde{\gamma}^W} = I_{L^2(0,1)}(L^2(0,1)).$$

Proof. Let us consider the simple function

$$g(x) = \alpha \mathbb{1}_{[0,s)}(x) + \beta \mathbb{1}_{[0,t)} = (\alpha+\beta)\mathbb{1}_{[0,s)}(x) + \beta \mathbb{1}_{[s,t)}(x),$$

$s < t \in [0,1]$. Let (W_t) be the Brownian motion defined in (6.3.4). Then by (6.3.1)

$$I_{C([0,1])}(g) = \int_0^1 g(t)dW_t = (\alpha+\beta)W_s + \beta(W_t - W_s) = \alpha W_s + \beta W_t.$$

Therefore,

$$I_{C([0,1])}(g)(f) = \alpha f(s) + \beta f(t)$$

for γ^W–a.e. $f \in C([0,1])$. On the other hand, setting

$$\mu = \alpha\delta_s + \beta\delta_t,$$

since $j(\mu)(f) = \alpha f(s) + \beta f(t)$ for any $f \in C([0,1])$, we obtain $I_{C([0,1])}(g) = j(\mu)$ γ^W–a.e. Moreover, by (6.3.3)

$$\begin{aligned}\|g\|^2_{L^2(0,1)} &= \mathbb{E}\left[\Big(I_{C([0,1])}(g)\Big)^2\right] = \|I_{C([0,1])}(g)\|^2_{L^2(C([0,1]),\gamma^W)} \\ &= \|j(\mu)\|^2_{L^2(C([0,1]),\gamma^W)}.\end{aligned}$$

For any simple function $g \in \mathscr{S}([0,1])$

$$g(x) = \sum_{i=1}^n c_i \mathbb{1}_{[0,t_i)}(x),$$

a similar computation yields $\|g\|_{L^2(0,1)} = \|j(\mu)\|_{L^2(C([0,1]),\gamma^W)}$ with

$$\mu = \sum_{i=1}^n c_i \delta_{t_i},$$

and $I_{C([0,1])}(g) = j(\mu)$ γ^W–a.e.. Approximating any $g \in L^2(0,1)$ by a sequence of simple functions g_n, $I_{C([0,1])}(g_n)$ converges to $I_{C([0,1])}(g)$ in $L^2(C([0,1]),\gamma^W)$ by the Itô isometry (6.3.3). Since $I_{C([0,1])}(g_n) \in j(\mathscr{M}([0,1]))$ for every $n \in \mathbb{N}$, we have $I_{C([0,1])}(g) \in X^*_{\gamma^W}$.

This proves that $I_{C([0,1])}(L^2(0,1)) \subset X^*_{\gamma^W}$.

For the reverse inclusion, we use Lemma 6.1.1. Let us take $\mu \in \mathscr{M}([0,1])$ and let (μ_n) be the approximating sequence defined by (6.1.1). Then γ^W–a.e. $j(\mu_n) = I_{C([0,1])}(g_n)$, where

$$\begin{aligned}
g_n(x) =& \mu(\{1\}) + \sum_{i=0}^{2^n-1} \mu\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) 1\!\!1_{[0,\frac{i+1}{2^n})}(x) \\
=& \mu(\{1\}) + \sum_{i=0}^{2^n-1} \mu\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) \sum_{j=0}^{i} 1\!\!1_{[\frac{j}{2^n},\frac{j+1}{2^n})}(x) \\
=& \sum_{j=0}^{2^n-1} \mu(\{1\}) 1\!\!1_{[\frac{j}{2^n},\frac{j+1}{2^n})}(x) + \sum_{j=0}^{2^n-1}\sum_{i=j}^{2^n-1} \mu\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) 1\!\!1_{[\frac{j}{2^n},\frac{j+1}{2^n})}(x) \\
=& \sum_{j=0}^{2^n-1} c_{n,j} 1\!\!1_{[\frac{j}{2^n},\frac{j+1}{2^n})}(x),
\end{aligned} \tag{6.3.6}$$

where

$$c_{n,j} = \mu(\{1\}) + \sum_{i=j}^{2^n-1} \mu\left(\left[\frac{i}{2^n}, \frac{i+1}{2^n}\right)\right) = \mu\left(\left[\frac{j}{2^n}, 1\right]\right).$$

The functions $j(\mu_n)$ converge to $j(\mu)$ in $L^2(C([0,1]),\gamma^W)$ since

$$\begin{aligned}
&\|j(\mu_n) - j(\mu)\|^2_{L^2(C([0,1]),\gamma^W)} = \|j(\mu_n - \mu)\|^2_{L^2(C([0,1]),\gamma^W)} = B_{\gamma^W}(\mu_n - \mu, \mu_n - \mu) \\
&= \int_{[0,1]^2} \min\{x,y\}\big((\mu_n - \mu)\otimes(\mu_n - \mu)\big)(d(x,y)) \\
&= \int_{[0,1]^2} \min\{x,y\}(\mu_n \otimes \mu_n)(d(x,y)) - 2\int_{[0,1]^2} \min\{x,y\}(\mu_n \otimes \mu)(d(x,y)) \\
&\quad + \int_{[0,1]^2} \min\{x,y\}(\mu \otimes \mu)(d(x,y)).
\end{aligned}$$

As μ_n converges to μ weakly, $\mu_n \otimes \mu$ and $\mu_n \otimes \mu_n$ converge to $\mu \otimes \mu$ weakly, we obtain

$$\lim_{n\to+\infty} \|j(\mu_n) - j(\mu)\|^2_{L^2(C([0,1]),\gamma^W)} = 0.$$

Hence $j(\mu)$ is the limit in $L^2(C([0,1]),\gamma^W)$ of $j(\mu_n) = I_{C([0,1])}(g_n)$, whence $j(\mathscr{M}([0,1])) \subset I_{C([0,1])}(L^2(0,1))$. Hence $X^*_{\gamma^W} \subset I_{C([0,1])}(L^2(0,1))$, and this concludes the proof of the equality $X^*_{\gamma^W} = I_{C([0,1])}(L^2(0,1))$.

We now prove equality $X^*_{\tilde{\gamma}^W} = I_{L^2(0,1)}(L^2(0,1))$. For this purpose, let us define $J : L^2(0,1) \to L^2(0,1)$

$$J(f)(s) := \int_s^1 f(\tau)d\tau.$$

If $f = \mathbb{1}_{[0,t)}$, then $J(\mathbb{1}_{[0,t)})(s) = h_t(s) = (t-s)\mathbb{1}_{[0,t)}(s)$. On the other hand

$$j(f)(g) = j(\mathbb{1}_{[0,t)})(g) = \int_0^t g(\tau)d\tau.$$

We claim that $I_{L^2(0,1)}(J(\mathbb{1}_{[0,t)})) = j(\mathbb{1}_{[0,t)})$. In order to prove that, we approximate h_t using the simple functions

$$h_n(s) = \sum_{i=0}^{n-1} t\frac{n-i-1}{n}\mathbb{1}_{[\frac{i}{n}t, \frac{(i+1)}{n}t)}(s).$$

Since h_n converges to h_t in $L^2(0,1)$, $I_{L^2(0,1)}(h_n)$ converges to $I_{L^2(0,1)}(h_t)$ in $L^2(L^2(0,1), \tilde{\gamma}^W)$ thanks to (6.3.3). Let us consider now $\omega \in C_0([0,1])$ and $g = \iota(\omega)$; then

$$\begin{aligned}
I_{L^2(0,1)}(h_n)(g) &= \sum_{i=0}^{n-1}\int_0^t t\frac{n-i-1}{n}\mathbb{1}_{[\frac{it}{n}, \frac{(i+1)t}{n})}(s)d\widetilde{W}_s(g) \\
&= \sum_{i=0}^{n-1} t\frac{n-i-1}{n}\Big(\widetilde{W}_{\frac{(i+1)t}{n}}(g) - \widetilde{W}_{\frac{it}{n}}(g)\Big) \\
&= \sum_{i=0}^{n-1} t\frac{n-i-1}{n}\Big(W_{\frac{(i+1)t}{n}}(\omega) - W_{\frac{it}{n}}(\omega)\Big) \\
&= \sum_{i=0}^{n-1} t\frac{n-i-1}{n}\left(\omega\Big(\frac{(i+1)t}{n}\Big) - \omega\Big(\frac{it}{n}\Big)\right) = \sum_{i=1}^{n-1}\omega\Big(\frac{it}{n}\Big)\frac{t}{n}.
\end{aligned}$$

As a consequence,

$$\lim_{n\to\infty} I_{L^2(0,1)}(h_n)(g) = \int_0^t g(\tau)d\tau, \qquad \forall g \in \iota(C_0([0,1])).$$

Since $\gamma^W(C_0([0,1])) = 1$, we can conclude that $I_{L^2(0,1)}(h_t)(g) = \int_0^t g(\tau)d\tau$ for $\widetilde{\gamma}^W$–a.e. $g \in L^2(0,1)$.

Let us now show the equality $X^*_{\tilde{\gamma}^W} = I_{L^2(0,1)}(L^2(0,1))$. From what we have proved, it follows

$$X^*_{\tilde{\gamma}^W} = \overline{j(L^2(0,1))} \subset I_{L^2(0,1)}(L^2(0,1)).$$

On the other hand, we have that $C^1_c(0,1)$ is dense both in $L^2(0,1)$ and $J(L^2(0,1))$. Then

$$I_{L^2(0,1)}(L^2(0,1)) = \overline{I_{L^2(0,1)}(C^1_c(0,1))} \subset \overline{I_{L^2(0,1)}(J(L^2(0,1)))} = \overline{j(L^2(0,1))} = X^*_{\tilde{\gamma}^W}.$$

$\square$

Remark 6.3.4. By Proposition 3.1.2, a function $f \in C([0,1])$ belongs to the Cameron–Martin space H if and only if it belongs to the range of R_{γ^W}, namely if and only if $f = R_{\gamma^W}(I_{C([0,1])}(g))$ for some $g \in L^2(0,1)$. In this case, by (3.1.4) we have

$$|f|_H = \|I_{C([0,1])}(g)\|_{L^2(C([0,1]),\gamma^W)},$$

and by the Itô isometry (6.3.3)

$$|f|_H = \|g\|_{L^2(0,1)}.$$

The same argument holds true for $f \in L^2(0,1)$.

We close this lecture by characterising the spaces

$$R_{\gamma^W}(j(\mathscr{M}([0,1])))$$

and

$$R_{\widetilde{\gamma}^W}(j(L^2(0,1))) = Q(L^2(0,1)).$$

The latter is easier to describe. We have indeed the following result.

Proposition 6.3.5. *Let $\widetilde{\gamma}^W$ be the Wiener measure on $L^2(0,1)$. Then*

$$Q(L^2(0,1)) = \{u \in H^1_0([0,1]) \cap H^2(0,1) : u'(1) = 0\}$$

and for any $f \in L^2(0,1)$, $u = Qf$ is the solution of the problem on $(0,1)$

$$\begin{cases} u'' = -f \\ u(0) = 0, \\ u'(1) = 0. \end{cases} \tag{6.3.7}$$

Proof. If $u = Qf$, then

$$u(x) = Qf(x) = \int_0^1 \min\{x,y\} f(y)dy.$$

Then u is weakly differentiable and

$$u'(x) = \int_x^1 f(y)dy.$$

Hence u' admits a continuous version such that $u'(1) = 0$; u' is also a.e. differentiable with

$$u''(x) = -f(x).$$

On the other hand, arguing as in the proof of Proposition 6.2.1, if u is a solution of (6.3.7), integrating twice we obtain

$$u(x) = \int_0^1 \min\{x,y\} f(y)dy = Qf(x),$$

and this completes the proof. □

To prove a similar result in the case of γ^W on $C([0,1])$, we need the following lemma.

Lemma 6.3.6. *Let $v \in L^2(0,1)$ be such that $v' = \mu \in \mathscr{M}([0,1])$ in the sense of distributions, i.e.,*

$$\int_0^1 v(x)\varphi'(x)dx = -\int_{[0,1]} \varphi(x)\mu(dx), \qquad \forall \varphi \in C_c^1(0,1).$$

Then there exists $c \in \mathbb{R}$ such that for a.e. $x \in (0,1)$

$$v(x) = \mu((0,x]) + c. \tag{6.3.8}$$

Proof. Let us set $w(x) = \mu((0,x])$. We claim that $w' = \mu$ is the sense of distributions. Indeed, for any $\varphi \in C_c^1(0,1)$, by the Fubini Theorem 1.1.17

$$\begin{aligned}\int_0^1 w(x)\varphi'(x)dx &= \int_0^1 \mu((0,x])\varphi'(x)dx = \int_0^1 \left(\int_{(0,1)} \mathbb{1}_{(0,x]}(y)\mu(dy)\right)\varphi'(x)dx \\ &= \int_{(0,1)} \left(\int_0^1 \mathbb{1}_{(0,x]}(y)\varphi'(x)\,dx\right)\mu(dy) = -\int_{(0,1)} \varphi(y)\mu(dy).\end{aligned}$$

As a consequence, the weak derivative of $v - w$ is zero and the conclusion holds. □

Definition 6.3.7. *Let $v \in L^2(0,1)$ be such that $v' \in \mathscr{M}([0,1])$ in the sense of distributions. Then, writing $v(x) = \mu((0,x]) + c$ for a.e $x \in (0,1)$ as in (6.3.8), we set*

$$v(1^-) := c + \mu((0,1)).$$

We close this lecture with the following proposition.

Proposition 6.3.8. *Let γ^W be the Wiener measure on $C([0,1])$. Then*

$$\begin{aligned}R_{\gamma^W}(j(\mathscr{M}([0,1]))) = \big\{u \in H_0^1([0,1]) :\exists \mu \in \mathscr{M}([0,1]) \text{ s.t. } u'' = -\mu \text{ on } (0,1) \\ \text{in the sense of distributions, } u'(1^-) = \mu(\{1\})\big\}\end{aligned}$$

and $u = R_{\gamma^W}(j(\mu))$ if and only if u is the solution of the following problem on $(0,1)$

$$\begin{cases} u'' = -\mu, \\ u(0) = 0, \\ u'(1^-) = \mu(\{1\}). \end{cases} \tag{6.3.9}$$

Proof. Let $u = R_{\gamma^W}(j(\mu))$. Then for any $\nu \in \mathscr{M}([0,1])$ we have

$$\begin{aligned}\nu(u) = \nu(R_{\gamma^W}(j(\mu))) &= \int_{C([0,1])} j(\mu)(f)j(\nu)(f)\gamma^W(df) \\ &= \int_{C([0,1])} \int_{[0,1]} f(x)\mu(dx) \int_{[0,1]} f(y)\nu(dy)\gamma^W(df) \\ &= \int_{[0,1]^2} \left(\int_{C([0,1])} W_x(f)W_y(f)\gamma^W(df)\right)(\mu\otimes\nu)(d(x,y)) \\ &= \int_{[0,1]^2} \min\{x,y\}(\mu\otimes\nu)(d(x,y)),\end{aligned}$$

where we have used the fact that W_t is a standard Brownian motion. Indeed, since if $x < y$, W_x and $W_y - W_x$ are independent,

$$\begin{aligned}\int_{C([0,1])} W_x(f)W_y(f)\gamma^W(df) &= \int_{C([0,1])} W_x(f)\big(W_y(f) - W_x(f)\big)\gamma^W(df)+ \\ &\quad + \int_{C([0,1])} W_x(f)^2\gamma^W(df) \\ &= \int_{C([0,1])} W_x(f)\gamma^W(df)\cdot \int_{C([0,1])} \big(W_y(f) - W_x(f)\big)\gamma^W(df)+ \\ &\quad + \int_{C([0,1])} W_x(f)^2\gamma^W(df) = x.\end{aligned}$$

Then

$$u(y) = R_{\gamma^W}(j(\mu))(y) = \int_{[0,1]} \min\{x, y\}\mu(dx). \tag{6.3.10}$$

We claim that for a.e. $y \in (0, 1)$

$$u'(y) = \mu((y, 1]) = -\mu((0, y]) + \mu((0, 1]) \tag{6.3.11}$$

and that $u' \in L^2(0, 1)$. Indeed, for any $\varphi \in C^1_c(0, 1)$,

$$\begin{aligned}\int_0^1 u(x)\varphi'(x)dx &= \int_{[0,1]} \int_0^1 \min\{x, y\}\varphi'(x)dx\,\mu(dy) \\ &= -\int_{[0,1]} \int_0^y \varphi(x)dx\,\mu(dy) = -\int_0^1 \varphi(x) \int_{[0,1]} \mathbb{1}_{(0,y)}(x)\mu(dy)\,dx \\ &= -\int_0^1 \varphi(x) \int_{[0,1]} \mathbb{1}_{(x,1]}(y)\mu(dy)\,dx = -\int_0^1 \varphi(x)\mu((x, 1])\,dx.\end{aligned}$$

It is readily verified that $u'(x) = \mu((x, 1])$ belongs to $L^2(0, 1)$; in addition $u'' = -\mu$ on $(0, 1)$ in the sense of distributions. Indeed for any $\varphi \in C^1_c(0, 1)$

$$\begin{aligned}\int_0^1 u'(x)\varphi'(x)dx &= \int_{[0,1]} \int_0^1 \mathbb{1}_{(0,y)}(x)\varphi'(x)dx\,\mu(dy) \\ &= \int_{[0,1]} \varphi(y)\mu(dy) = \int_{(0,1)} \varphi(y)\mu(dy),\end{aligned}$$

where the last equality follows from the fact that $\varphi(0) = \varphi(1) = 0$. Comparing (6.3.11) with (6.3.8) in Lemma 6.3.6, we obtain

$$\lim_{y\to 1^-} u'(y) = \mu(\{1\}).$$

On the other hand, if u is a solution of (6.3.9), then since $u'' = -\mu$, by (6.3.8) in Lemma 6.3.6 we have for a.e. $x \in (0, 1)$

$$u'(x) = -\mu((0, x]) + c = \mu((x, 1)) - \mu((0, 1)) + c.$$

Then $u'(1^-) = \mu(\{1\})$ if and only if $c = \mu(\{1\}) + \mu((0,1))$, and in this case $u'(x) = \mu((x,1])$ for a.e. $x \in (0,1)$. Integrating u' between 0 and x, we get

$$\begin{aligned} u(x) &= \int_0^x \mu((t,1])dt = \int_0^1 1\!\!1_{(0,x)}(t)\mu((t,1])dt = \int_0^1 \int_{[0,1]} 1\!\!1_{(0,x)}(t) 1\!\!1_{(t,1]}(y)\mu(dy)\, dt \\ &= \int_{[0,1]} \int_0^1 1\!\!1_{(t,1]}(x) 1\!\!1_{(t,1]}(y) dt\, \mu(dy) = \int_{[0,1]} \min\{x,y\}\mu(dy), \end{aligned}$$

and then $u = R_{\gamma W}(j(\mu))$. This completes the proof. $\square$

6.4 Exercises

Exercise 6.1. Let $\mu \in \mathscr{M}([0,1])$ and let μ_n be the sequence defined by (6.1.1); prove that $|\mu_n|([0,1]) \le |\mu|([0,1])$.

Exercise 6.2. Verify that the eigenvalues of Q in Theorem 6.2.1 are given by (6.2.4) and that the eigenfunctions with unit norm are given by (6.2.5).

Exercise 6.3. Prove that the stochastic processes defined in (6.3.4) are standard Brownian motions, i.e., they satisfy Definition 5.2.4.

Exercise 6.4. Prove that the function g_n in formula (6.3.6) is given by

$$g_n(x) = \mu\left(\left[\frac{[2^n x]}{2^n}, 1\right]\right),$$

where $[2^n x]$ is the integer part of $2^n x$. Prove that if $\mu \in \mathscr{M}([0,1])$, $j(\mu) = I_{C([0,1])}(g)$ where

$$g(x) = \mu([x,1]), \qquad \text{for a.e. } x \in [0,1].$$

Exercise 6.5. Prove that for any $\mu \in \mathscr{M}([0,1])$, the function

$$u(y) = \int_{[0,1]} \min\{x,y\}\mu(dx), \qquad y \in [0,1],$$

is continuous.

Lecture 7

Finite dimensional approximations

In this Lecture we present some techniques that allow to get infinite dimensional results through finite dimensional arguments and suitable limiting procedures. They rely on factorising X as the direct sum of a finite dimensional subspace F and a topological complement X_F. The finite dimensional space F is a subspace of the Cameron-Martin space H. To define the projection onto F, we use an othonormal basis of H, so that we get at the same time an orthogonal decomposition $H = F \oplus F^{\perp}$ of H. Throughout this Lecture, X is a separable Banach space endowed with a centred Gaussian measure γ.

7.1 Cylindrical functions

In analogy to cylindrical sets discussed in Section 2.1, cylindrical functions play an important role in the infinite dimensional Gaussian analysis.

Definition 7.1.1 (Cylindrical functions)**.** *We say that $\varphi : X \to \mathbb{R}$ is a* cylindrical function *if there are $n \in \mathbb{N}$, $\ell_1, \dots, \ell_n \in X^*$ and a function $\psi : \mathbb{R}^n \to \mathbb{R}$ such that $\varphi(x) = \psi(\ell_1(x), \dots, \ell_n(x))$ for all $x \in X$. For $k \in \mathbb{N}$, we write $\varphi \in \mathcal{F}C_b^k(X)$ (resp. $\varphi \in \mathcal{F}C_b^{\infty}(X)$), and we say that φ is a cylindrical k times (resp. infinitely many times) boundedly differentiable function, if, with the above notation, $\psi \in C_b^k(\mathbb{R}^n)$ (resp. $\psi \in C_b^{\infty}(\mathbb{R}^n)$).*

We fix now an orthonormal basis $\{\hat{h}_i : \ i \in \mathbb{N}\}$ of X^*_{γ}. By Lemma 3.1.8, we may assume that each $\hat{h}_i$ belongs to $j(X^*)$, i.e., $\hat{h}_i = j(\ell_i)$ with $\ell_i \in X^*$. We recall that the set $\{h_i : \ i \in \mathbb{N}\}$, with $h_i = R_{\gamma}\hat{h}_i$, is an orthonormal basis of H.

We need a preliminary result; let us denote by

$$\mathscr{N} = \{B \in \mathscr{E}(X) : \gamma(B) = 0\}$$

the family of Borel sets of null measure. We recall that by Theorem 2.1.1 $\mathscr{E}(X) = \mathscr{B}(X)$.

Proposition 7.1.2. *The following equality holds:*

$$\mathscr{E}(X) = \sigma(\mathscr{E}(X, \{\ell_i : \ i \in \mathbb{N}\}) \cup \mathscr{N}),$$

where $\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\})\cup\mathscr{N}=\{F=E\cup B:\ E\in\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\}),\ B\in\mathscr{N}\}$. *In addition, for any* $E\in\mathscr{E}(X)$ *and* $\varepsilon>0$ *there exist* $F\in\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\})$ *and* $B\in\mathscr{N}$ *with* $\gamma(E\triangle(F\cup B)<\varepsilon$.

Proof. The inclusion $\mathscr{E}(X)\supset\sigma(\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\})\cup\mathscr{N})$ is trivial, so let us prove the reverse inclusion. Let $x^*\in X^*$ and let x^*_n be a finite linear combination of elements of $\{\ell_i:\ i\in\mathbb{N}\}$ such that $j(x^*_n)\to j(x^*)$ in $L^2(X,\gamma)$ and $j(x^*_n)\to j(x^*)$ a.e. Then there exists $N\in\mathscr{N}$ such that $1\!\!1_{X\setminus N}j(x^*_n)\to 1\!\!1_{X\setminus N}j(x^*)$ pointwise. The function $1\!\!1_{X\setminus N}j(x^*_n)$ is $\sigma(\mathscr{E}(X,\{\ell_i\}_{i\in\mathbb{N}})\cup\mathscr{N})$–measurable for all $n\in\mathbb{N}$, and therefore $1\!\!1_{X\setminus N}j(x^*)$ is $\sigma(\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\})\cup\mathscr{N})$–measurable. The $\mathscr{E}(X)$–measurability of $j(x^*)$ implies that $1\!\!1_N j(x^*)$ is also $\sigma(\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\})\cup\mathscr{N})$–measurable, hence $j(x^*)=1\!\!1_{X\setminus N}j(x^*)+1\!\!1_N j(x^*)$ is $\sigma(\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\})\cup\mathscr{N})$–measurable. In conclusion, x^* is $\sigma(\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\})\cup\mathscr{N})$–measurable too. Since this holds for all $x^*\in X^*$, we obtain the desired inclusion.

For the last assertion, notice that the Carathéodory extension of the restriction of γ to $\mathscr{E}(X,\{\ell_i:\ i\in\mathbb{N}\})\cup\mathscr{N}$ is γ again. □

We define

$$P_nx=\sum_{i=1}^n \hat h_i(x)h_i,\quad n\in\mathbb{N},\ x\in X. \tag{7.1.1}$$

Note that every P_n is a projection, since by (2.3.6) $\hat h_i(h_l)=\delta_{il}$. Moreover, if $x\in H$, then $\hat h_i(x)=[x,h_i]_H$ so that P_nx is just a natural extension to X of the orthogonal projection of H onto span $\{h_1,\dots,h_n\}$.

We state (without proof) a deep result on finite dimensional approximations.

Theorem 7.1.3. *For* γ*-a.e.* $x\in X$, $\lim\limits_{n\to\infty}P_nx=x$.

The proof of theorem 7.1.3 may be the subject of one of the projects of Phase 2. However, it is easy if X is a Hilbert space, $\gamma=\mathscr{N}(0,Q)$, and we choose as usual an orthonormal basis $\{e_i:\ i\in\mathbb{N}\}$ of X consisting of eigenvectors of Q, $Qe_k=\lambda_ke_k$. Let us first assume that Q is nondegenerate, i.e. $\lambda_k>0$ for any $k\in\mathbb{N}$. Then $\{h_i=\lambda_i^{1/2}e_i,\ i\in\mathbb{N}\}$ is an orthonormal basis of H and we have $\hat h_i(x)=\langle x,e_i\rangle/\lambda_i^{1/2}$ for every $x\in X$. Indeed, for every $x\in H$,

$$\hat h_i(x)h_i=[x,h_i]_Hh_i=\langle Q^{-1/2}x,Q^{-1/2}Q^{1/2}e_i\rangle Q^{1/2}e_i=\langle x,e_i\rangle e_i. \tag{7.1.2}$$

Since for every $x\in X$ we have $x=\sum_{k=1}^\infty\langle x,e_k\rangle e_k$ and the partial sums of this series are in H, the space H is dense in X. Therefore, equality (7.1.2) holds for every $x\in X$ and P_nx is the orthogonal (in X) projection of x onto span $\{e_1,\dots,e_n\}=$ span $\{h_1,\dots,h_n\}$, which goes to x as $n\to\infty$ for every $x\in X$. Let now Q be degenerate and set

$$X_1=\overline{\operatorname{span}\{e_k:\lambda_k>0\}},\quad X_2=X_1^\perp.$$

We can then define γ_1 on X_1 as

$$\gamma_1=\bigotimes_{\lambda_k>0}\mathscr{N}(0,\lambda_k)$$

and $\gamma_2=\delta_0$ on X_2. A direct computation shows that γ and $\gamma_1\otimes\gamma_2$ have the same covariance operator, and then they are equal.

7.2 Some more notions from Probability Theory

In this section we recall some further notions of probability theory, in particular conditional expectation. We use the notation of Lecture 5.

Let us introduce the notion of *conditional expectation*.

Theorem 7.2.1. *Given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a sub-$\sigma$-algebra $\mathscr{G} \subset \mathscr{F}$ and $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, there exists a unique a random variable $Y \in L^1(\Omega, \mathscr{G}, \mathbb{P})$ such that*

$$\int_A Y \, d\mathbb{P} = \int_A X \, d\mathbb{P}, \qquad \forall A \in \mathscr{G}. \tag{7.2.1}$$

Such random variable is called expectation of X conditioned by $\mathscr{G}$*, and it is denoted by $Y = \mathbb{E}(X|\mathscr{G})$. Moreover, $|\mathbb{E}(X|\mathscr{G})| \le \mathbb{E}(|X|\,|\mathscr{G})$.*

Proof. The map $B \mapsto \int_B X \, d\mathbb{P}$, $B \in \mathscr{G}$, defines a measure that is absolutely continuous with respect to the restriction of $\mathbb{P}$ to $\mathscr{G}$. The assertions then follow from the Radon-Nikodym Theorem 1.1.11. □

Remark 7.2.2. Using approximations by simple functions, we have that (7.2.1) implies

$$\int_\Omega g X \, d\mathbb{P} = \int_\Omega g \, \mathbb{E}(X|\mathscr{G}) d\mathbb{P}$$

for any bounded $\mathscr{G}$–measurable functions $g : \Omega \to \mathbb{R}$.

We list some useful properties of conditional expectation. The proofs are easy consequences of the definition and are left as an exercise, see Exercise 7.3.

Proposition 7.2.3. *The conditional expectation satisfies the following properties.*

1. *If $\mathscr{G} = \{\emptyset, \Omega\}$, then $\mathbb{E}(X|\mathscr{G}) = \mathbb{E}[X]$.*
2. *$\mathbb{E}[\mathbb{E}(X|\mathscr{G})] = \mathbb{E}[X]$.*
3. *For any X, Y and $\alpha, \beta \in \mathbb{R}$, $\mathbb{E}(\alpha X + \beta Y|\mathscr{G}) = \alpha\mathbb{E}(X|\mathscr{G}) + \beta\mathbb{E}(Y|\mathscr{G})$.*
4. *If $X \le Y$, then $\mathbb{E}(X|\mathscr{G}) \le \mathbb{E}(Y|\mathscr{G})$; in particular, if $X \ge 0$, then $\mathbb{E}(X|\mathscr{G}) \ge 0$.*
5. *If $\mathscr{H} \subset \mathscr{G}$ is a sub-σ-algebra of $\mathscr{G}$, then*
$$\mathbb{E}(\mathbb{E}(X|\mathscr{G})|\mathscr{H}) = \mathbb{E}(X|\mathscr{H}).$$
6. *If X is $\mathscr{G}$-measurable, then $\mathbb{E}(X|\mathscr{G}) = X$.*
7. *If $X, Y, X \cdot Y \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ and X is $\mathscr{G}$-measurable, then*
$$\mathbb{E}(X \cdot Y|\mathscr{G}) = X \cdot \mathbb{E}(Y|\mathscr{G}).$$
8. *If X is independent of $\mathscr{G}$, then $\mathbb{E}(X|\mathscr{G}) = \mathbb{E}[X]$.*

The following result allows to handle conditional expectations in L^p spaces, $1 \le p < \infty$.

Theorem 7.2.4 (Jensen)**.** *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $\mathscr{G} \subset \mathscr{F}$ be a sub-σ-algebra, let $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ be a real random variable, and let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex C^1 function such that $\varphi(X) \in L^1(\Omega, \mathscr{F}, \mathbb{P})$. Then,*

$$\mathbb{E}(\varphi \circ X|\mathscr{G}) \ge \varphi \circ \mathbb{E}(X|\mathscr{G}). \tag{7.2.2}$$

Proof. As φ is convex, we have that for any $x, y \in \mathbb{R}$

$$\varphi(x) \ge \varphi(y) + \varphi'(y)(x - y).$$

We use this inequality with $x = X$ and $y = \mathbb{E}(X|\mathscr{G})$ and we obtain

$$\varphi(X) \ge \varphi(\mathbb{E}(X|\mathscr{G})) + \varphi'(\mathbb{E}(X|\mathscr{G}))(X - \mathbb{E}(X|\mathscr{G})). \tag{7.2.3}$$

Since $\varphi(\mathbb{E}(X|\mathscr{G}))$ is $\mathscr{G}$–measurable, by property 6 of Proposition 7.2.3 we have that $\mathbb{E}\big(\mathbb{E}(\varphi(X)|\mathscr{G})|\mathscr{G}\big) = \mathbb{E}(\varphi(X)|\mathscr{G})$. In the same way, $\mathbb{E}(X|\mathscr{G})$ is $\mathscr{G}$–measurable and then $\mathbb{E}(X - \mathbb{E}(X|\mathscr{G})|\mathscr{G}) = 0$. Since also $\varphi'(\mathbb{E}(X|\mathscr{G}))$ is $\mathscr{G}$–measurable, by property 7 of Proposition 7.2.3 we also have

$$\mathbb{E}\Big(\varphi'(\mathbb{E}(X|\mathscr{G}))\big(X - \mathbb{E}(X|\mathscr{G})\big)\Big|\mathscr{G}\Big) = \varphi'(\mathbb{E}(X|\mathscr{G})) \cdot \mathbb{E}\big(X - \mathbb{E}(X|\mathscr{G})\big|\mathscr{G}\big) = 0.$$

Then, taking conditional expectation in (7.2.3), we have

$$\mathbb{E}(\varphi(X)|\mathscr{G}) \ge \mathbb{E}(\varphi(\mathbb{E}(X|\mathscr{G}))|\mathscr{G}) + \mathbb{E}\Big(\varphi'(\mathbb{E}(X|\mathscr{G}))(X - \mathbb{E}(X|\mathscr{G}))\Big|\mathscr{G}\Big) = \varphi(\mathbb{E}(X|\mathscr{G})).$$

$\square$

Corollary 7.2.5. *Let $X \in L^p(\Omega, \mathscr{F}, \mathbb{P})$, $1 \le p \le \infty$, be a real random variable. Then, its conditional expectation $\mathbb{E}(X|\mathscr{G})$ given by Theorem 7.2.1 belongs to $L^p(\Omega, \mathscr{F}, \mathbb{P})$ as well, and $\|\mathbb{E}(X, \mathscr{G})\|_{L^p(\Omega,\mathscr{F},\mathbb{P})} \le \|X\|_{L^p(\Omega,\mathscr{F},\mathbb{P})}$.*

Proof. Let us first consider the case $1 \le p < \infty$: Theorem 7.2.4 with $\varphi(x) = |x|^p$ yields

$$\int_\Omega |\mathbb{E}(X|\mathscr{G})|^p \, d\mathbb{P} \le \int_\Omega \mathbb{E}(|X|^p|\mathscr{G}) \, d\mathbb{P} = \int_\Omega |X|^p \, d\mathbb{P}. \tag{7.2.4}$$

The case $p = \infty$ follows by 4. of Proposition 7.2.3. $\square$

Notice that the properties of the conditional expectation listed in Proposition 7.2.3 hold also in $L^p(X, \gamma)$, $1 \le p < \infty$.

7.3 Factorisation of the Gaussian measure

In this section we describe an important decomposition of γ as the product of two Gaussian measures on subspaces. The projections onto finite dimensional subspaces generate a canonical decomposition of the Gaussian measure as follows. Let $F \subset R_\gamma(j(X^*))$ be an n–dimensional subspace and let us denote by P_F the projection onto X with image F (which is given by P_n of (7.1.1) with a suitable choice of an orthonormal basis of H). Define the measure $\gamma_F = \gamma \circ P_F^{-1}$ and notice that $\gamma_F(F) = 1$ since $P_F^{-1}(F) = X$. For any $\zeta \in X^*$

$$\begin{aligned}\widehat{\gamma_F}(\zeta) &= \int_X \exp\{i\zeta(P_F(x))\}\,\gamma(dx) = \int_X \exp\{iP_F^*\zeta(x)\}\,\gamma(dx)\\ &= \exp\left\{-\frac{1}{2}B_\gamma(P_F^*\zeta, P_F^*\zeta)\right\},\end{aligned}$$

hence γ_F is a centred Gaussian measure by Corollary 2.2.7(i), with

$$\begin{aligned}B_{\gamma_F}(\zeta_1,\zeta_2) =&B_\gamma(P_F^*\zeta_1, P_F^*\zeta_2) = \int_X P_F^*\zeta_1(x)P_F^*\zeta_2(x)\gamma(dx)\\ &= \int_X \zeta_1(P_F x)\zeta_2(P_F x)\,\gamma(dx) = \int_X \zeta_1(z)\zeta_2(z)\,\gamma_F(dz)\\ &= \int_F \zeta_1(z)\zeta_2(z)\,\gamma_F(dz) = \langle \zeta_1,\zeta_2\rangle_{L^2(F,\gamma_F)},\end{aligned} \tag{7.3.1}$$

for any $\zeta_1, \zeta_2 \in X^*$. In the same way we define the measure $\gamma_F^\perp = \gamma \circ (I - P_F)^{-1}$ and notice that $\gamma_F^\perp(X_F) = 1$ where $X_F := \ker P_F$. This measure is again a centred Gaussian measure with

$$\begin{aligned}B_{\gamma_F^\perp}(\zeta_1,\zeta_2) =&B_\gamma((I-P_F)^*\zeta_1, (I-P_F)^*\zeta_2) = \int_X (I-P_F)^*\zeta_1(x)(I-P_F)^*\zeta_2(x)\gamma(dx)\\ &= \int_X \zeta_1((I-P_F)x)\zeta_2((I-P_F)x)\gamma(dx) = \int_{X_F} \zeta_1(y)\zeta_2(y)\gamma_F^\perp(dy)\\ &= \langle \zeta_1,\zeta_2\rangle_{L^2(X_F,\gamma_F^\perp)},\end{aligned} \tag{7.3.2}$$

for any $\zeta_1, \zeta_2 \in X^*$. The explicit computations of B_{γ_F} and $B_{\gamma_F^\perp}$ imply that the Cameron–Martin spaces of γ_F and $\gamma_F^\perp$ are respectively equal to F and $F^\perp$, the last being the orthogonal complement of F in H.

Since γ is centred, we have that $j(f)(x) = f(x)$ for any $f \in X^*$. To simplify the notation, we shall write f instead of $j(f)$ also when considered as an element of X^*_γ; in this way we may think to X^* as a subset of X^*_γ. Let us assume that $F = \text{span}\{h_1, \ldots, h_n\}$ with $h_1, \ldots, h_n$ orthonormal and such that $h_k \in R_\gamma(X^*)$. In this way we may use the explicit expression for P_F given by (7.1.1). We can state and prove the following result.

Lemma 7.3.1. *For any $f \in X^*$, we have*

$$P_F(R_\gamma(f)) = R_\gamma(P_F^* f), \quad (I-P_F)(R_\gamma f) = R_\gamma((I-P_F)^* f).$$

As a consequence

$$|R_\gamma(f)|_H^2 = \|P_F^* f\|^2_{L^2(F,\gamma_F)} + \|(I-P_F)^* f\|^2_{L^2(X_F,\gamma_F^\perp)}. \tag{7.3.3}$$

Proof. We know that for any $g \in X^*$, by (7.1.1) and Remark 2.3.7

$$g(P_F(R_\gamma(f))) = \sum_{k=1}^n \hat{h}_k(R_\gamma(f))g(h_k) = \sum_{k=1}^n \langle f, \hat{h}_k\rangle_{L^2(X,\gamma)} g(h_k).$$

On the other hand, we also have

$$P_F^* f(x) = f(P_F x) = \sum_{k=1}^n \hat{h}_k(x) f(h_k) = \sum_{k=1}^n \langle f, \hat{h}_k\rangle_{L^2(X,\gamma)} \hat{h}_k(x).$$

Hence for any $g \in X^*$

$$\begin{aligned} g(R_\gamma(P_F^* f)) =& g\left(R_\gamma\Big(\sum_{k=1}^n \langle f, \hat{h}_k\rangle_{L^2(X,\gamma)}\hat{h}_k\Big)\right) \\ =& g\left(\sum_{k=1}^n \langle f, \hat{h}_k\rangle_{L^2(X,\gamma)} h_k\right) = \sum_{k=1}^n \langle f, \hat{h}_k\rangle_{L^2(X,\gamma)} g(h_k), \end{aligned}$$

and then $P_F(R_\gamma(f)) = R_\gamma(P_F^* f)$. In addition

$$(I-P_F)R_\gamma(f) = R_\gamma(f) - P_F(R_\gamma(f)) = R_\gamma(f) - R_\gamma(P_F^* f) R_\gamma((I-P_F^*)f).$$

Since $H = F \oplus F^\perp$, for $f \in X^*$ we have

$$\begin{aligned} |R_\gamma f|_H^2 =& |P_F R_\gamma f|_H^2 + |(I-P_F)R_\gamma f|_H^2 \\ =& \|P_F^* f\|^2_{L^2(F,\gamma_F)} + \|(I-P_F)^* f\|^2_{L^2(X_F,\gamma_F^\perp)}. \end{aligned}$$

□

We have the following result.

Proposition 7.3.2. *Let $\tilde{\gamma}_F$ the restriction of γ_F to $\mathscr{B}(F)$ and $\tilde{\gamma}_F^\perp$ the restriction of $\gamma_F^\perp$ to $\mathscr{B}(X_F)$. Then equality $\tilde{\gamma}_F \otimes \tilde{\gamma}_F^\perp = \gamma$ holds.*

Proof. We use the fact that $X = F \oplus X_F$ and then for any $\xi \in X^*$

$$\begin{aligned} \widehat{\tilde{\gamma}_F \otimes \tilde{\gamma}_F^\perp}(\xi) =& \int_{F\times X_F} \exp\{i\xi(z+y)\} \tilde{\gamma}_F \otimes \tilde{\gamma}_F^\perp(d(z,y)) \\ =& \int_F \exp\{i\xi(z)\}\gamma_F(dz) \cdot \int_{X_F} \exp\{i\xi(y)\}\gamma_F^\perp(dy) \\ =& \exp\left\{-\frac{1}{2}\Big(B_{\gamma_F}(\xi,\xi) + B_{\gamma_F^\perp}(\xi,\xi)\Big)\right\}. \end{aligned}$$

Taking into account (7.3.1) and (7.3.2), we obtain that

$$
\begin{aligned}
B_{\gamma_F}(\xi,\xi)+B_{\gamma_F^\perp}(\xi,\xi) &= \int_F \xi(z)^2\gamma_F(dz)+\int_{X_F}\xi(y)^2\gamma_F^\perp(dy)\\
&= \int_X\Big(\xi(P_Fx)^2+\xi((I-P_F)x)^2\Big)\gamma(dx)\\
&= \|P_F^*\xi\|^2_{L^2(F,\gamma_F)}+\|(I-P_F)^*\xi\|^2_{L^2(X_F,\gamma_F^\perp)}\\
&= |R_\gamma(\xi)|_H^2 = B_\gamma(\xi,\xi),
\end{aligned}
$$

where we have used identity (7.3.3). □

As a consequence, by the Fubini theorem, setting for every $A\in\mathscr{B}(X)$ and $z\in F$ (as in Remark 1.1.16) $A_z=\{y\in X_F:\ (z,y)\in A\}$, we have $A_z\in\mathscr{B}(X_F)$; in the same way, setting, for any $y\in X_F$, $A^y=\{z\in F:\ (z,y)\in A\}$, $A^y\in\mathscr{B}(F)$ and we have

$$
\gamma(A)=\int_F \gamma_F^\perp(A_z)\,\gamma_F(dz)=\int_{X_F}\gamma_F(A^y)\,\gamma_F^\perp(dy).
$$

7.4 Cylindrical approximations

Now we are ready to study the approximation of a function via cylindrical ones, taking advantage of the tools just presented.

We fix an orthonormal basis $\{h_k,\ k\in\mathbb{N}\}$ of H, $h_k=R_\gamma\hat h_k$ with $\hat h_k\in j(X^*)$ for all $k\in\mathbb{N}$, see Lemma 3.1.8. For every $f\in L^p(X,\gamma)$, $n\in\mathbb{N}$, we define $\mathbb{E}_nf$ as the conditional expectation of f with respect to the σ-algebra Σ_n generated by the random variables $\hat h_1,\ \dots,\ \hat h_n$. Using Proposition 7.3.2, we can explicitly characterise the expectation of a function $f\in L^p(X,\gamma)$ conditioned to Σ_n.

Proposition 7.4.1. *Let $1\le p\le\infty$. For every $f\in L^p(X,\gamma)$ and $n\in\mathbb{N}$ we have*

$$
(\mathbb{E}_nf)(x)=\int_X f(P_nx+(I-P_n)y)\gamma(dy),\quad x\in X. \tag{7.4.1}
$$

Proof. Let us define

$$
f_n(x)=\int_X f(P_nx+(I-P_n)y)\gamma(dy),\quad n\in\mathbb{N},\ x\in X.
$$

Using the factorisation $\gamma=\tilde\gamma_F\otimes\tilde\gamma_F^\perp$, we may also write

$$
f_n(x)=\int_{X_F} f(P_nx+y)\tilde\gamma_F^\perp(dy).
$$

Since for any $B \in \Sigma_n$, $1\!\!1_B(x) = 1\!\!1_B(P_n x)$, we have

$$
\begin{aligned}
\int_B f(x)\gamma(dx) &= \int_X 1\!\!1_B(P_n x) f(P_n x + (I - P_n)x)\gamma(dx) \\
&= \int_{F\times X_F} 1\!\!1_B(z) f(z+y)\tilde{\gamma}_F \otimes \tilde{\gamma}_F^{\perp}(d(z,y)) \\
&= \int_F 1\!\!1_B(z)\Big(\int_{X_F} f(z+y)\tilde{\gamma}_F^{\perp}(dy)\Big)\,\tilde{\gamma}_F(dz) \\
&= \int_F 1\!\!1_B(z)\Big(\int_X f(z+y)\gamma_F^{\perp}(dy)\Big)\,\tilde{\gamma}_F(dz) \\
&= \int_F 1\!\!1_B(z)\Big(\int_X f(z+(I-P_n)y)\gamma(dy)\Big)\,\tilde{\gamma}_F(dz) \\
&= \int_X 1\!\!1_B(z)\Big(\int_X f(z+(I-P_n)y)\gamma(dy)\Big)\,\gamma_F(dz) \\
&= \int_X 1\!\!1_B(P_n x)\Big(\int_X f(P_n x+(I-P_n)y)\gamma(dy)\Big)\,\gamma(dx) \\
&= \int_X 1\!\!1_B(x) f_n(x)\gamma(dx).
\end{aligned}
$$

By Theorem 7.2.1 we deduce that $f_n = \mathbb{E}(f|\Sigma_n)$.

□

Let us come back to the space $\mathbb{R}^\infty$ described in Subsection 4.1. Through $\mathbb{R}^\infty$, we give a description of $\mathscr{E}(X)$.

Lemma 7.4.2. *A set $E \subset X$ belongs to $\mathscr{E}(X)$ if and only if there are $B \in \mathscr{B}(\mathbb{R}^\infty)$ and a sequence $(f_n)_{n\in\mathbb{N}} \subset X^*$ such that*

$$
E = \Big\{x \in X : \ f(x) := (f_n(x)) \in B\Big\}. \tag{7.4.2}
$$

Proof. For every fixed sequence $(f_n) \subset X^*$ the sets of the form (7.4.2) are a σ-algebra, see Exercise 7.5. Then, the family of the sets as in (7.4.2) are in turn a σ-algebra (let us call it $\mathscr{F}$) and the cylinders belong to $\mathscr{F}$, whence $\mathscr{E}(X) \subset \mathscr{F}$.

On the other hand, for any fixed sequence $f = (f_n) \subset X^*$, the family $\mathscr{G}_f$ consisting of all the Borel subsets $B \subset \mathbb{R}^\infty$ such that the set E described in (7.4.2) belongs to $\mathscr{E}(X)$ contains all the cylinders in $\mathbb{R}^\infty$, hence $\mathscr{G}_f \supset \mathscr{E}(\mathbb{R}^\infty)$. But, since the coordinate functions in $\mathbb{R}^\infty$ are continuous and separate the points, from Theorem 2.1.1 it follows that $\mathscr{B}(\mathbb{R}^\infty) = \mathscr{E}(\mathbb{R}^\infty)$. Therefore, the family of sets $E \subset X$ given by (7.4.2) with $B \in \mathscr{G}_f$ is contained in $\mathscr{E}(X)$ for every f as above. Then $\mathscr{E}(X) \supset \mathscr{F}$ and the proof is complete. □

Lemma 7.4.2 easily implies further useful approximation results.

Lemma 7.4.3. *For every $A \in \mathscr{E}(X)$ and $\varepsilon > 0$ there are a cylinder with compact base C and a compact set $B \subset \mathbb{R}^\infty$ such that $\gamma(C \triangle A) < \varepsilon$ and the set E defined via (7.4.2) verifies $E \subset A$ and $\gamma(A \setminus E) < \varepsilon$.*

Proof. Let A be as in (7.4.2). For every $\varepsilon > 0$ there is a cylinder C_0 such that $\gamma(A \triangle C_0) < \varepsilon/2$: for instance, define $B_k = \{y \in \mathbb{R}^\infty : \ y_j = f_j(x), \ x \in A, \ j \le k\}$ and $C_k = f^{-1}(B_k)$, and take $C_0 = C_k$ with k large enough. Since $C_0 = P^{-1}(D_0)$ for some $D_0 \in \mathscr{B}(\mathbb{R}^n)$ and a linear continuous operator $P : X \to \mathbb{R}^n$, it suffices to take a compact set $K \subset D_0$ such that $\gamma \circ P^{-1}(D_0 \setminus K) < \varepsilon/2$ and $C = P^{-1}(K)$.

By Proposition 1.1.6 the measure $\gamma \circ f^{-1}$ is Radon on $\mathbb{R}^\infty$, hence for every $\varepsilon > 0$ there is a compact set $K \subset B$ such that $\gamma \circ f^{-1}(B \setminus K) < \varepsilon$ and it suffices to choose $E = f^{-1}(K)$. □

Proposition 7.4.4. *For every $1 \le p < \infty$ and $f \in L^p(X, \gamma)$ the sequence $\mathbb{E}_n f$ converges to f in $L^p(X, \gamma)$ and γ-a.e. in X.*

Proof. Let us fix $f \in L^p(X, \gamma)$. We know that for any $\varepsilon > 0$ there exists a simple function s_ε,

$$s_\varepsilon = \sum_{i=1}^m c_i 1\!\!1_{A_i}, \qquad A_i \in \mathscr{B}(X), c_i \in \mathbb{R} \setminus \{0\},$$

such that $\|f - s_\varepsilon\|_{L^p(X,\gamma)} < \varepsilon$. By Proposition 7.1.2, for any $i = 1, \dots, m$ there exists $\tilde{A}_i \in \mathscr{E}(X, \{\ell_i\}_{i \in \mathbb{N}})$ with $\gamma(A_i \Delta \tilde{A}_i) = 0$. Here $\ell_i \in X^*$ is such that $j(\ell_i) = \hat{h}_i$ for any $i \in \mathbb{N}$. Since $\mathscr{E}(X, \{\ell_i\}_{i \in \mathbb{N}})$ is the σ–algebra generated by the algebra

$$\Big\{\mathscr{E}(X, F_n) : n \in \mathbb{N}, F_n = \{\ell_1, \dots, \ell_n\}\Big\},$$

for any $i = 1, \dots, m$ there exists n_i and $C_i \in \mathscr{E}(X, F_{n_i})$ with $\gamma(\tilde{A}_i \Delta C_i) \le \frac{\varepsilon^p}{m^p |c_i|^p}$. The choice of the sets C_i implies that, defining

$$\tilde{s}_\varepsilon = \sum_{i=1}^m c_i 1\!\!1_{C_i},$$

we have

$$\begin{aligned}
\|s_\varepsilon - \tilde{s}_\varepsilon\|_{L^p(X,\gamma)} &\le \sum_{i=1}^m |c_i| \|1\!\!1_{A_i} - 1\!\!1_{C_i}\|_{L^p(X,\gamma)} \\
&= \sum_{i=1}^m |c_i| \gamma(A_i \Delta C_i)^p = \varepsilon.
\end{aligned}$$

Let $n \ge \max\{n_i : i = 1, \dots, m\}$. Since $\tilde{s}_\varepsilon$ is $\mathscr{E}(X, F_n)$–measurable, by property 6 of Proposition 7.2.3 we have $\mathbb{E}_n \tilde{s}_\varepsilon = \tilde{s}_\varepsilon$ and then

$$\begin{aligned}
\|f - \mathbb{E}_n f\|_{L^p(X,\gamma)} \le &\|f - s_\varepsilon\|_{L^p(X,\gamma)} + \|s_\varepsilon - \tilde{s}_\varepsilon\|_{L^p(X,\gamma)} + \\
&+ \|\tilde{s}_\varepsilon - \mathbb{E}_n s_\varepsilon\|_{L^p(X,\gamma)} + \|\mathbb{E}_n s_\varepsilon - \mathbb{E}_n f\|_{L^p(X,\gamma)} \\
\le &\|f - s_\varepsilon\|_{L^p(X,\gamma)} + \|s_\varepsilon - \tilde{s}_\varepsilon\|_{L^p(X,\gamma)} + \\
&+ \|\mathbb{E}_n(\tilde{s}_\varepsilon - s_\varepsilon)\|_{L^p(X,\gamma)} + \|\mathbb{E}_n(s_\varepsilon - f)\|_{L^p(X,\gamma)} \\
\le &2\|f - s_\varepsilon\|_{L^p(X,\gamma)} + 2\|s_\varepsilon - \tilde{s}_\varepsilon\|_{L^p(X,\gamma)} < 4\varepsilon,
\end{aligned}$$

where we have used the contractivity property of the conditional expectation. The proof is then completed.

□

As a consequence of the above results, we have the following approximation theorem. Notice that the conditional expectations $\mathbb{E}_n f$ of a function f is invariant under translations along $\ker P_n$, hence it can be identified with a function defined on $F = P_nX$ setting $f_n(y) = \mathbb{E}_n f(x)$, $y \in F$, $y = P_nx$.

Theorem 7.4.5. *For every* $1 \le p < \infty$ *the space* $\mathcal{F}C_b^\infty(X)$ *is dense in* $L^p(X, \gamma)$.

Proof. Fix p and $f \in L^p(X, \gamma)$. Assume first that $f \in L^\infty(X, \gamma)$ (this hypothesis will be removed later). Set

$$f_n(\xi) = \int_X f\Big(\sum_{j=1}^n \xi_j h_j + (I - P_n)y\Big)\,\gamma(dy), \qquad \xi \in \mathbb{R}^n,$$

and notice that $f_n \in L^p(\mathbb{R}^n, \gamma \circ P_n^{-1})$. Given $\varepsilon > 0$, fix $n \in \mathbb{N}$ such that

$$\|f - f_n\|_{L^p(X,\gamma)} < \varepsilon.$$

Since $f_n \in L^p(\mathbb{R}^n, \gamma \circ P_n^{-1})$, there exists $g \in C_b^\infty(\mathbb{R}^n)$ such that

$$\|f_n - g\|_{L^p(\mathbb{R}^n, \gamma\circ P_n^{-1})} < \varepsilon.$$

Each f_n can be approximated in $L^p(\mathbb{R}^n, \gamma \circ P_n^{-1})$ by a sequence $(\psi_{n,j})$ of functions in $C_b^\infty(\mathbb{R}^n)$, e.g. by convolution. Defining the $\mathcal{F}C_b^\infty(X)$ functions $g_{n.j}(x) = \psi_{n,j}(P_nx)$, it is easily checked that the diagonal sequence $g_{n,n}$ converges to f in $L^p(X, \gamma)$. In order to remove the assumption that f is bounded, given any $f \in L^p(X, \gamma)$, just consider a sequence of truncations $f_k = \max\{-k, \min\{k, f\}\}$, $k \in \mathbb{N}$, and proceed as before. □

7.5 Exercises

Exercise 7.1. Let X be a separable Banach space endowed with a centred Gaussian measure γ. Prove that for any choice $h_1, \dots, h_d \in H$, the map $P : X \to \mathbb{R}^d$, $P(x) = (\hat{h}_1(x), \dots, \hat{h}_d(x))$ is a Gaussian random variable with law $\gamma \circ P^{-1} = \mathscr{N}(0, Q)$, $Q_{i,j} = [h_i, h_j]_H$.

Exercise 7.2. Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a convex function. Prove that there are two sequences $(a_n) \subset \mathbb{R}^d$ and $(b_n) \subset \mathbb{R}$ such that

$$\varphi(x) = \sup_{n\in\mathbb{N}}\{a_n \cdot x + b_n\}.$$

Use this fact to prove Theorem 7.2.4 for any convex function $\varphi : \mathbb{R} \to \mathbb{R}$.

Exercise 7.3. Prove the properties of conditional expectation stated in Proposition 7.2.3.

Exercise 7.4. Prove that if $\Omega = (0,1)^2$ with $\mathscr{F} = \mathscr{B}((0,1)^2)$ and $\mathbb{P} = \lambda_2$ the Lebesgue measure in Ω, then by considering $\mathscr{G} = \mathscr{B}((0,1)) \times (0,1)$

$$\mathbb{E}(X|\mathscr{G})(x,y) = \int_0^1 X(x,t)d\lambda_1(t) \qquad \forall\, y \in (0,1).$$

Exercise 7.5. Prove that for every fixed sequence $(f_n) \subset X^*$ the family of sets defined in (7.4.2) is a σ-algebra.

Exercise 7.6. Prove that if $\varphi \in C^\infty(X)$ has compact support in an infinite dimensional Banach space then $\varphi \equiv 0$.

Lecture 8

Zero-One law and Wiener chaos decomposition

In this Lecture we introduce the Hermite polynomials, which provide an orthonormal basis in $L^2(X,\gamma)$. Accordingly, $L^2(X,\gamma)$ is decomposed as the Hilbert sum of the (mutually orthogonal) subspaces $\mathcal{X}_k$ generated by the polynomials of degree $k \in \mathbb{N}$, see Proposition 8.1.9. Knowing explicitly an orthonormal basis in this not elementary setting is a real luxury! The term *chaos* has been introduced by Wiener in [28] and the structure that we discuss here is usually called *Wiener chaos*. Of course, the Hermite polynomials are used in several proofs, including that of the zero-one law. The expression "zero-one law" is used in different probabilistic contexts, where the final statement is that a certain event has probability either 0 or 1. In our case we show that every measurable subspace has measure either 0 or 1.

We work as usual in a separable Banach space X endowed with a centred Gaussian measure γ. The symbols $R_\gamma,\ X_\gamma^*,\ H$ have the usual meaning.

8.1 Hermite polynomials

As first step, we introduce the Hermite polynomials and we present their main properties. We shall encounter them in many occasions; further properties will be presented when needed.

8.1.1 Hermite polynomials in finite dimension

To start with, we introduce the one dimensional Hermite polynomials.

Definition 8.1.1. *The sequence of Hermite polynomials in $\mathbb{R}$ is defined by*

$$H_k(x) = \frac{(-1)^k}{\sqrt{k!}} \exp\{x^2/2\} \frac{d^k}{dx^k} \exp\{-x^2/2\}, \quad k \in \mathbb{N} \cup \{0\},\ x \in \mathbb{R}. \tag{8.1.1}$$

Then, $H_0(x) \equiv 1$, $H_1(x) = x$, $H_2(x) = (x^2-1)/\sqrt{2}$, $H_3(x) = (x^3-3x)/\sqrt{6}$, etc. Some properties of Hermite polynomials are listed below. Their proofs are easy, and left as exercises, see Exercise 8.1.

Lemma 8.1.2. *For every $k \in \mathbb{N}$, H_k is a polynomial of degree k, with positive leading coefficient. Moreover, for every $x \in \mathbb{R}$,*

$$\begin{cases} (i) \quad H_k'(x) = \sqrt{k}H_{k-1}(x) = xH_k(x) - \sqrt{k+1}H_{k+1}(x), \\ \\ (ii) \quad H_k''(x) - xH_k'(x) = -kH_k(x). \end{cases} \tag{8.1.2}$$

Note that formula (ii) says that H_k is an eigenfunction of the one dimensional *Ornstein–Uhlenbeck operator* $D^2 - xD$, with eigenvalue $-k$. This operator will play an important role in the next lectures.

Proposition 8.1.3. *The set of the Hermite polynomials is an orthonormal Hilbert basis in $L^2(\mathbb{R}, \gamma_1)$.*

Proof. We introduce the auxiliary analytic function

$$F : \mathbb{R}^2 \to \mathbb{R}, \quad F(t,x) := e^{-t^2/2+tx}.$$

Since

$$F(x,t) = \sum_{k=0}^{\infty} \frac{1}{k!}\Big(-\frac{t^2}{2} + tx\Big)^k,$$

for every $x \in \mathbb{R}$ the Taylor expansion of F with respect to t, centred at $t = 0$, converges for every $t \in \mathbb{R}$ and we write it as

$$\begin{aligned} F(t,x) = e^{x^2/2}e^{-(t-x)^2/2} &= e^{x^2/2}\sum_{n=0}^{\infty} \frac{t^n}{n!}\frac{\partial^n}{\partial t^n}e^{-(t-x)^2/2}\Big|_{t=0} \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!}\, e^{x^2/2}(-1)^n\frac{d^n}{dx^n}e^{-x^2/2} = \sum_{n=0}^{\infty} \frac{t^n}{\sqrt{n!}}\, H_n(x). \end{aligned}$$

So, for t, $s \in \mathbb{R}$ we have

$$F(t,x)F(s,x) = e^{-(t^2+s^2)/2+(t+s)x} = \sum_{n,m=0}^{\infty} \frac{t^n}{\sqrt{n!}}\frac{s^m}{\sqrt{m!}}H_n(x)H_m(x).$$

Integrating with respect to x in $\mathbb{R}$ and recalling that $\int_{\mathbb{R}} e^{\lambda x}\gamma_1(dx) = e^{\lambda^2/2}$ for every $\lambda \in \mathbb{R}$ we get

$$\int_{\mathbb{R}} F(t,x)F(s,x)\,\gamma_1(dx) = e^{-(t^2+s^2)/2}\int_{\mathbb{R}} e^{(t+s)x}\gamma_1(dx) = e^{ts} = \sum_{n=0}^{\infty}\frac{t^n s^n}{n!},$$

as well as

$$\int_{\mathbb{R}} F(t,x)F(s,x)\,\gamma_1(dx) = \sum_{n,m=0}^{\infty} \frac{t^n}{\sqrt{n!}}\frac{s^m}{\sqrt{m!}} \int_{\mathbb{R}} H_n(x)H_m(x)\,\gamma_1(dx).$$

Comparing the series gives, for every n, $m \in \mathbb{N} \cup \{0\}$,

$$\int_{\mathbb{R}} H_n(x)H_m(x)\,\gamma_1(dx) = \delta_{n,m},$$

which shows that the set of the Hermite polynomials is orthonormal.

Let now $f \in L^2(\mathbb{R}, \gamma_1)$ be orthogonal to all the Hermite polynomials. Since the linear span of $\{H_k : \ k \leq n\}$ is the set of all polynomials of degree $\leq n$, f is orthogonal to all powers x^n. Then, all the derivatives of the holomorphic function

$$g(z) = \int_{\mathbb{R}} \exp\{ixz\} f(x)\, d\gamma_1(x)$$

vanish at $z = 0$, showing that $g \equiv 0$. For $z = t \in \mathbb{R}$, $g(t)$ is nothing but (a multiple of) the Fourier transform of $x \mapsto f(x)e^{-x^2/2}$, which therefore vanishes a.e. So, $f(x) = 0$ a.e., and the proof is complete. □

Next, we define d-dimensional Hermite polynomials.

Definition 8.1.4. *If $\alpha = (\alpha_1, \ldots, \alpha_d) \in (\mathbb{N} \cup \{0\})^d$ is a multiindex, we define the polynomial H_α by*

$$H_\alpha(x) = H_{\alpha_1}(x_1) \cdots H_{\alpha_d}(x_d), \qquad x = (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{8.1.3}$$

Proposition 8.1.5. *The system of Hermite polynomials is an orthonormal Hilbert basis in $L^2(\mathbb{R}^d, \gamma_d)$. Moreover, for every multiindex $\alpha = (\alpha_1, \ldots, \alpha_d)$ the following equality holds,*

$$\Delta H_\alpha(x) - \langle x, \nabla H_\alpha(x)\rangle = -\Big(\sum_{j=1}^{d} \alpha_j\Big) H_\alpha(x). \tag{8.1.4}$$

Proof. Since γ_d is the product measure of d copies of γ_1, and every H_α is a product of one dimensional Hermite polynomials, Proposition 8.1.3 yields $\langle H_\alpha, H_\beta\rangle_{L^2(\mathbb{R}^d,\gamma_d)} = 1$ if $\alpha = \beta$ and $\langle H_\alpha, H_\beta\rangle_{L^2(\mathbb{R}^d,\gamma_d)} = 0$ if $\alpha \neq \beta$. Completeness may be shown by recurrence on d. By Proposition 8.1.3 the statement holds for $d = 1$. Assume that the statement holds for $d = n-1$, and let $f \in L^2(\mathbb{R}^n, \gamma_n)$ be orthogonal to all Hermite polynomials in $\mathbb{R}^n$. The Hermite polynomials in $\mathbb{R}^n$ are all the functions of the form $H_\alpha(x_1, \ldots, x_n) = H_k(x_1)H_\beta(x_2, \ldots, x_n)$ with $k \in \mathbb{N} \cup \{0\}$ and $\beta \in (\mathbb{N} \cup \{0\})^{n-1}$. So, for every $k \in \mathbb{N} \cup \{0\}$ and $\beta \in (\mathbb{N} \cup \{0\})^{n-1}$ we have

$$0 = \langle f, H_\alpha\rangle_{L^2(\mathbb{R}^n,\gamma_n)} = \int_{\mathbb{R}} \left(H_k(x_1) \int_{\mathbb{R}^{n-1}} f(x_1, y)H_\beta(y)\gamma_{n-1}(dy) \right)\gamma_1(dx_1).$$

Then, the function $g(x_1) = \int_{\mathbb{R}^{n-1}} f(x_1, y)H_\beta(y)\gamma_{n-1}(dy)$ is orthogonal in $L^2(\mathbb{R}, \gamma_1)$ to all H_k. By Proposition 8.1.3 it vanishes for a.e. x_1, which means that for a.e. $x_1 \in \mathbb{R}$ the function $f(x_1,\cdot)$ is orthogonal, in $L^2(\mathbb{R}^{n-1},\gamma_{n-1})$, to all Hermite polynomials H_β. By the recurrence assumption, $f(x_1,y)$ vanishes for a.e. $y\in\mathbb{R}^{n-1}$.

For $d=1$ equality (8.1.4) has already been stated in Lemma 8.1.2. For $d\geq 2$ we have

$$
\begin{aligned}
D_jH_\alpha(x) &= H'_{\alpha_j}(x_j)\prod_{h\neq j}H_{\alpha_h}(x_h)\\
\Delta H_\alpha(x) &= \sum_{j=1}^d H''_{\alpha_j}(x_j)\prod_{h\neq j}H_{\alpha_h}(x_h)=\sum_{j=1}^d\Big[x_jH'_{\alpha_j}(x_j)-\alpha_jH_{\alpha_j}(x_j)\Big]\prod_{h\neq j}H_{\alpha_h}(x_h)\\
&=\sum_{j=1}^d x_jD_jH_\alpha(x)-\Big(\sum_{j=1}^d\alpha_j\Big)H_\alpha(x)=\langle x,\nabla H_\alpha(x)\rangle-\Big(\sum_{j=1}^d\alpha_j\Big)H_\alpha(x).
\end{aligned}
$$

□

Let us denote by $\mathcal{X}_k$ the linear span of all Hermite polynomials of degree k. It is a finite dimensional subspace of $L^2(\mathbb{R}^d,\gamma_d)$, hence it is closed. For $f\in L^2(\mathbb{R}^d,\gamma_d)$, we denote by $I_k(f)$ the orthogonal projection of f on $\mathcal{X}_k$, given by

$$
I_k(f)=\sum_{|\alpha|=k}\langle f,H_\alpha\rangle H_\alpha. \tag{8.1.5}
$$

We recall that if $\alpha=(\alpha_1,\ldots,\alpha_n)$ then $|\alpha|=\alpha_1+\ldots+\alpha_n$, so that the degree of H_α is $|\alpha|$. By Proposition 8.1.5 we have

$$
f=\sum_{k=0}^\infty I_k(f), \tag{8.1.6}
$$

where the series converges in $L^2(\mathbb{R}^d,\gamma_d)$.

8.1.2 The infinite dimensional case

Let us define Hermite polynomials in infinite dimension.

Let us fix an orthonormal basis $\{\hat h_j:\ j\in\mathbb{N}\}$ of X^*_γ with $\hat h_j=\ell_j\in X^*$ so that $\{h_j:\ j\in\mathbb{N}\}$ is an orthonormal basis of H. We introduce the set Λ of multi-indices $\alpha\in(\mathbb{N}\cup\{0\})^{\mathbb{N}}$, $\alpha=(\alpha_j)$, with finite length $|\alpha|=\sum_{j=1}^\infty\alpha_j<\infty$. Λ is just the set of all $\mathbb{N}\cup\{0\}$-valued sequences, that are eventually 0.

Definition 8.1.6. (Hermite polynomials) *For every $\alpha\in\Lambda$, $\alpha=(\alpha_j)$, we set*

$$
H_\alpha(x)=\prod_{j=1}^\infty H_{\alpha_j}(\hat h_j(x)),\quad x\in X. \tag{8.1.7}
$$

Note that only a finite number of terms in the above product are different from 1. So, every H_α is a smooth function with polynomial growth at infinity, namely $|H_\alpha(x)|\leq C(1+\|x\|^{|\alpha|})$. Therefore, $H_\alpha\in L^p(X,\gamma)$ for every $1\leq p<\infty$.

Theorem 8.1.7. *The set $\{H_\alpha :\ \alpha \in \Lambda\}$ is an orthonormal basis of $L^2(X,\gamma)$.*

Proof. Let us first show the orthogonality. Let α, β be in Λ, and let $d \in \mathbb{N}$ be such that $\alpha_j = \beta_j = 0$ for every $j > d$. We have (by Exercise 7.1)

$$\begin{aligned}\int_X H_\alpha\, H_\beta\, d\gamma &= \int_X \prod_{j=1}^d H_{\alpha_j}(\hat h_j(x)) H_{\beta_j}(\hat h_j(x))\, \gamma(dx)\\ &= \int_{\mathbb{R}^d} \prod_{j=1}^d H_{\alpha_j}(\xi_j) H_{\beta_j}(\xi_j) \gamma_d(d\xi)\end{aligned}$$

which is equal to 1 if $\alpha_j = \beta_j$ for every j (namely, if $\alpha = \beta$), otherwise it vanishes. The statement follows.

Next, let us prove that the linear span of the H_α with $\alpha \in \Lambda$ is dense in $L^2(X,\gamma)$. By Theorem 7.4.5, the cylindrical functions of the type $f(x) = \varphi(\hat h_1(x), \dots, \hat h_d(x))$ with $d \in \mathbb{N}$ and $\varphi \in C_b(\mathbb{R}^d, \gamma_d)$ are dense in $L^2(X,\gamma)$. So, it is sufficient to approximate such functions. To this aim, we recall that the linear span of the Hermite polynomials in $\mathbb{R}^d$ is dense in $L^2(\mathbb{R}^d, \gamma_d)$, by Proposition 8.1.5; more precisely the sequence

$$\sum_{k=0}^n I_k^{(d)}(\varphi) = \sum_{k=0}^n \sum_{\alpha \in (\mathbb{N}\cup\{0\})^d,\, |\alpha| = k} \langle \varphi, H_\alpha \rangle_{L^2(\mathbb{R}^d,\gamma_d)} H_\alpha$$

converges to φ in $L^2(\mathbb{R}^d, \gamma_d)$ as $n \to \infty$. Set

$$f_n(x) := \sum_{k=0}^n \sum_{\alpha \in (\mathbb{N}\cup\{0\})^d,\, |\alpha| = k} \langle \varphi, H_\alpha \rangle_{L^2(\mathbb{R}^d,\gamma_d)} H_\alpha(\hat h_1(x), \dots, \hat h_d(x)), \quad n \in \mathbb{N},\ x \in X.$$

Since $\gamma \circ (\hat h_1, \dots, \hat h_d)^{-1}$ is the standard Gaussian measure γ_d in $\mathbb{R}^d$,

$$\|f - f_n\|_{L^2(X,\gamma)} = \Big\|\varphi - \sum_{k=0}^n I_k^{(d)}(\varphi)\Big\|_{L^2(\mathbb{R}^d,\gamma_d)}, \quad n \in \mathbb{N},$$

so that $f_n \to f$ in $L^2(X,\gamma)$. □

Definition 8.1.8. *For every $k \in \mathbb{N} \cup \{0\}$ we set*

$$\mathcal{X}_k = \overline{\operatorname{span}\{H_\alpha :\ \alpha \in \Lambda,\ |\alpha| = k\}},$$

where the closure is in $L^2(X,\gamma)$.

For $k = 0$, $\mathcal{X}_0$ is the subset of $L^2(X,\gamma)$ consisting of constant functions. In contrast with the case $X = \mathbb{R}^d$, for any fixed length $k \in \mathbb{N}$ there are infinitely many Hermite polynomials H_α with $|\alpha| = k$, so that $\mathcal{X}_k$ is infinite dimensional. For $k = 1$, $\mathcal{X}_1$ is the closure of the linear span of the functions $\hat h_j$, $j \in \mathbb{N}$, that are the Hermite polynomials H_α with $|\alpha| = 1$. Therefore, it coincides with X_γ^*.

Proposition 8.1.9. *(The Wiener Chaos decomposition)*

$$L^2(X,\gamma) = \bigoplus_{k\in\mathbb{N}\cup\{0\}} \mathcal{X}_k.$$

Proof. Since $\langle H_\alpha, H_\beta\rangle_{L^2(X,\gamma)} = 0$ for $\alpha \neq \beta$, the subspaces $\mathcal{X}_k$ are mutually orthogonal. Moreover, they span $L^2(X,\gamma)$ by Theorem 8.1.7. □

As in finite dimension, we denote by I_k the orthogonal projection onto $\mathcal{X}_k$. So,

$$I_k(f) = \sum_{\alpha\in\Lambda,\,|\alpha|=k} \langle f, H_\alpha\rangle_{L^2(X,\gamma)} H_\alpha, \quad f \in L^2(X,\gamma), \tag{8.1.8}$$

$$f = \sum_{k=0}^{\infty} I_k(f), \quad f \in L^2(X,\gamma), \tag{8.1.9}$$

where the series converge in $L^2(X,\gamma)$.

8.2 The zero-one law

We start this section by presenting an important technical notion that we need later, that of *completion* of a σ-algebra.

Definition 8.2.1. *Let $\mathscr{F}$ be a σ-algebra of subsets of X and let γ be a measure on $(X,\mathscr{F})$. The* completion *of $\mathscr{F}$ is the family*

$$\mathscr{F}_\gamma = \Big\{E \subset X :\ \exists\, B_1, B_2 \in \mathscr{F}\ \ \textit{such that}\ B_1 \subset E \subset B_2,\ \gamma(B_2 \setminus B_1) = 0\Big\}.$$

We leave as an exercise to verify that $\mathscr{F}_\gamma$ is a σ-algebra. The measure γ is extended to $\mathscr{F}_\gamma$ in the natural way. From now on, unless otherwise specified, a set $E \subset X$ is called *measurable* if it belongs to the completed σ-algebra $\mathscr{B}(X)_\gamma = \mathscr{E}(X)_\gamma$. The main result of this section is the following.

Theorem 8.2.2. *If V is a measurable affine subspace of X*[(1)]*, then $\gamma(V) \in \{0,1\}$.*

We need some preliminary results.

Proposition 8.2.3. *If $A \in \mathscr{B}(X)_\gamma$ is such that $\gamma(A+h) = \gamma(A)$ for all $h \in H$, then $\gamma(A) \in \{0,1\}$.*

Proof. Let $\{h_j\}_{j\in\mathbb{N}}$ be an orthonormal basis of H. Then, for every $n \in \mathbb{N}$ the function

$$F(t_1,\dots,t_n) = \gamma(A - t_1h_1 + \ldots - t_nh_n) = \int_A \exp\Big\{\sum_{j=1}^n t_j\hat{h}_j(x) - \frac{1}{2}\sum_{j=1}^n t_j^2\Big\}\gamma(dx)$$

[(1)] By measurable affine subspace we mean a set $V = V_0 + x_0$, with V_0 measurable (linear) subspace and $x_0 \in X$.

is constant. Therefore, for all $\alpha_1, \dots, \alpha_n$ not all 0 we get

$$\frac{\partial^{\alpha_1+\ldots+\alpha_n} F}{\partial t_1^{\alpha_1} \ldots \partial t_n^{\alpha_n}}(0, \dots, 0) = 0.$$

Arguing as in the proof of Proposition 8.1.3 we get

$$\frac{\partial^{\alpha_1+\ldots+\alpha_n}}{\partial t_1^{\alpha_1} \ldots \partial t_n^{\alpha_n}} \exp\Big\{\sum_{j=1}^n t_j \hat{h}_j(x) - \frac{1}{2}\sum_{j=1}^n t_j^2\Big\}\Big|_{t_1=\cdots=t_n=0} = H_{\alpha_1}(\hat{h}_1(x)) \cdot \ldots \cdot H_{\alpha_n}(\hat{h}_n(x))$$

(where H_{α_j} are the 1-dimensional Hermite polynomials), whence

$$\int_X H_{\alpha_1}(\hat{h}_1(x)) \cdot \ldots \cdot H_{\alpha_n}(\hat{h}_n(x)) 1\!\!1_A(x)\gamma(dx) = 0.$$

It follows that the function $1\!\!1_A$ is orthogonal to all nonconstant Hermite polynomials and then by Theorem 8.1.7 it is constant, i.e., either $1\!\!1_A = 0$ or $1\!\!1_A = 1$ a.e. $\square$

Corollary 8.2.4. *If $A \in \mathscr{B}(X)_\gamma$ is such that $\gamma(A \setminus (A + h)) = 0$ for every $h \in H$, then $\gamma(A) \in \{0, 1\}$.*

Proof. Since if $h \in H$ also $-h \in H$, we deduce that $\gamma(A \setminus (A - h)) = 0$ for all $h \in H$ and then by Theorem 3.1.5 $\gamma((A + h) \setminus A) = 0$. In conclusion

$$\gamma(A + h) = \gamma(A), \qquad \forall h \in H$$

and we conclude by applying Proposition 8.2.3. $\square$

Corollary 8.2.5. *If f is a measurable function such that $f(x + h) = f(x)$ a.e., for all $h \in H$, then there exists $c \in \mathbb{R}$ such that $f(x) = c$ for a.e. $x \in X$.*

Proof. By Proposition 8.2.3, for every $t \in \mathbb{R}$ either $\gamma(\{x \in X : f(x) < t\}) = 1$ or $\gamma(\{x \in X : f(x) < t\}) = 0$. Since the function $t \mapsto \gamma(\{x \in X : f(x) < t\})$ is increasing, there exists exactly one $c \in \mathbb{R}$ such that $\gamma(\{x \in X : f(x) < t\}) = 0$ for all $t \le c$ and $\gamma(\{x \in X : f(x) < t\}) = 1$ for all $t > c$. Then,

$$\gamma(\{x \in X : f(x) = c\}) = \lim_{n\to\infty} \gamma\Big(\Big\{x \in X : c - \frac{1}{n} \le f(x) < c + \frac{1}{n}\Big\}\Big) = 1.$$

$\square$

Now we prove our main theorem.

Proof. of Theorem 8.2.2 Let us assume first that V is a linear subspace. If $\gamma(V) = 0$ there is nothing to prove. If $\gamma(V) > 0$, then there exists $A, B \in \mathscr{B}(X)$ with $A \subset V \subset B$ and $\gamma(B \setminus A) = 0$. By Proposition 3.1.6, there exists $r > 0$ such that

$$B^H(0, r) \subset A - A \subset V - V = V,$$

and then $H \subset V$. Hence $V + h = V$ for every $h \in H$ and by Proposition 8.2.3, $\gamma(V) = 1$.

Let now V be an affine subspace. Then, there is x_0 such that $V_0 = V + x_0$ is a vector subspace, hence, applying the result for V_0 to the measure γ_{x_0} we obtain $\gamma(V) = \gamma_{x_0}(V + x_0) = \gamma_{x_0}(V_0) \in \{0, 1\}$. $\square$

8.3 Measurable linear functionals

In this section we give the notion of measurable linear functionals and we prove that such functions are just the elements of X_γ^*.

Definition 8.3.1 (Measurable linear functionals)**.** *We say that $f : X \to \mathbb{R}$ is a* measurable linear functional *or* γ–measurable linear functional *if there exist a measurable subspace $V \subset X$ with $\gamma(V) = 1$ and a γ-measurable function $f_0 : X \to \mathbb{R}$ such that f_0 is linear on V and $f = f_0$ γ-a.e.*

In the above definition $f = f_0$ γ-a.e., so we may modify any measurable linear functional on a negligible set in such a way that the modification is still mesurable, as the σ-algebra $\mathscr{B}(X)$ has been completed, it is defined *everywhere* on a full-measure subspace V and it is linear on V. This will be always done in what follows. As by Theorem 3.1.9(ii), which is easily checked to hold for all measurable subspaces and not only for Borel measurable subspaces, the Cameron-Martin space H is contained in V, all measurable linear functionals will be defined everywhere and linear on H.

Example 8.3.2. Let us exhibit two simple examples of measurable linear functionals which are not continuous, except trivial cases.

(i) Let $f : \mathbb{R}^\infty \to \mathbb{R}$ be the functional defined by

$$f(x) = \sum_{k=1}^{\infty} c_k x_k$$

where $(c_k) \in \ell^2$. Here, as usual $\mathbb{R}^\infty$ is endowed with a countable product of standard 1-dimensional Gaussian measure, see (4.1.1). Indeed, the series defining f converges γ-a.e. in $\mathbb{R}^\infty$. If $\{k : c_k \neq 0\} = \infty$, only the restriction of f to $\mathbb{R}^\infty_c$ is continuous.

(ii) Let X be a Hilbert space endowed with the Gaussian measure $\gamma = \mathscr{N}(0, Q)$, where Q is a selfadjoint positive trace-class operator with eigenvalues $\{\lambda_k : \ k \in \mathbb{N}\}$. Let $\{e_k : \ k \in \mathbb{N}\}$ be an orthonormal basis of eigenvectors of Q in X with $Qe_k = \lambda_k e_k$ for all $k \in \mathbb{N}$. Fix a sequence $(c_k) \subset \mathbb{R}$ such that the series $\sum_k c_k^2 \lambda_k$ is convergent and define the functional

$$f(x) = \sum_{k=1}^{\infty} c_k \langle x, e_k \rangle_X.$$

Then, f is a measurable linear functional on X which is not continuous if $(c_k) \notin \ell^2$, see Exercise 8.4.

We shall call *proper measurable linear functionals* the measurable linear functionals that are linear on X.

Proposition 8.3.3. *Let f be a measurable linear functional and let V be a full measure subspace such that f is linear on V. If $X \setminus V \neq \emptyset$ then there is a modification of f on the γ-negligible set $X \setminus V$ which is proper.*

Proof. If V is a complemented subspace, just put $f = 0$ on the complementary space. If not, we use the existence of a vector (or Hamel) basis in X, i.e., an infinite (indeed, uncountable) linearly independent set of generators, see [12, Theorem 1.4.5]. Notice that the existence of such a basis is equivalent to the axiom of choice or Zorn Lemma. Fix a Hamel basis of V, say $\mathcal{B} = \{e_\alpha : \alpha \in \mathbb{A}\}$ for a suitable set if indices $\mathbb{A}$. Then, complete $\mathcal{B}$ in order to get a basis of X and extend $f_{|V}$ setting $f = 0$ on the added generators. The extension of $f_{|V}$ is different from f on a γ-negligible set and is linear on the whole of X. □

The first result on the measurable linear functionals is the following.

Proposition 8.3.4. *If $f : X \to \mathbb{R}$ is a measurable linear functional, then its restriction to H is continuous with respect to the norm of H.*

Proof. Setting $V_n = \{f \le n\}$, $n \in \mathbb{N}$, since $X = \cup_n V_n$, there is $n_0 \in \mathbb{N}$ such that $\gamma(V_{n_0}) > 0$. By Lemma 3.1.6 there is $r > 0$ such that $B_H(0, r) \subset V_{n_0} - V_{n_0}$, and therefore

$$\sup_{h \in B_H(0,r)} |f(h)| \le 2n_0.$$

□

For the statement of Proposition 8.3.4 to be meaningful, f has to be defined *everywhere* on the subspace V in definition 8.3.1, because H is negligible. Nevertheless, proper functionals are uniquely determined by their values on H.

Lemma 8.3.5. *Let f be a proper measurable linear functional. If $f \in X^*_\gamma$ then*

$$f(h) = [R_\gamma f, h]_H = \int_X f(x)\hat{h}(x)\,\gamma(dx), \qquad \forall h \in H. \tag{8.3.1}$$

Proof. The second equality is nothing but the definition of inner product in H. In order to prove the first one, consider a sequence $(f_n) \subset X^*$ converging to f in $L^2(X, \gamma)$ and fix $h \in H$. By (2.3.6), writing as usual $h = R_\gamma \hat{h}$, we have

$$f_n(h) = f_n(R_\gamma \hat{h}) = \int_X f_n(x)\hat{h}(x)\;\gamma(dx).$$

The right hand side converges to the right hand side of (8.3.1), hence (up to a subsequence that we do not relabel) $f_n \to f$ a.e. Then

$$L = \{x \in X : \; f(x) = \lim_{n\to\infty} f_n(x)\}$$

is a measurable linear subspace of full measure, hence L contains H thanks to Proposition 3.1.9(ii). Therefore, $f(h) = \lim_{n\to\infty} f_n(h)$ and this is true for all $h \in H$. □

Corollary 8.3.6. *If (f_n) is a sequence of proper measurable linear functionals converging to 0 in measure, then their restrictions $f_{n|H}$ converge to 0 uniformly on the bounded subsets of H.*

Proof. Let us first show that the convergence in measure defined in (1.1.4) implies the convergence in $L^2(X,\gamma)$. Indeed, if $f_n \to 0$ in measure then

$$\exp\Big\{-\frac{1}{2}\|f_n\|^2_{L^2(X,\gamma)}\Big\} = \hat\gamma(f_n) \to 1,$$

whence $\|f_n\|_{L^2(X,\gamma)} \to 0$. Therefore, by Lemma 8.3.5

$$|f_n(h)| \le \int_X |f_n(x)||\hat h(x)|\,\gamma(dx) \le \|f_n\|_{L^2(X,\gamma)}|h|_H, \qquad \forall h \in H.$$

To show that the sequence $(f_{n|H})$ converges uniformly on the bounded sets, it is enough to consider the unit ball:

$$\sup_{h\in H, |h|_H\le 1} |f_{n|H}(h)| \le \|f_n\|_{L^2(X,\gamma)} \to 0.$$

□

Proposition 8.3.7. *If f and g are two measurable linear functionals, then either $\gamma(\{f = g\}) = 1$ or $\gamma(\{f = g\}) = 0$. We have $\gamma(\{f = g\}) = 1$ if and only if $f = g$ on H.*

Proof. According to Proposition 8.3.3, we may assume that f and g are proper; then $L = \{f = g\}$ is a measurable vector space. By Theorem 8.2.2 either $\gamma(L) = 0$, or $\gamma(L) = 1$. If $\gamma(L) = 1$ then $H \subset L$ by Proposition 3.1.9(ii) and then $f = g$ on H. Conversely, if $f = g$ in H then the measurable function $\varphi := f - g$ verifies $\varphi(x + h) = \varphi(x)$ for every $h \in H$. By Corollary 8.2.5 $\varphi = c$ a.e., but as φ is linear, $c = 0$. □

Notice that, as a consequence of Proposition 8.3.7, if a measurable linear functional vanishes on a dense subspace of H then it vanishes a.e. Indeed, any measurable linear functional is continuous on H, hence if it vanishes on a dense set then it vanishes everywhere in H.

Theorem 8.3.8. *The following conditions are equivalent.*

(i) $f \in X^*_\gamma$.

(ii) There is a sequence $(f_n)_{n\in\mathbb{N}} \subset X^*$ *that converges to* f *in measure.*

(iii) f *is a measurable linear functional.*

Proof. (i) $\implies$ (ii) is obvious.
(ii) $\implies$ (iii) If $(f_n) \subset X^*$ converges to f in measure, then (up to subsequences that we do not relabel) $f_n \to f$ a.e. and therefore defining

$$V = \{x \in X : \exists \lim_{n\to\infty} f_n(x)\},$$

V is a measurable subspace and $\gamma(V) = 1$, hence we may define also the functional

$$f_0(x) = \lim_{n\to\infty} f_n(x), \qquad x \in V.$$

V and f_0 satisfy the conditions of Definition 8.3.1 and therefore f is a measurable linear functional.
(iii) $\Longrightarrow$ (i) Let f be a measurable linear functional; then by Proposition 8.3.4 its restriction to H is linear and continuous with respect to the norm of H. By Riesz-Frechet Representation Theorem, there exists $g \in X^*_\gamma$ such that

$$f(h) = [R_\gamma g, h]_H, \qquad \forall h \in H.$$

By (3.1.7) $[R_\gamma g, h]_H = g(h)$, and then $f = g$ on H. By the implications (i)$\Longrightarrow$(ii) and (ii)$\Longrightarrow$(iii) we know that g is a measurable linear functional, and then by Proposition 8.3.7 we deduce that $f = g$ γ–a.e. and then $f \in X^*_\gamma$. □

8.4 Exercises

Exercise 8.1. Prove the equalities (8.1.2).

Exercise 8.2. Verify that the family $\mathscr{F}_\gamma$ introduced in Definition 8.2.1 is a σ-algebra. Prove also that the measure γ, extended to $\mathscr{F}_\gamma$ by $\gamma(E) = \gamma(B_1) = \gamma(B_2)$ for E, B_1, B_2 as in Definition 8.2.1, is still a measure.

Exercise 8.3. Prove that if A is a measurable set such that $A + rh_j = A$ up to γ-negligible sets with $r \in \mathbb{Q}$ and $\{h_j : \ j \in \mathbb{N}\}$ an orthonormal basis of H, then $\gamma(A) \in \{0, 1\}$.
Hint: Use the continuity of the map $h \mapsto \gamma(A + h)$ in H.

Exercise 8.4. Prove that the functionals f defined in Example 8.3.2 enjoy the stated properties.
Hint: For the case (ii), prove that $f \in L^2(X, \gamma)$.

Lecture 9

Sobolev Spaces I

9.1 The finite dimensional case

We consider here the standard Gaussian measure $\gamma_d = \mathscr{N}(0, I_d)$ in $\mathbb{R}^d$. As in the case of the Lebesgue measure λ_d, for $1 \le p < \infty$ there are several equivalent definitions of the Sobolev space $W^{1,p}(\mathbb{R}^d, \gamma_d)$. It may be defined as the set of the functions in $L^p(\mathbb{R}^d, \gamma_d)$ having weak derivatives $D_i f$, $i = 1, \dots, d$ in $L^p(\mathbb{R}^d, \gamma_d)$, or as the completion of a set of smooth functions in the Sobolev norm,

$$\|f\|_{W^{1,p}(\mathbb{R}^d,\gamma_d)} := \left(\int_{\mathbb{R}^d} |f|^p d\gamma_d \right)^{1/p} + \left(\int_{\mathbb{R}^d} |\nabla f|^p d\gamma_d \right)^{1/p}. \tag{9.1.1}$$

Such approaches are equivalent. We will follow the second one, which is easily extendable to the infinite dimensional case, and in the infinite dimensional case seems to be the simplest one. To begin with, we exhibit an integration formula for functions in $C^1_b(\mathbb{R}^d)$, the space of bounded continuously differentiable functions with bounded first order derivatives.

Lemma 9.1.1. *For every $f \in C^1_b(\mathbb{R}^d)$ and for every $i = 1, \dots, d$ we have*

$$\int_{\mathbb{R}^d} \frac{\partial f}{\partial x_i}(x)\, \gamma_d(dx) = \int_{\mathbb{R}^d} x_i f(x)\, \gamma_d(dx). \tag{9.1.2}$$

The proof is left as an exercise. Applying Lemma 9.1.1 to the product fg we get the integration by parts formula

$$\int_{\mathbb{R}^d} f \frac{\partial g}{\partial x_i}\, d\gamma_d = -\int_{\mathbb{R}^d} g \frac{\partial f}{\partial x_i}\, d\gamma_d + \int_{\mathbb{R}^d} f(x)g(x)x_i\, \gamma_d(dx), \quad f,\ g \in C^1_b(\mathbb{R}^d), \tag{9.1.3}$$

which is the starting point of the theory of Sobolev spaces.

We recall the definition of a closable operator, and of the closure of a closable operator.

Definition 9.1.2. *Let E, F be Banach spaces and let $L : D(L) \subset E \to F$ be a linear operator. L is called* closable *(in E) if there exists a linear operator $\overline{L} : D(\overline{L}) \subset E \to F$ whose graph is the closure of the graph of L in $E \times F$. Equivalently, L is closable if*

$$(x_n) \subset D(L), \ \lim_{n\to\infty} x_n = 0 \text{ in } E, \ \lim_{n\to\infty} Lx_n = z \text{ in } F \Longrightarrow z = 0. \tag{9.1.4}$$

If L is closable, the domain of the closure $\overline{L}$ of L is the set

$$D(\overline{L}) = \Big\{x \in E : \ \exists (x_n) \subset D(L), \ \lim_{n\to\infty} x_n = x, \ \ Lx_n \text{ converges in } F\Big\}$$

and for $x \in D(\overline{L})$ we have

$$\overline{L}x = \lim_{n\to\infty} Lx_n,$$

for every sequence $(x_n) \subset D(L)$ such that $\lim_{n\to\infty} x_n = x$. Condition (9.1.4) guarantees that $\lim_{n\to\infty} Lx_n$ is independent of the sequence (x_n). Since $\overline{L}$ is a closed operator, its domain is a Banach space with the graph norm $x \mapsto \|x\|_E + \|\overline{L}x\|_F$.

For every $1 \le p < \infty$ we set as usual $p' = p/(p-1)$ if $1 < p < \infty$, $p' = \infty$ if $p = 1$.

Lemma 9.1.3. *For any $1 \le p < \infty$, the operator $\nabla : D(\nabla) = C^1_b(\mathbb{R}^d) \to L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$ is closable in $L^p(\mathbb{R}^d, \gamma_d)$.*

Proof. Let $f_n \in C^1_b(\mathbb{R}^d)$ be such that $f_n \to 0$ in $L^p(\mathbb{R}^d, \gamma_d)$ and $\nabla f_n \to G = (g_1, \dots g_d)$ in $L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$. For every $i = 1, \dots, d$ and $\varphi \in C^1_c(\mathbb{R}^d)$ we have

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \frac{\partial f_n}{\partial x_i} \varphi \, d\gamma_d = \int_{\mathbb{R}^d} g_i \varphi \, d\gamma_d,$$

since

$$\int_{\mathbb{R}^d} \Big| \Big(\frac{\partial f_n}{\partial x_i} - g_i \Big) \varphi \Big| \, d\gamma_d \le \|\partial f_n / \partial x_i - g_i\|_{L^p(\mathbb{R}^d, \gamma_d)} \|\varphi\|_{L^{p'}(\mathbb{R}^d, \gamma_d)}.$$

On the other hand,

$$\int_{\mathbb{R}^d} \frac{\partial f_n}{\partial x_i} \varphi \, d\gamma_d = -\int_{\mathbb{R}^d} f_n \frac{\partial \varphi}{\partial x_i} \, d\gamma_d + \int_{\mathbb{R}^d} x_i f_n(x) \varphi(x) \, \gamma_d(dx), \quad n \in \mathbb{N},$$

so that, since $f_n \to 0$ in $L^p(\mathbb{R}^d, \gamma_d)$ and the functions $x \mapsto \partial\varphi/\partial x_i(x)$, $x \mapsto x_i \varphi(x)$ are bounded,

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \frac{\partial f_n}{\partial x_i} \varphi \, d\gamma_d = 0.$$

So,

$$\int_{\mathbb{R}^d} g_i \, \varphi \, d\gamma_d = 0, \quad \varphi \in C^1_c(\mathbb{R}^d)$$

which implies $g_i = 0$ a.e. □

Lemma 9.1.3 allows to define the Sobolev spaces of order 1, as follows.

Definition 9.1.4. *For every $1 \le p < \infty$, $W^{1,p}(\mathbb{R}^d, \gamma_d)$ is the domain of the closure of $\nabla : C_b^1(\mathbb{R}^d) \to L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$ in $L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$ (still denoted by ∇). Therefore, $f \in L^p(\mathbb{R}^d, \gamma_d)$ belongs to $W^{1,p}(\mathbb{R}^d, \gamma_d)$ iff there exists a sequence of functions $f_n \in C_b^1(\mathbb{R}^d)$ such that $f_n \to f$ in $L^p(\mathbb{R}^d, \gamma_d)$ and ∇f_n converges in $L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$, and in this case, $\nabla f = \lim_{n\to\infty} \nabla f_n$. Moreover we set $\partial f/\partial x_i(x) := \nabla f(x) \cdot e_i$, $i = 1, \dots d$.*

$W^{1,p}(\mathbb{R}^d, \gamma_d)$ is a Banach space with the graph norm

$$\begin{aligned} \|f\|_{W^{1,p}(\mathbb{R}^d,\gamma_d)} &:= \|f\|_{L^p(\mathbb{R}^d,\gamma_d)} + \|\nabla f\|_{L^p(\mathbb{R}^d,\gamma_d;\mathbb{R}^d)} \\ &= \left(\int_{\mathbb{R}^d} |f|^p d\gamma_d\right)^{1/p} + \left(\int_{\mathbb{R}^d} |\nabla f|^p d\gamma_d\right)^{1/p}. \end{aligned} \tag{9.1.5}$$

One could give a more abstract definition of the Sobolev spaces, as the completion of $C_b^1(\mathbb{R}^d)$ in the norm (9.1.1). Since the norm (9.1.1) is stronger than the L^p norm, every element of the completion may be identified in an obvious way with an element f of $L^p(\mathbb{R}^d, \gamma_d)$. However, to define ∇f we need to know that for any sequence (f_n) of C_b^1 functions such that $f_n \to f$ in $L^p(\mathbb{R}^d, \gamma_d)$ and ∇f_n is a Cauchy sequence in $L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$, the sequence of gradients (∇f_n) converges to the same limit in $L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$. In other words, we need Lemma 9.1.3.

Several properties of the spaces $W^{1,p}(\mathbb{R}^d, \gamma_d)$ follow easily.

Proposition 9.1.5. *Let $1 < p < \infty$. Then*

(i) the integration formula (9.1.2) *holds for every $f \in W^{1,p}(X, \gamma_d)$, $i = 1, \dots, d$;*

(ii) if $\theta \in C_b^1(\mathbb{R}^d)$ and $f \in W^{1,p}(\mathbb{R}^d, \gamma_d)$, then $\theta \circ f \in W^{1,p}(\mathbb{R}^d, \gamma_d)$, and $\nabla(\theta \circ f) = (\theta' \circ f)\nabla f$;

(iii) if $f \in W^{1,p}(\mathbb{R}^d, \gamma_d)$, $g \in W^{1,q}(\mathbb{R}^d, \gamma_d)$ with $1/p + 1/q = 1/s \le 1$, then $fg \in W^{1,s}(\mathbb{R}^d, \gamma_d)$ and
$$\nabla(fg) = g\nabla f + f\nabla g;$$

(iv) $W^{1,p}(\mathbb{R}^d, \gamma_d)$ is reflexive;

(v) if $f_n \to f$ in $L^p(\mathbb{R}^d, \gamma_d)$ and $\sup_{n\in\mathbb{N}} \|f_n\|_{W^{1,p}(\mathbb{R}^d,\gamma_d)} < \infty$, then $f \in W^{1,p}(\mathbb{R}^d, \gamma_d)$.

Proof. Statement (i) follows just approximating f by a sequence of functions belonging to $C_b^1(\mathbb{R}^d)$, using (9.1.2) for every approximating function f_n and letting $n \to \infty$.

Statement (ii) follows approximating $\theta \circ f$ by $\theta \circ f_n$, if $f_n \in C_b^1(\mathbb{R}^d)$ is such that $f_n \to f$ in $L^p(\mathbb{R}^d, \gamma_d)$ and $\nabla f_n \to \nabla f$ in $L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$.

Statement (iii) follows easily from the definition, approximating fg by $f_n g_n$ if $f_n \in C_b^1(\mathbb{R}^d)$ are such that $f_n \to f$ in $L^p(\mathbb{R}^d, \gamma_d)$, $\nabla f_n \to \nabla f$ in $L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$, $g_n \to g$, in $L^q(\mathbb{R}^d, \gamma_d)$, $\nabla g_n \to \nabla g$ in $L^q(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$.

The proof of (iv) is similar to the standard proof of the reflexivity of $W^{1,p}(\mathbb{R}^d, \lambda_d)$. The mapping $u \mapsto Tu = (u, \nabla u)$ is an isometry from $W^{1,p}(\mathbb{R}^d, \gamma_d)$ to the product space $E := L^p(\mathbb{R}^d, \gamma_d) \times L^p(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$, which implies that the range of T is closed in E. Now,

$L^p(\mathbb{R}^d,\gamma_d)$ and $L^p(\mathbb{R}^d,\gamma_d;\mathbb{R}^d)$ are reflexive so that E is reflexive, and $T(W^{1,p}(\mathbb{R}^d,\gamma_d))$ is reflexive too. Being isometric to a reflexive space, $W^{1,p}(\mathbb{R}^d,\gamma_d)$ is reflexive.

As a consequence of reflexivity, if a sequence (f_n) is bounded in $W^{1,p}(\mathbb{R}^d,\gamma_d)$ a subsequence f_{n_k} converges weakly to an element g of $W^{1,p}(\mathbb{R}^d,\gamma_d)$ as $k\to\infty$. Since $f_{n_k}\to f$ in $L^p(\mathbb{R}^d,\gamma_d)$, $f=g$ and statement (v) is proved. □

Note that the argument of the proof of (ii) works as well for $p=1$, and statement (ii) is in fact true also for $p=1$. Even statement (i) holds for $p=1$, but the fact that $x\mapsto x_if(x)\in L^1(\mathbb{R}^d,\gamma_d)$ for every $f\in W^{1,1}(\mathbb{R}^d,\gamma_d)$ is not obvious, and will not be considered in these lectures.

Instead, $W^{1,1}(\mathbb{R}^d,\gamma_d)$ is not reflexive, and statement (v) does not hold for $p=1$ (see Exercise 9.2).

The next characterisation is useful to recognise whether a given function belongs to $W^{1,p}(\mathbb{R}^d,\gamma_d)$. We recall that $L^p_{loc}(\mathbb{R}^d)$ (resp. $W^{1,p}_{loc}(\mathbb{R}^d)$) is the space of all (equivalence classes of) functions f such that the restriction of f to any ball B belongs to $L^p(B,\lambda_d)$ (resp. $W^{1,p}(B,\lambda_d)$). Equivalently, $f\in L^p_{loc}(\mathbb{R}^d)$ (resp. $f\in W^{1,p}_{loc}(\mathbb{R}^d)$) if $f\theta\in L^p(\mathbb{R}^d,\lambda_d)$ (resp. $f\theta\in W^{1,p}(\mathbb{R}^d,\lambda_d)$) for every $\theta\in C^\infty_c(\mathbb{R}^d)$. For $f\in W^{1,p}_{loc}(\mathbb{R}^d)$ we denote by D_if the weak derivative of f with respect to x_i, $i=1,\dots d$.

Proposition 9.1.6. *For every $1\le p<\infty$,*

$$W^{1,p}(\mathbb{R}^d,\gamma_d)=\Big\{f\in W^{1,p}_{loc}(\mathbb{R}^d):\ f,\ D_if\in L^p(\mathbb{R}^d,\gamma_d),\ i=1,\dots d\Big\}.$$

Moreover, for every $f\in W^{1,p}(\mathbb{R}^d,\gamma_d)$ and $i=1,\dots,d$, $\partial f/\partial x_i$ coincides with the weak derivative D_if.

Proof. Let $f\in W^{1,p}(\mathbb{R}^d,\gamma_d)$. Then for every $g\in C^1_c(\mathbb{R}^d)$, (9.1.3) still holds: indeed, it is sufficient to approximate f by a sequence of functions belonging to $C^1_b(\mathbb{R}^d)$, to use (9.1.3) for every approximating function f_n, and to let $n\to\infty$.

This implies that $\partial f/\partial x_i$ is equal to the weak derivative D_if. Indeed, for every $\varphi\in C^\infty_c(\mathbb{R}^d)$, setting $g(x)=\varphi(x)e^{|x|^2/2}(2\pi)^{d/2}$, (9.1.3) yields

$$\int_{\mathbb{R}^d} f\frac{\partial\varphi}{\partial x_i}\,dx=\int_{\mathbb{R}^d} f\left(\frac{\partial g}{\partial x_i}-x_ig\right)d\gamma_d=-\int_{\mathbb{R}^d}\frac{\partial f}{\partial x_i}g\,d\gamma_d=-\int_{\mathbb{R}^d}\frac{\partial f}{\partial x_i}\varphi\,dx.$$

So, $\partial f/\partial x_i=D_if$, for every $i=1,\dots,d$. Since $L^p(\mathbb{R}^d,\gamma_d)\subset L^p_{loc}(\mathbb{R}^d)$, the inclusion $W^{1,p}(\mathbb{R}^d,\gamma_d)\subset\{f\in W^{1,p}_{loc}(\mathbb{R}^d):\ f,\ D_if\in L^p(\mathbb{R}^d,\gamma_d),\ i=1,\dots d\}$ is proved.

Conversely, let $f\in W^{1,p}_{loc}(\mathbb{R}^d)$ be such that $f,\ D_if\in L^p(\mathbb{R}^d,\gamma_d)$ for $i=1,\dots d$. Fix any function $\theta\in C^\infty_c(\mathbb{R}^d)$ such that $\theta\equiv 1$ in $B(0,1)$ and $\theta\equiv 0$ outside $B(0,2)$. For every $n\in\mathbb{N}$, we define

$$f_n(x):=\theta(x/n)f(x),\quad x\in\mathbb{R}^d.$$

Each f_n belongs to $W^{1,p}(\mathbb{R}^d,\gamma_d)$, because the restriction of f to $B(0,2n)$ may be approximated by a sequence (φ_k) of C^1 functions in $W^{1,p}(B(0,2n),\lambda_d)$, and the sequence (u_k) defined by $u_k(x)=\theta(x/n)\varphi_k(x)$ for $|x|\le 2n$, $u_k(x)=0$ for $|x|\ge 2n$ is contained in

$C_b^1(\mathbb{R}^d)$, it is a Cauchy sequence in the norm (9.1.5), and it converges to f_n in $L^p(\mathbb{R}^d, \gamma_d)$ since

$$\int_{\mathbb{R}^d} |u_k - f_n|^p d\gamma_d = \int_{B(0,2n)} |\theta(x/n)(f(x) - \varphi_k(x))|^p \gamma_d(dx) \le \frac{\|\theta\|_\infty}{(2\pi)^{d/2}} \int_{B(0,2n)} |f - \varphi_k|^p dx.$$

In its turn, the sequence (f_n) converges to f in $L^p(\mathbb{R}^d, \gamma_d)$, by the Dominated Convergence Theorem. Moreover, for every $i = 1, \ldots, d$ we have $\partial f_n/\partial x_i(x) = n^{-1} D_i\theta(x/n) f(x) + \theta(x/n) D_i f(x)$, so that $\partial f_n/\partial x_i$ converges to $D_i f$ in $L^p(\mathbb{R}^d, \gamma_d)$, still by the Dominated Convergence Theorem. Therefore, $f \in W^{1,p}(\mathbb{R}^d, \gamma_d)$. □

By Proposition 9.1.6, if a C^1 function f is such that f, $D_i f$ belong to $L^p(\mathbb{R}^d, \gamma_d)$ for every $i = 1, \ldots, d$, then $f \in W^{1,p}(\mathbb{R}^d, \gamma_d)$. In particular, all polynomials belong to $W^{1,p}(\mathbb{R}^d, \gamma_d)$, for every $1 \le p < \infty$.

9.2 The Bochner integral

We only need the first notions of the theory of integration for Banach space valued functions. We refer to the books [8], [29, Ch. V] for a detailed treatment.

Let $(\Omega, \mathscr{F})$ be a measurable space and let $\mu : \mathscr{F} \to [0, \infty)$ be a positive finite measure. We shall define integrals and L^p spaces of Y-valued functions, where Y is any separable real Banach space, with norm $\|\cdot\|_Y$.

In the following sections, Ω will be a Banach space X endowed with a Gaussian measure, and Y will be either X or the Cameron–Martin space H. However, the definitions and the basic properties are the same for a Gaussian measure and for a general positive finite measure.

As in the scalar valued case, the *simple functions* are functions of the type

$$F(x) = \sum_{i=1}^n 1\!\!1_{\Gamma_i}(x) y_i, \quad x \in \Omega,$$

with $n \in \mathbb{N}$, $\Gamma_i \in \mathscr{F}$, $y_i \in Y$ for every $i = 1, \ldots, n$ and $\Gamma_i \cap \Gamma_j = \emptyset$ for $i \ne j$. In this case, the integral of F is defined by

$$\int_\Omega F(x)\mu(dx) := \sum_{i=1}^n \mu(\Gamma_i) y_i. \tag{9.2.1}$$

It is easily seen that the integral is linear, namely for every α, $\beta \in \mathbb{R}$ and for every couple of simple functions F_1, F_2

$$\int_\Omega (\alpha F_1(x) + \beta F_2(x))\mu(dx) = \alpha \int_\Omega F_1(x)\mu(dx) + \beta \int_X F_2(x)\mu(dx) \tag{9.2.2}$$

and it satisfies

$$\left\| \int_\Omega F(x)\gamma(dx) \right\|_Y \le \int_\Omega \|F(x)\|_Y \gamma(dx), \tag{9.2.3}$$

for every simple function F (notice that $x \mapsto \|F(x)\|_Y$ is a simple real valued function).

Definition 9.2.1. *A function $F:\Omega\to Y$ is called strongly measurable if there exists a sequence of simple functions (F_n) such that $\lim_{n\to\infty}\|F(x)-F_n(x)\|_Y=0$, for μ-a.e. $x\in\Omega$.*

Notice that if Y is separable then this notion coincides with the general notion of measurable function given in Definition 1.1.7, see [27, Proposition I.1.9]. If $Y=\mathbb{R}$, see Exercise 9.3. Also, notice that if F is strongly measurable, then $\|F(\cdot)\|_Y$ is a real valued measurable function. The following theorem is a consequence of an important result is due to Pettis (e.g. [8, Thm. II.2]).

Theorem 9.2.2. *A function $F:\Omega\to Y$ is strongly measurable if and only if for every $f\in Y^*$ the composition $f\circ F:\Omega\to\mathbb{R}$, $x\mapsto f(F(x))$, is measurable.*

As a consequence, if Y is a separable Hilbert space and $\{y_k:\ k\in\mathbb{N}\}$ is an orthonormal basis of Y, then $F:\Omega\to Y$ is strongly measurable if and only if the real valued functions $x\mapsto\langle F(x),y_k\rangle_Y$ are measurable.

Definition 9.2.3. *A strongly measurable function $F:\Omega\to Y$ is called Bochner integrable if there exists a sequence of simple functions (F_n) such that*

$$\lim_{n\to\infty}\int_\Omega\|F(x)-F_n(x)\|_Y\mu(dx)=0.$$

In this case, the sequence $\int_\Omega F_n d\mu$ is a Cauchy sequence in Y by estimate (9.2.3), and we define

$$\int_\Omega F(x)\,\mu(dx):=\lim_{n\to\infty}\int_\Omega F_n(x)\,\mu(dx)$$

(of course, the above limit is independent of the defining sequence (F_n)). The following result is known as the Bochner Theorem.

Proposition 9.2.4. *A measurable function $F:\Omega\to Y$ is Bochner integrable if and only if*

$$\int_\Omega\|F(x)\|_Y\mu(dx)<\infty.$$

Proof. If F is integrable, for every sequence of simple functions (F_n) in Definition 9.2.3 we have

$$\int_\Omega\|F(x)\|_Y\mu(dx)\le\int_\Omega\|F(x)-F_n(x)\|_Y\mu(dx)+\int_\Omega\|F_n(x)\|_Y\mu(dx),$$

which is finite for n large enough.

To prove the converse, if $\int_\Omega\|F(x)\|_Y\mu(dx)<\infty$ we construct a sequence of simple functions (F_n) that converge pointwise to F and such that $\lim_{n\to\infty}\int_\Omega\|F(x)-F_n(x)\|_Y\mu(dx)=0$.

Let $\{y_k:\ k\in\mathbb{N}\}$ be a dense subset of Y. Set

$$\begin{aligned}\theta_n(x)&:=\min\{\|F(x)-y_k\|_Y:\ k=1,\dots,n\},\\ k_n(x)&:=\min\{k\le n:\ \theta_n(x)=\|F(x)-y_k\|_Y\},\end{aligned}$$

and

$$F_n(x) := y_{k_n(x)}, \quad x \in X.$$

Then every θ_n is a real valued measurable function. This implies that F_n is a simple function, because it takes the values $y_1, \dots y_n$, and for every $k = 1, \dots, n$, $F_n^{-1}(y_k)$ is the measurable set $\Gamma_k = \{x \in \Omega : \ \theta_n(x) = \|F(x) - y_k\|_Y\}$.

For every x the sequence $\|F_n(x) - F(x)\|_Y$ decreases monotonically to 0 as $n \to \infty$. Moreover, for every $n \in \mathbb{N}$,

$$\|F_n(x) - F(x)\|_Y \le \|y_1 - F(x)\|_Y \le \|y_1\|_Y + \|F(x)\|_Y, \quad x \in X. \tag{9.2.4}$$

By the Dominated Convergence Theorem (recall that μ is a finite measure) or else, by the Monotone Convergence Theorem,

$$\lim_{n\to\infty} \int_\Omega \|F_n(x) - F(x)\|_Y \mu(dx) = 0.$$

$\square$

If $F : \Omega \to Y$ is integrable, for every $E \in \mathscr{F}$ the function $1\!\!1_E F$ is integrable, and we set

$$\int_E F(x)\mu(dx) = \int_\Omega 1\!\!1_E(x) F(x)\, \mu(dx).$$

The Bochner integral is linear with respect to F, namely for every α, $\beta \in \mathbb{R}$ and for every integrable F_1, F_2, (9.2.2) holds. Moreover, it enjoys the following properties.

Proposition 9.2.5. *Let $F : \Omega \to Y$ be a Bochner integrable function. Then*

(i) $\| \int_\Omega F(x)\, \mu(dx)\|_Y \le \int_\Omega \|F(x)\|_Y \mu(dx)$*;*

(ii) $\lim_{\mu(E)\to 0} \int_E F(x)\, \mu(dx) = 0$*;*

(iii) If (E_n) is a sequence of pairwise disjoint measurable sets in Ω and $E = \cup_{n\in\mathbb{N}} E_n$, then

$$\int_E F(x)\, \mu(dx) = \sum_{n\in\mathbb{N}} \int_{E_n} F(x)\, \mu(dx);$$

(iv) For every $f \in Y^*$, the real valued function $x \mapsto f(F(x))$ is in $L^1(\Omega, \mu)$, and

$$f\bigg(\int_\Omega F(x)\, \mu(dx) \bigg) = \int_\Omega f(F(x))\, \mu(dx). \tag{9.2.5}$$

Proof. (i) Let (F_n) be a sequence of simple functions as in Definition 9.2.3. By (9.2.3) for every $n \in \mathbb{N}$ we have $\| \int_\Omega F_n(x)\, \mu(dx)\|_Y \le \int_\Omega \|F_n(x)\|_Y \mu(dx)$. Then,

$$\begin{aligned}
\bigg\| \int_\Omega F(x)\, \mu(dx) \bigg\|_Y &= \bigg\| \lim_{n\to\infty} \int_\Omega F_n(x)\, \mu(dx) \bigg\|_Y \le \limsup_{n\to\infty} \int_\Omega \|F_n(x)\|_Y\, \mu(dx) \\
&\le \lim_{n\to\infty} \int_\Omega \|F_n(x) - F(x)\|_Y\, \mu(dx) + \int_\Omega \|F(x)\|_Y\, \mu(dx) \\
&= \int_\Omega \|F(x)\|_Y\, \mu(dx).
\end{aligned}$$

Statement (ii) means: for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\mu(E) \le \delta$ implies $\| \int_E F(x)\,\mu(dx)\|_Y \le \varepsilon$. Since $\lim_{\mu(E)\to 0} \int_E \|F(x)\|_Y\,\mu(dx) = 0$, statement (ii) is a consequence of (i).

Let us prove statement (iii). Since, for every n,

$$\left\| \int_{E_n} F(x)\,\mu(dx) \right\|_Y \le \int_{E_n} \|F(x)\|_Y \mu(dx),$$

the series $\sum_{n\in\mathbb{N}} \int_{E_n} F(x)\,\mu(dx)$ converges in Y, and its norm does not exceed

$$\int_\Omega \|F(x)\|_Y \mu(dx).$$

Since the Bochner integral is finitely additive,

$$\left\| \int_E F(x)\,\mu(dx) - \sum_{n=1}^m \int_{E_n} F(x)\,\mu(dx) \right\|_Y = \left\| \int_{\cup_{n=m+1}^\infty E_n} F(x)\,\mu(dx) \right\|_Y$$

where $\lim_{m\to\infty} \mu(\cup_{n=m+1}^\infty E_n) = 0$. By statement (ii), the right-hand side vanishes as $m \to \infty$, and statement (iii) follows.

Let us prove statement (iv). Note that (9.2.5) holds obviously for simple functions. Let (F_n) be the sequence of functions in the proof of Proposition 9.2.4. Then,

$$\begin{aligned} f\Big(\int_\Omega F(x)\,\mu(dx) \Big) &= f\Big(\lim_{n\to\infty} \int_\Omega F_n(x)\,\mu(dx) \Big) \\ &= \lim_{n\to\infty} f\Big(\int_\Omega F_n(x)\,\mu(dx) \Big) = \lim_{n\to\infty} \int_\Omega f(F_n(x))\,\mu(dx). \end{aligned}$$

On the other hand, the sequence $(f(F_n(x)))$ converges pointwise to $f(F(x))$, and by (9.2.4)

$$\begin{aligned} |f(F_n(x))| &\le \|f\|_{Y^*} \|F_n(x)\|_Y \le \|f\|_{Y^*} (\|F_n(x) - F(x)\|_Y + \|F(x)\|_Y) \\ &\le \|f\|_{Y^*} (\|y_1\|_Y + 2\|F(x)\|_Y). \end{aligned}$$

By the Dominated Convergence Theorem,

$$\lim_{n\to\infty} \int_\Omega f(F_n(x))\,\mu(dx) = \int_\Omega f(F(x))\,\mu(dx),$$

and the statement follows. □

Remark 9.2.6. As a consequence of (iv), if Y is a separable Hilbert space and $\{y_k :\ k \in \mathbb{N}\}$ is an orthonormal basis of Y, for every Bochner integrable $F : \Omega \to Y$ the real valued functions $x \mapsto \langle F(x), y_k \rangle_Y$ belong to $L^1(\Omega, \mu)$, and we have

$$\int_\Omega F(x)\mu(dx) = \sum_{k=1}^\infty \int_\Omega \langle F(x), y_k \rangle_Y \mu(dx)\, y_k.$$

The L^p spaces of Y-valued functions are defined as in the scalar case. Namely, for every $1 \le p < \infty$, $L^p(\Omega, \mu; Y)$ is the space of the (equivalence classes) of Bochner integrable functions $F : \Omega \to Y$ such that

$$\|F\|_{L^p(\Omega,\mu;Y)} := \left(\int_X \|F(x)\|_Y^p \mu(dx) \right)^{1/p} < \infty.$$

The proof that $L^p(\Omega, \mu; Y)$ is a Banach space with the above norm is the same as in the real valued case. If $p = 2$ and Y is a Hilbert space, $L^p(\Omega, \mu; Y)$ is a Hilbert space with the scalar product

$$\langle F, G\rangle_{L^2(\Omega,\mu;Y)} := \int_\Omega \langle F(x), G(x)\rangle_Y \mu(dx).$$

As usual, we define

$$L^\infty(\Omega, \mu; Y) := \Big\{ F : \Omega \to Y \text{ measurable s.t. } \|F\|_{L^\infty(\Omega,\mu;Y)} < \infty \Big\},$$

where

$$\|F\|_{L^\infty(\Omega,\mu;Y)} := \inf\Big\{ M > 0 : \mu(\{x : \|F(x)\|_Y > M\}) = 0 \Big\}.$$

Notice that if Y is a separable Hilbert space, which is our setting, the space $L^p(\Omega, \mu; Y)$ is reflexive for $1 < p < \infty$, see [8, Section IV.1].

The first example of Bochner integral that we met in these lectures was the mean a of a Gaussian measure γ on a separable Banach space X. By Proposition 2.3.3, there exists a unique $a \in X$ such that $a_\gamma(f) = f(a)$, for every $f \in X^*$. Since γ is a Borel measure, every continuous $F : X \to X$ is measurable; in particular $F(x) := x$ is measurable, hence strongly measurable. By the Fernique Theorem and Proposition 9.2.4 it belongs to $L^p(X, \gamma; X)$ for every $1 \le p < \infty$, and we have

$$a = \int_X x\, \gamma(dx).$$

Indeed, for every $f \in X^*$, we have

$$f\left(\int_X x\, \gamma(dx) \right) = \int_X f(x)\, \gamma(dx) = a_\gamma(f),$$

by (9.2.5). Therefore, $a = \int_X x\, \gamma(dx)$.

9.3 The infinite dimensional case

9.3.1 Differentiable functions

Definition 9.3.1. *Let X, Y be normed spaces. Let $\overline{x} \in X$ and let Ω be a neighbourhood of $\overline{x}$. A function $f : \Omega \to Y$ is called (Fréchet) differentiable at $\overline{x}$ if there exists $\ell \in \mathcal{L}(X, Y)$ such that*

$$\|f(\overline{x} + h) - f(\overline{x}) - \ell(h)\|_Y = o(\|h\|_X) \quad \text{as } h \to 0 \text{ in } X.$$

In this case, ℓ is unique, and we set $f'(\overline{x}) := \ell$.

Several properties of differentiable functions may be proved as in the case $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$. First, if f is differentiable at $\overline{x}$ it is continuous at $\overline{x}$. Moreover, for every $v \in X$ the directional derivative

$$\frac{\partial f}{\partial v}(\overline{x}) := Y - \lim_{t\to 0} \frac{f(\overline{x}+tv) - f(\overline{x})}{t}$$

exists and is equal to $f'(\overline{x})(v)$.

If $Y = \mathbb{R}$ and $f : X \to \mathbb{R}$ is differentiable at $\overline{x}$, $f'(\overline{x})$ is an element of X^*. In particular, if $f \in X^*$ then f is differentiable at every $\overline{x}$ and f' is constant, with $f'(\overline{x})(v) = f(v)$ for every $\overline{x}$, $v \in X$. If $f \in \mathcal{F}C^1_b(X)$, $f(x) = \varphi(\ell_1(x), \dots, \ell_n(x))$ with $\ell_k \in X^*$, $\varphi \in C^1_b(\mathbb{R}^n)$, f is differentiable at every $\overline{x}$ and

$$f'(\overline{x})(v) = \sum_{k=1}^{n} \frac{\partial \varphi}{\partial \xi_k}((\ell_1(\overline{x}), \dots, \ell_n(\overline{x}))\ell_k(v), \quad \overline{x},\ v \in X.$$

If f is differentiable at x for every x in a neighbourhood of $\overline{x}$, it may happen that the function $X \to \mathcal{L}(X, Y)$, $x \mapsto f'(x)$ is differentiable at $\overline{x}$, too. In this case, the derivative is denoted by $f''(\overline{x})$, and it is an element of $\mathcal{L}(X, \mathcal{L}(X, Y))$. The higher order derivatives are defined recursively, in the same way.

If $f : X \to \mathbb{R}$ is twice differentiable at $\overline{x}$, $f''(\overline{x})$ is an element of $\mathcal{L}(X, X^*)$, which is canonically identified with the space of the continuous bilinear forms $\mathcal{L}^{(2)}(X)$: indeed, if $v \in \mathcal{L}(X, X^*)$, the function $X^2 \to \mathbb{R}$, $(x, y) \mapsto v(x)(y)$, is linear both with respect to x and with respect to y and it is continuous, so that it is a bilinear form; conversely every bilinear continuous form $a : X^2 \to \mathbb{R}$ gives rise to the element $v \in \mathcal{L}(X, X^*)$ defined by $v(x)(y) = a(x, y)$. Moreover,

$$\|v\|_{\mathcal{L}(X,X^*)} = \sup_{x\neq 0,\, y\neq 0} \frac{|v(x)(y)|}{\|x\|_X\, \|y\|_X} = \sup_{x\neq 0,\, y\neq 0} \frac{|a(x,y)|}{\|x\|_X\, \|y\|_X} = \|a\|_{\mathcal{L}^{(2)}(X)}.$$

Similarly, if $f : X \to \mathbb{R}$ is k times differentiable at $\overline{x}$, $f^{(k)}(\overline{x})$ is identified with an element of the space $\mathcal{L}^{(k)}(X)$ of the continuous k-linear forms.

Definition 9.3.2. *Let $k \in \mathbb{N}$. We denote by $C^k_b(X)$ the set of bounded and k times continuously differentiable functions $f : X \to \mathbb{R}$, with bounded $\|f^{(j)}\|_{\mathcal{L}^{(j)}(X)}$ for every $j = 1, \dots, k$. It is normed by*

$$\|f\|_{C^k_b(X)} = \sum_{j=0}^{k} \sup_{x\in X} \|f^{(j)}(x)\|_{\mathcal{L}^{(j)}(X)},$$

where we set $f^{(0)}(x) = f(x)$. Moreover we set

$$C^\infty_b(X) = \bigcap_{k\in\mathbb{N}} C^k_b(X).$$

If X is a Hilbert space and $f : X \to \mathbb{R}$ is differentiable at $\overline{x}$, there exists a unique $y \in X$ such that $f'(\overline{x})(x) = \langle x, y\rangle$, for every $x \in X$. We set

$$\nabla f(\overline{x}) := y.$$

From now on, X is a separable Banach space endowed with a norm $\|\cdot\|$ and with a Gaussian centred non degenerate measure γ, and H is its Cameron-Martin space defined in Lecture 3.

Definition 9.3.3. *A function $f : X \to \mathbb{R}$ is called H-differentiable at $\overline{x} \in X$ if there exists $\ell_0 \in H^*$ such that*

$$|f(\overline{x} + h) - f(\overline{x}) - \ell_0(h)| = o(|h|_H) \quad \textit{as } h \to 0 \textit{ in } H.$$

If f is H-differentiable at $\overline{x}$, the operator ℓ_0 in the definition is called H-derivative of f at $\overline{x}$, and there exists a unique $y \in H$ such that $\ell_0(h) = [h, y]_H$ for every $h \in H$. We set

$$\nabla_H f(\overline{x}) := y.$$

Definition 9.3.3 differs from 9.3.1 in that the increments are taken only in H.

Lemma 9.3.4. *If f is differentiable at $\overline{x}$, then it is H-differentiable at $\overline{x}$, with H-derivative given by $h \mapsto f'(\overline{x})(h)$ for every $h \in H$. Moreover, we have*

$$\nabla_H f(\overline{x}) = R_\gamma f'(\overline{x}). \tag{9.3.1}$$

Proof. Setting $\ell = f'(\overline{x})$ we have

$$\lim_{|h|_H \to 0} \frac{|f(\overline{x} + h) - f(\overline{x}) - \ell(h)|}{|h|_H} = \lim_{|h|_H \to 0} \frac{|f(\overline{x} + h) - f(\overline{x}) - \ell(h)|}{\|h\|} \frac{\|h\|}{|h|_H} = 0,$$

because H is continuously embedded in X so that the ratio $\|h\|/|h|_H$ is bounded by a constant independent of h. This proves the first assertion. To prove (9.3.1), we recall that for every $\varphi \in X^*$ we have $\varphi(h) = [R_\gamma \varphi, h]_H$ for each $h \in H$; in particular, taking $\varphi = f'(\overline{x})$ we obtain $f'(\overline{x})(h) = [R_\gamma f'(\overline{x}), h]_H = [\nabla_H f(\overline{x}), h]_H$ for each $h \in H$, and therefore $\nabla_H f(\overline{x}) = R_\gamma f'(\overline{x})$. $\square$

If f is just H-differentiable at $\overline{x}$, the directional derivative $\frac{\partial f}{\partial v}(\overline{x})$ exists for every $v \in H$, and it is given by $[\nabla_H f(\overline{x}), v]_H$. Fixed any orthonormal basis $\{h_n :\ n \in \mathbb{N}\}$ of H, we set

$$\partial_i f(\overline{x}) := \frac{\partial f}{\partial h_i}(\overline{x}), \quad i \in \mathbb{N}.$$

So, we have

$$\nabla_H f(\overline{x}) = \sum_{i=1}^{\infty} \partial_i f(\overline{x}) h_i, \tag{9.3.2}$$

where the series converges in H.

We warn the reader that if X is a Hilbert space and f is differentiable at $\overline{x}$, the gradient and the H-gradient of f at $\overline{x}$ do not coincide in general. If $\gamma = \mathscr{N}(0, Q)$, identifying X^* with X as usual, Lemma 9.3.4 implies that $\nabla_H f(\overline{x}) = Q\nabla f(\overline{x})$.

We recall that if γ is non degenerate, then Q is positive definite. Fixed any orthonormal basis $\{e_j : \ j \in \mathbb{N}\}$ of X consisting of eigenvectors of Q, $Qe_j = \lambda_j e_j$, then a canonical orthonormal basis of H is $\{h_j : \ j \in \mathbb{N}\}$, with $h_j = \sqrt{\lambda_j} e_j$, and we have

$$\partial_j f(\overline{x}) = \sqrt{\lambda_j}\frac{\partial f}{\partial e_j}(\overline{x}), \quad j \in \mathbb{N}.$$

9.3.2 Sobolev spaces of order 1

As in finite dimension, the starting point to define the Sobolev spaces is an integration formula for C^1_b functions.

Proposition 9.3.5. *For every $f \in C^1_b(X)$ and $h \in H$ we have*

$$\int_X \frac{\partial f}{\partial h}\, d\gamma = \int_X f\, \hat{h}\, d\gamma. \tag{9.3.3}$$

Consequently, for every f, $g \in C^1_b(X)$ and $h \in H$ we have

$$\int_X \frac{\partial f}{\partial h} g\, d\gamma = -\int_X \frac{\partial g}{\partial h} f\, d\gamma + \int_X f\, g\, \hat{h}\, d\gamma. \tag{9.3.4}$$

Proof. By the Cameron–Martin Theorem 3.1.5, for every $t \in \mathbb{R}$ we have

$$\int_X f(x + th)\, \gamma(dx) = \int_X f(x) e^{t\hat{h}(x) - t^2|h|^2_H/2} \gamma(dx),$$

so that, for $0 < |t| \le 1$,

$$\int_X \frac{f(x + th) - f(x)}{t}\, \gamma(dx) = \int_X f(x)\, \frac{e^{t\hat{h}(x) - t^2|h|^2_H/2} - 1}{t}\, \gamma(dx).$$

As $t \to 0$, the integral in the left-hand side converges to $\int_X \partial f/\partial h\, d\gamma$, by the Dominated Convergence Theorem. Concerning the right-hand side, $(e^{t\hat{h}(x) - t^2|h|^2_H/2} - 1)/t \to \hat{h}(x)$ for every $x \in X$. For $|t| \le 1$, we estimate

$$\begin{aligned}\left|\frac{e^{t\hat{h}(x) - t^2|h|^2_H/2} - 1}{t}\right| &= \left|\frac{e^{-t^2|h|^2_H/2}(e^{t\hat{h}(x)} - 1)}{t} + \frac{e^{-t^2|h|^2_H/2} - 1}{t}\right| \\ &\le |\hat{h}(x)|e^{|\hat{h}(x)|}| + \sup_{0<t\le 1}\left|\frac{e^{-t^2|h|^2_H/2} - 1}{t}\right|,\end{aligned}$$

where the function $x \mapsto |\hat{h}(x)|e^{|\hat{h}(x)|}$ belongs to $L^1(X, \gamma)$ since $\hat{h}$ is a Gaussian random variable. So, applying the Dominated Convergence Theorem we get the statement. $\square$

Notice that formula (9.3.3) is a natural extension of (9.1.2) to the infinite dimensional case. In $(\mathbb{R}^d, \gamma_d)$ the equality $H = \mathbb{R}^d$ holds, and for every $h \in \mathbb{R}^d$ we have $\hat{h}(x) = h \cdot x = [h, x]_H$.

We proceed as in finite dimension to define the Sobolev spaces of order 1. Next step is to prove that some gradient operator, defined on a set of good enough functions, is closable in $L^p(X, \gamma)$. In our general setting the only available gradient is ∇_H. We shall use the following lemma, whose proof is left as an exercise, being a consequence of the results of Lecture 7.

Lemma 9.3.6. *Let $\psi \in L^1(X, \gamma)$ be such that*

$$\int_X \psi\, \varphi\, d\gamma = 0, \quad \varphi \in \mathcal{F}C_b^1(X).$$

Then $\psi = 0$ a.e.

Proposition 9.3.7. *For every $1 \leq p < \infty$, the operator $\nabla_H : D(\nabla_H) = \mathcal{F}C_b^1(X) \to L^p(X, \gamma; H)$ is closable as an operator from $L^p(X, \gamma)$ to $L^p(X, \gamma; H)$.*

Proof. Let $1 < p < \infty$. Let $f_n \in \mathcal{F}C_b^1(X)$ be such that $f_n \to 0$ in $L^p(X, \gamma)$ and $\nabla_H f_n \to G$ in $L^p(X, \gamma; H)$. For every $h \in H$ and $\varphi \in \mathcal{F}C_b^1(X)$ we have

$$\lim_{n\to\infty} \int_X \frac{\partial f_n}{\partial h} \varphi\, d\gamma = \int_X [G(x), h]_H \varphi(x)\, \gamma(dx),$$

since

$$\int_X |(\partial f_n/\partial h - [G(x), h]_H)\varphi|\, d\gamma \leq |h|_H^p \bigg(\int_X |\nabla_H f_n - G|_H^p d\gamma \bigg)^{1/p} \|\varphi\|_{L^{p'}(X,\gamma)}.$$

On the other hand,

$$\int_X \frac{\partial f_n}{\partial h} \varphi\, d\gamma = -\int_X f_n \frac{\partial \varphi}{\partial h}\, d\gamma + \int_X f_n \varphi \hat{h}\, d\gamma, \quad n \in \mathbb{N},$$

so that, since $f_n \to 0$ in $L^p(X, \gamma)$ and $\partial\varphi/\partial h$, $\hat{h}\varphi \in L^{p'}(X, \gamma)$,

$$\lim_{n\to\infty} \int_X \frac{\partial f_n}{\partial h} \varphi\, d\gamma = 0.$$

So,

$$\int_X [G(x), h]_H \varphi(x)\, \gamma(dx) = 0, \quad \varphi \in \mathcal{F}C_b^1(X), \tag{9.3.5}$$

and by Lemma 9.3.6, $[G(x), h]_H = 0$ a.e. Fix any orthonormal basis $\{h_k :\ k \in \mathbb{N}\}$ of H. Then $\bigcup_{k\in\mathbb{N}} \{x : [G(x), h_k]_H \neq 0\}$ is negligible so that $G(x) = 0$ a.e.

Let now $p = 1$. The above procedure does not work, since $\hat{h}\varphi \notin L^\infty(X, \gamma)$ in general, although it belongs to $L^q(X, \gamma)$ for every $q > 1$. Let (f_n) be a sequence in $\mathcal{F}C_b^1(X)$, $f_n \to 0$ in $L^1(X, \gamma)$, $\nabla_H f_n \to G$ in $L^1(X, \gamma; H)$. We want to show that $G = 0$.

Without loss of generality we may assume that $f_n \to 0$, $\nabla_H f_n \to G$ a.e., and that there exists $g \in L^1(X,\gamma)$ such that $|\nabla_H f_n|_H \le g$ a.e. for any $n \in \mathbb{N}$ (just take a subsequence which we do not relabel such that $\sum_n \|f_n\|_{L^1(X,\gamma)} < \infty$ and set $g = \sum_n |f_n|$).

Let $\theta \in C_b^1(\mathbb{R})$ be such that $\theta(0) = 0$, $\theta'(0) = 1$. Then $\theta \circ f_n \to \theta(0) = 0$ a.e., and therefore in $L^p(X,\gamma)$ for all $1 \le p < \infty$, because θ is bounded. Also

$$\nabla_H(\theta \circ f_n) = (\theta' \circ f_n)\nabla_H f_n \to \theta'(0)G = G$$

a.e. and therefore in $L^1(X,\gamma;H)$ by the Dominated Convergence Theorem. These convergences imply that the proof of $G = 0$ can be carried out in the same way as for $1 < p < \infty$. Indeed, for any $h \in H$, using the integration by parts formula (9.3.4), for every n we have

$$\int_X \frac{\partial(\theta \circ f_n)}{\partial h}\varphi\, d\gamma = \int_X (\theta \circ f_n)[\hat{h}\varphi - \frac{\partial \varphi}{\partial h}]\, d\gamma.$$

Letting $n \to \infty$, we obtain

$$\int_X [G(x), h]_H \varphi\, d\gamma = 0, \quad \varphi \in \mathcal{F}C_b^1(X).$$

□

The proof of Proposition 9.3.7 for $p = 1$ is more complicated than the proof of Lemma 9.1.3, where we could use compactly supported functions φ.

Remark 9.3.8. Note that in the proof of Proposition 9.3.7 we proved that for every $h \in H$ the linear operator $\partial_h : D(\partial_h) = \mathcal{F}C_b^1(X) \to L^p(X,\gamma)$ is closable as an operator from $L^p(X,\gamma)$ to $L^p(X,\gamma;H)$.

We are now ready to define the Sobolev spaces of order 1 and the generalized H-gradients.

Definition 9.3.9. *For every $1 \le p < \infty$, $W^{1,p}(X,\gamma)$ is the domain of the closure of $\nabla_H : \mathcal{F}C_b^1(X) \to L^p(X,\gamma;H)$ in $L^p(X,\gamma)$ (still denoted by ∇_H). Therefore, an element $f \in L^p(X,\gamma)$ belongs to $W^{1,p}(X,\gamma)$ iff there exists a sequence of functions $f_n \in \mathcal{F}C_b^1(X)$ such that $f_n \to f$ in $L^p(X,\gamma)$ and $\nabla_H f_n$ converges in $L^p(X,\gamma;H)$, and in this case, $\nabla_H f = \lim_{n\to\infty} \nabla_H f_n$.*

$W^{1,p}(X,\gamma)$ is a Banach space with the graph norm

$$\|f\|_{W^{1,p}} := \|f\|_{L^p(X,\gamma)} + \|\nabla_H f\|_{L^p(X,\gamma;H)} = \left(\int_X |f|^p d\gamma\right)^{1/p} + \left(\int_X |\nabla_H f|_H^p d\gamma\right)^{1/p}. \tag{9.3.6}$$

For $p = 2$, $W^{1,2}(X,\gamma)$ is a Hilbert space with the natural inner product

$$\langle f, g\rangle_{W^{1,2}} := \int_X f\, g\, d\gamma + \int_X [\nabla_H f, \nabla_H g]_H d\gamma,$$

which gives an equivalent norm.

For every fixed orthonormal basis $\{h_j :\ j \in \mathbb{N}\}$ of H, and for every $f \in W^{1,p}(X,\gamma)$, we set

$$\partial_j f(x) := [\nabla_H f(x), h_j]_H, \quad j \in \mathbb{N}.$$

More generally, for every $h \in H$ we set

$$\partial_h f(x) := [\nabla_H f(x), h]_H.$$

By definition,

$$\int_X |\nabla_H f|_H^p d\gamma = \int_X \bigg(\sum_{j=1}^{\infty} [\nabla_H f, h_j]_H^2\bigg)^{p/2} d\gamma = \int_X \bigg(\sum_{j=1}^{\infty} (\partial_j f)^2\bigg)^{p/2} d\gamma.$$

Moreover, if $f_n \in \mathcal{F}C_b^1(X)$ is such that $f_n \to f$ in $L^p(X,\gamma)$ and $\nabla_H f_n$ converges in $L^p(X,\gamma;H)$, then

$$\lim_{n\to\infty} [\nabla_H f_n, h_j]_H = \lim_{n\to\infty} \partial_j f_n = \partial_j f, \quad \text{ in } \quad L^p(X,\gamma).$$

As in finite dimension, several properties of the spaces $W^{1,p}(X,\gamma)$ follow easily.

Proposition 9.3.10. *Let $1 < p < \infty$. Then*

(i) the integration formula (9.3.3) *holds for every $f \in W^{1,p}(X,\gamma)$, $h \in H$;*

(ii) if $\theta \in C_b^1(X;\mathbb{R})$ and $f \in W^{1,p}(X,\gamma)$, then $\theta \circ f \in W^{1,p}(X,\gamma)$, and $\nabla_H(\theta \circ f) = (\theta' \circ f)\nabla_H f$;

(iii) if $f \in W^{1,p}(X,\gamma)$, $g \in W^{1,q}(X,\gamma)$ with $1/p + 1/q = 1/s \leq 1$, then $fg \in W^{1,s}(X,\gamma)$ and

$$\nabla_H(fg) = \nabla_H f\, g + f \nabla_H g;$$

(iv) $W^{1,p}(X,\gamma)$ is reflexive;

(v) if $f_n \to f$ in $L^p(X,\gamma)$ and $\sup_{n\in\mathbb{N}} \|f_n\|_{W^{1,p}(X,\gamma)} < \infty$, then $f \in W^{1,p}(X,\gamma)$.

Proof. The proof is just a rephrasing of the proof of Proposition 9.1.5.

Statement (i) follows approximating f by a sequence of functions belonging to $\mathcal{F}C_b^1(X)$, using (9.3.3) for every approximating function f_n and letting $n \to \infty$.

Statement (ii) follows approximating $\theta \circ f$ by $\theta \circ f_n$, if $(f_n) \subset \mathcal{F}C_b^1(X)$ is such that $f_n \to f$ in $L^p(X,\gamma)$ and $\nabla_H f_n \to \nabla_H f$ in $L^p(X,\gamma;H)$.

Statement (iii) follows from the definition, approximating fg by $f_n g_n$ if (f_n), $(g_n) \subset \mathcal{F}C_b^1(X)$, are such that $f_n \to f$ in $L^p(X,\gamma)$, $\nabla_H f_n \to \nabla_H f$ in $L^p(X,\gamma;H)$, $g_n \to g$ in $L^{p'}(X,\gamma)$, $\nabla_H g_n \to \nabla_H g$ in $L^{p'}(X,\gamma;H)$. Then $\lim_{n\to\infty} f_n g_n = fg$ in $L^s(X,\gamma)$, and the sequence $(\nabla_H(f_n g_n))$ converges to $g\nabla_H f + f\nabla_H g$ in $L^s(X,\gamma;H)$.

Let us prove (iv). The mapping $u \mapsto Tu = (u, \nabla_H u)$ is an isometry from $W^{1,p}(X,\gamma)$ to the product space $E := L^p(X,\gamma) \times L^p(X,\gamma;H)$, which implies that the range of T is closed in E. Now, $L^p(X,\gamma)$ and $L^p(X,\gamma;H)$ are reflexive (e.g. [8, Ch. IV]) so that

E is reflexive, and $T(W^{1,p}(X,\gamma))$ is reflexive too. Being isometric to a reflexive space, $W^{1,p}(X,\gamma)$ is reflexive.

As a consequence of reflexivity, if a sequence (f_n) is bounded in $W^{1,p}(X,\gamma)$ a subsequence f_{n_k} converges weakly to an element g of $W^{1,p}(X,\gamma)$ as $k\to\infty$. Since $f_{n_k}\to f$ in $L^p(X,\gamma)$, $f=g$ and statement (v) is proved. □

As in finite dimension, statement (ii) holds as well for $p=1$.

Remark 9.3.11. Let X be a Hilbert space and let $\gamma=\mathscr{N}(0,Q)$ with $Q>0$. For every $f\in\mathcal{F}C^1_b(X)$ we have $\nabla_H f(x)=Q\nabla f(x)$, so that

$$|\nabla_H f(x)|_H^2=\langle Q^{-1/2}Q\nabla f(x),Q^{-1/2}Q\nabla f(x)\rangle=\|Q^{1/2}\nabla f(x)\|^2,$$

and

$$\|f\|_{W^{1,p}(X,\gamma)}=\|f\|_{L^p(X,\gamma)}+\left(\int_X\|Q^{1/2}\nabla f(x)\|^p d\gamma\right)^{1/p}.$$

Fixed any orthonormal basis $\{e_j:\ j\in\mathbb{N}\}$ of X consisting of eigenvectors of Q, $Qe_j=\lambda_j e_j$, then a canonical basis of H is $\{h_j:\ j\in\mathbb{N}\}$, with $h_j=\sqrt{\lambda_j}e_j$, $\partial_j f(x)=\sqrt{\lambda_j}\partial f/\partial e_j$, and

$$\|f\|_{W^{1,p}(X,\gamma)}=\|f\|_{L^p(X,\gamma)}+\left(\int_X\left(\sum_{j=1}^\infty\lambda_j\left(\frac{\partial f}{\partial e_j}\right)^2\right)^{p/2}d\gamma\right)^{1/p}.$$

One can consider Sobolev spaces $\widetilde{W}^{1,p}(X,\gamma)$ defined as in Definition 9.3.9, with the gradient ∇ replacing the H-gradient ∇_H. Namely, the proof of Proposition 9.3.7 yields that the operator $\nabla:\mathcal{F}C^1_b(X)\to L^p(X,\gamma;X)$ is closable; we define $\widetilde{W}^{1,p}(X,\gamma)$ as the domain of its closure, still denoted by ∇. This choice looks even simpler and more natural; the norm in $\widetilde{W}^{1,p}$ is the graph norm of ∇ and it is given by

$$\begin{aligned}\|f\|_{\widetilde{W}^{1,p}(X,\gamma)} &:= \left(\int_X|f|^p d\gamma\right)^{1/p}+\left(\int_X\|\nabla f\|^p d\gamma\right)^{1/p}\\ &= \left(\int_X|f|^p d\gamma\right)^{1/p}+\left(\int_X\left(\sum_{j=1}^\infty\left(\frac{\partial f}{\partial e_j}\right)^2\right)^{1/p}d\gamma\right)^{1/p}.\end{aligned}\tag{9.3.7}$$

Since $\lim_{k\to\infty}\lambda_k=0$, our Sobolev space $W^{1,p}(X,\gamma)$ strictly contains $\widetilde{W}^{1,p}(X,\gamma)$, and the embedding $\widetilde{W}^{1,p}(X,\gamma)\subset W^{1,p}(X,\gamma)$ is continuous.

9.4 Exercises

Exercise 9.1. Prove Lemma 9.1.1.

Exercise 9.2. (i) Prove that statement (v) of Proposition 9.1.5 is false for $p=1$, $d=1$. (Hint: use Proposition 9.1.6, and the sequence (f_n) defined by $f_n(x)=0$ for $x\le 0$, $f_n(x)=nx$ for $0\le x\le 1/n$, $f_n(x)=1$ for $x\ge 1/n$).

(ii) Using (i), prove that $W^{1,1}(\mathbb{R}, \gamma_1)$ is not reflexive.

Exercise 9.3. Let $(\Omega, \mathscr{F})$ be a measurable space, and let μ be a positive finite measure in Ω. Prove that a function $f : \Omega \to \mathbb{R}$ is measurable if and only if it is the pointwise a.e. limit of a sequence of simple functions.

Exercise 9.4. Prove Lemma 9.3.6.

Lecture 10

Sobolev Spaces II

In this Lecture we go on in the description of the Sobolev spaces $W^{1,p}(X,\gamma)$, and we define the Sobolev spaces $W^{2,p}(X,\gamma)$. We give approximation results through the cylindrical functions $\mathbb{E}_n f$, and we introduce the divergence of vector fields; formally, the divergence operator is the adjoint of the H-gradient. We use the notation of Lecture 9. So, X is a separable Banach space endowed with a centred nondegenerate Gaussian measure γ, and if $\{h_j:\ j\in\mathbb{N}\}\subset R_\gamma(X^*)$ is an orthonormal basis of the Cameron-Martin space H, then for every $f\in W^{1,p}(X,\gamma)$ we denote by $\partial_j f(x)=[\nabla_H f(x),h_j]_H$ the generalised derivative of f in the direction h_j.

10.1 Further properties of $W^{1,p}$ spaces

Let $f\in W^{1,p}(X,\gamma)$, $1<p<\infty$. For every $h\in H$, $\partial_h f$ plays the role of weak derivative of f in the h direction. Indeed, by Proposition 9.3.10, for every $f\in W^{1,p}(X,\gamma)$ and $\varphi\in C^1_b(X)$, applying formula (9.3.3) to the product $f\varphi$ we get

$$\int_X(\partial_h\varphi)f\,d\gamma=-\int_X\varphi(\partial_h f)\,d\gamma+\int_X\varphi\,f\,\hat h\,d\gamma.$$

The Sobolev spaces may be defined through the weak derivatives. Given $f\in L^p(X,\gamma)$ and $h\in H$, a function $g\in L^1(X,\gamma)$ is called *weak derivative* of f in the direction of h if

$$\int_X(\partial_h\varphi)f\,d\gamma=-\int_X\varphi\,g\,d\gamma+\int_X\varphi\,f\,\hat h\,d\gamma,\quad\forall\varphi\in C^1_b(X).$$

The weak derivative is unique, because if $\int_X\varphi\,g\,d\gamma=0$ for every $\varphi\in C^1_b(X)$, then $g=0$ a.e. by Lemma 9.3.6.

We set

$$G^{1,p}(X,\gamma)=\Big\{f\in L^p(X,\gamma):\exists\Psi\in L^p(X,\gamma;H)\text{ such that for each }h\in H,\ [\Psi(\cdot),h]_H\text{ is the weak derivative of }f\text{ in the direction }h\Big\}.$$

If $f \in G^{1,p}(X,\gamma)$ and Ψ is the function in the definition, we set

$$D_H f := \Psi, \quad \|f\|_{G^{1,p}(X,\gamma)} = \|f\|_{L^p(X,\gamma)} + \|\Psi\|_{L^p(X,\gamma;H)}.$$

Theorem 10.1.1. *For every $p > 1$, $G^{1,p}(X,\gamma) = W^{1,p}(X,\gamma)$ and $D_H f = \nabla_H f$ for every $f \in W^{1,p}(X,\gamma)$.*

The proof may be found e.g. in [3, Cor. 5.4.7].

Let us come back to the approximation by conditional expectations introduced in Subsection 7.4. We already know that if $f \in L^p(X,\gamma)$ then $\mathbb{E}_n f \to f$ in $L^p(X,\gamma)$ as $n \to \infty$.

Proposition 10.1.2. *Let $1 \le p < \infty$ and let $f \in W^{1,p}(X,\gamma)$. Then, $\mathbb{E}_n f \in W^{1,p}(X,\gamma)$ for all $n \in \mathbb{N}$ and:*

(i) *for every $j \in \mathbb{N}$*

$$\partial_j(\mathbb{E}_n f) = \begin{cases} \mathbb{E}_n(\partial_j f) & \text{if } j \le n, \\ 0 & \text{if } j > n; \end{cases} \tag{10.1.1}$$

(ii) $\|\mathbb{E}_n f\|_{W^{1,p}(X,\gamma)} \le \|f\|_{W^{1,p}(X,\gamma)}$;

(iii) $\lim_{n\to\infty} \mathbb{E}_n f = f$ *in $W^{1,p}(X,\gamma)$.*

Proof. Let $f \in \mathcal{F}C_b^1(X)$. Since $P_n x = \sum_{i=1}^n \hat{h}_i(x) h_i$ and $\frac{\partial \hat{h}_i}{\partial h_j}(x) = \delta_{ij}$ for every x, for every $y \in X$ the function $x \mapsto f(P_n x + (I - P_n)y)$ has directional derivatives along all h_j, that vanish for $j > n$ and are equal to $\frac{\partial f}{\partial h_j}(P_n x + (I - P_n)y)$ for $j \le n$.
Since $x \mapsto \frac{\partial f}{\partial h_j}(P_n x + (I - P_n)y)$ is continuous and bounded by a constant independent of y, for $j \le n$ we get

$$\partial_j \mathbb{E}_n f(x) = \frac{\partial}{\partial h_j} \int_X f(P_n x + (I - P_n)y)\gamma(dy) = \int_X \frac{\partial f}{\partial h_j}(P_n x + (I - P_n)y)\gamma(dy).$$

In other words, (i) holds, and it yields

$$\nabla_H \mathbb{E}_n f(x) = \int_X P_n \nabla_H f(P_n x + (I - P_n)y)\gamma(dy), \qquad \forall x \in X. \tag{10.1.2}$$

So we have

$$\begin{aligned} \|\nabla_H \mathbb{E}_n f - \nabla_H f\|^p_{L^p(X,\gamma;H)} &= \int_X \left| \int_X (P_n \nabla_H f(P_n x + (I - P_n)y) - \nabla_H f(x))\gamma(dy) \right|_H^p \gamma(dx) \\ &\le \int_X \int_X |P_n \nabla_H f(P_n x + (I - P_n)y) - \nabla_H f(x)|_H^p \,\gamma(dy)\gamma(dx). \end{aligned}$$

Notice that

$$\lim_{n\to+\infty} |P_n \nabla_H f(P_n x + (I - P_n)y) - \nabla_H f(x)|_H = 0, \qquad \gamma \otimes \gamma - a.e.\ (x,y).$$

Indeed, recalling that $\|P_n\|_{\mathcal{L}(H)} \leq 1$,

$$\begin{aligned}
&|P_n\nabla_H f(P_n x + (I-P_n)y) - \nabla_H f(x)|_H \\
&\leq \Big|P_n\Big(\nabla_H f(P_n x + (I-P_n)y) - \nabla_H f(x)\Big)\Big|_H z + |P_n\nabla_H f(x) - \nabla_H f(x)|_H \\
&\leq |\nabla_H f(P_n x + (I-P_n)y) - \nabla_H f(x)|_H + |(P_n - I)\nabla_H f(x)|_H
\end{aligned}$$

and the first summand vanishes as $n \to +\infty$ for $\gamma \otimes \gamma$–a.e. (x, y) since by Theorem 7.1.3

$$\lim_{n\to+\infty} P_n x + (I - P_n)y = x$$

for $\gamma \otimes \gamma$-a.e. (x, y) and $\nabla_H f$ is continuous; the second addendum goes to 0 as $n \to \infty$ for every $x \in X$. Moreover,

$$|P_n\nabla_H f(P_n x + (I-P_n)y) - \nabla_H f(x)|_H \leq 2\sup_{z\in X} |\nabla_H f(z)|_H.$$

By the Dominated Convergence Theorem,

$$\lim_{n\to+\infty} \nabla_H \mathbb{E}_n f = \nabla_H f$$

in $L^p(X, \gamma; H)$, and taking into account Proposition 7.4.4,

$$\lim_{n\to+\infty} \mathbb{E}_n f = f$$

in $W^{1,p}(X, \gamma)$. So, f satisfies (iii). Moreover,

$$\begin{aligned}
\|\nabla_H \mathbb{E}_n f\|^p_{L^p(X,\gamma;H)} &= \int_X \left|\int_X P_n\nabla_H f(P_n x + (I-P_n)y)\gamma(dy)\right|^p_H \gamma(dx) \\
&\leq \int_X \int_X |P_n\nabla_H f(P_n x + (I-P_n)y)|^p_H \gamma(dy)\gamma(dx) \\
&\leq \int_X \int_X |\nabla_H f(P_n x + (I-P_n)y)|^p_H \gamma(dy)\gamma(dx) \\
&= \int_X |\nabla_H f(x)|^p_H \gamma(dx) \qquad (10.1.3)
\end{aligned}$$

where the last equality follows from Proposition 7.3.2.

Estimate (10.1.3) and (7.2.4) yield (ii) for $f \in \mathcal{F}C^1_b(X)$.

Let now $f \in W^{1,p}(X, \gamma)$, and let $(f_k) \subset \mathcal{F}C^1_b(X)$ be a sequence converging to f in $W^{1,p}(X, \gamma)$. By estimate (7.2.4) , for every $n \in \mathbb{N}$ the sequence $(\mathbb{E}_n f_k)_k$ converges to $\mathbb{E}_n f$ in $L^p(X, \gamma)$, and by (ii) $(\mathbb{E}_n f_k)_k$ is a Cauchy sequence in $W^{1,p}(X, \gamma)$. Therefore, $\mathbb{E}_n f \in W^{1,p}(X, \gamma)$ and

$$\nabla_H \mathbb{E}_n f = \lim_{k\to+\infty} \nabla_H \mathbb{E}_n f_k$$

in $L^p(X, \gamma; H)$ so that

$$\begin{aligned}
\|\nabla_H \mathbb{E}_n f\|_{L^p(X,\gamma;H)} &= \lim_{k\to+\infty} \|\nabla_H \mathbb{E}_n f_k\|_{L^p(X,\gamma;H)} \leq \lim_{k\to+\infty} \|\nabla_H f_k\|_{L^p(X,\gamma;H)} \\
&= \|\nabla_H f\|_{L^p(X,\gamma;H)}.
\end{aligned}$$

Therefore (ii) holds for every $f \in W^{1,p}(X,\gamma)$ and then (iii) follows from (ii) and from the density of $\mathcal{F}C_b^1(X)$ in $W^{1,p}(X,\gamma)$.
(i) follows as well and in fact we have

$$\nabla_H \mathbb{E}_n f = \mathbb{E}_n(P_n \nabla_H f) \qquad \forall n \in \mathbb{N},$$

where the right-hand side has to be understood as a Bochner H-valued integral. Indeed, by (10.1.2) we have

$$\nabla_H \mathbb{E}_n f_k(x) = \int_X P_n \nabla_H f_k(P_n x + (I - P_n)y)\gamma(dy)$$

for every $k \in \mathbb{N}$. The left hand side converges to $\nabla_H \mathbb{E}_n f$ in $L^p(X,\gamma;H)$ as $k \to +\infty$. The right hand side converges to $\mathbb{E}_n P_n \nabla_H f$ as $k \to +\infty$ since

$$\begin{aligned}
&\int_X \left| \int_X P_n \nabla_H (f_k - f)(P_n x + (I - P_n)y)\gamma(dy)\right|_H^p \gamma(dx) \\
&\le \int_X \int_X |\nabla_H (f_k - f)(P_n x + (I - P_n)y)|_H^p \,\gamma(dy)\gamma(dx) \\
&= \int_X |\nabla_H (f_k - f)(x)|_H^p \gamma(dx)
\end{aligned}$$

by Proposition 7.3.2. □

Regular L^p cylindrical functions with L^p gradient are in $W^{1,p}(X,\gamma)$, see Exercise 10.2. The simplest nontrivial examples of Sobolev functions are the elements of X_γ^*.

Lemma 10.1.3. *$X_\gamma^* \subset W^{1,p}(X,\gamma)$ for every $p \in [1,+\infty)$, and $\nabla_H \hat{h} = h$ (constant) for every $\hat{h} \in X_\gamma^*$.*

Proof. Fix $1 \le p < \infty$. For every $\hat{h} \in X_\gamma^*$, there exists $\ell_n \in X^*$ such that $\lim_{n\to\infty} \ell_n = \hat{h}$ in $L^2(X,\gamma)$. For every n, $m \in \mathbb{N}$ we have

$$\|\ell_n - \ell_m\|_{L^p(X,\gamma)}^p = \int_{\mathbb{R}} |\xi|^p \mathscr{N}(0, \|\ell_n - \ell_m\|_{L^2(X,\gamma)}^2)(d\xi) = \int_{\mathbb{R}} |\tau|^p \mathscr{N}(0,1)\,(d\tau)\|\ell_n - \ell_m\|_{L^2(X,\gamma)}^p$$

so that (ℓ_n) is a Cauchy sequence in $L^p(X,\gamma)$. Its L^2-limit $\hat{h}$ coincides with its L^p-limit, if $p \neq 2$.

As ℓ_n is in X^*, $\nabla_H \ell_n$ is constant and it coincides with $R_\gamma \ell_n$, see (9.3.1). Since $\lim_{n\to\infty} \ell_n = \hat{h}$ in $L^2(X,\gamma)$ and R_γ is an isometry from X_γ^* to H, $H - \lim_{n\to\infty} R_\gamma \ell_n = R_\gamma \hat{h} = h$. Therefore,

$$\int_X |\nabla_H \ell_n - h|_H^p d\gamma = |R_\gamma \ell_n - h|_H^p \to 0 \quad \text{as n} \to \infty.$$

It follows that $\hat{h} \in W^{1,p}(X,\gamma)$ and $\nabla_H \hat{h} = h$. □

An important example of Sobolev functions is given by Lipschitz functions. Since a Lipschitz function is continuous, it is Borel measurable.

Proposition 10.1.4. *If $f : X \to \mathbb{R}$ is Lipschitz continuous, then $f \in W^{1,p}(X, \gamma)$ for any $1 \le p < +\infty$.*

Proof. Let $L > 0$ be such that

$$|f(x) - f(y)| \le L\|x - y\| \qquad \forall\, x, y \in X.$$

Since $|f(x)| \le |f(0)|+L\|x\|$, by Theorem 2.3.1 (Fernique) $f \in L^p(X, \gamma)$ for any $1 \le p < \infty$.

Let us consider the conditional expectation $\mathbb{E}_n f$.

Let us notice that

$$\mathbb{E}_n f(x) = v_n(\hat{h}_1(x), \ldots, \hat{h}_n(x)),$$

with $v_n : \mathbb{R}^n \to \mathbb{R}$ an L_1–Lipschitz function since

$$\begin{aligned}
|v_n(z + \eta) - v_n(z)| =& \left|\mathbb{E}_n f\Big(\sum_{i=1}^n z_i h_i + \sum_{i=1}^n \eta_i h_i\Big) - \mathbb{E}_n f\Big(\sum_{i=1}^n z_i h_i\Big)\right| \\
\le& \int_X \left| f\Big(\sum_{i=1}^n z_i h_i + \sum_{i=1}^n \eta_i h_i\Big) + (I - P_n)y\Big) - f\Big(\sum_{i=1}^n z_i h_i + (I - P_n)y\Big)\right| \gamma(dy) \\
\le& L_1 \left|\sum_{i=1}^n \eta_i h_i\right|_H = L_1|\eta|_{\mathbb{R}^n},
\end{aligned}$$

where we have used (3.1.3), $\|h\| \le c|h|_H$ for $h \in H$, and we have set $L_1 := cL$. By the Rademacher Theorem, v_n is differentiable λ_n–a.e. in $\mathbb{R}^n$ and $|\nabla v_n(z)|_{\mathbb{R}^n} \le L_1$ for a.e. $z \in \mathbb{R}^n$. Hence $v_n \in W^{1,1}_{\rm loc}(\mathbb{R}^n)$, and

$$\int_{\mathbb{R}^n} |\nabla v_n(z)|^p_{\mathbb{R}^n} \gamma_n(dz) \le L_1^p.$$

We use now the map $T_n : X \to \mathbb{R}^n$, $T_n(x) = (\hat{h}_1(x), \ldots, \hat{h}_n(x))$. If $x \in X$ is a point such that v_n is differentiable at $T_n(x)$, then

$$\partial_h \mathbb{E}_n f(x) = \begin{cases} 0 & \text{if } h \in F_n^\perp \\ \nabla v_n(T_n(x)) \cdot T_n(h) & \text{if } h \in F_n, \end{cases}$$

where $F_n = \text{span}\{h_1, \ldots, h_n\}$ As a consequence, we can write

$$|\nabla_H \mathbb{E}_n f(x)|^2_H = \sum_{i=1}^\infty |\partial_i \mathbb{E}_n f(x)|^2 = \sum_{i=1}^n |\partial_i \mathbb{E}_n f(x)|^2 = |\nabla v_n(T_n(x))|^2_{\mathbb{R}^n}.$$

We claim that for γ–a.e. x v_n is differentiable at $T_n(x)$. Indeed, let $A \subset \mathbb{R}^n$ be such that $\lambda_n(A) = 0$ and v_n is differentiable at any point in $\mathbb{R}^n \setminus A$. Since $\gamma_n \ll \lambda_n$, $\gamma_n(A) = 0$ and then $\gamma(T_n^{-1}(A)) = 0$ because $\gamma \circ T_n^{-1} = \gamma_n$, see Exercise 2.4. Hence v_n is differentiable at any point $T_n(x)$, where $x \in X \setminus T_n^{-1}(A)$.

We know that $\mathbb{E}_n f \to f$ in $L^p(X,\gamma)$ and we have

$$\int_X |\nabla_H \mathbb{E}_n f(x)|_H^p \gamma(dx) = \int_X |\nabla v_n(T_n x)|_{\mathbb{R}^n}^p \gamma(dx) = \int_{\mathbb{R}^n} |\nabla v_n(z)|_{\mathbb{R}^n}^p \gamma_n(dz) \le L_1^p.$$

By Proposition 9.3.10(v) $f \in W^{1,p}(X,\gamma)$ for every $1 < p < \infty$ and by inclusion $f \in W^{1,1}(X,\gamma)$. □

Further properties of $W^{1,p}$ functions are presented in Exercises 10.3, 10.4, 10.5.

10.2 Sobolev spaces of H-valued functions

We recall the definition of Hilbert–Schmidt operators, see e.g. [11, §XI.6] for more information.

Definition 10.2.1. *Let H_1, H_2 be separable Hilbert spaces. A linear operator $A \in \mathcal{L}(H_1,H_2)$ is called a Hilbert–Schmidt operator if there exists an orthonormal basis $\{h_j : j \in \mathbb{N}\}$ of H_1 such that*

$$\sum_{j=1}^\infty \|Ah_j\|_{H_2}^2 < \infty. \tag{10.2.1}$$

If A is a Hilbert–Schmidt operator and $\{e_j : j \in \mathbb{N}\}$ is any orthonormal basis of H_1, $\{y_j : j \in \mathbb{N}\}$ is any orthonormal basis of H_2, then

$$\|Ae_j\|_{H_2}^2 = \sum_{k=1}^\infty \langle Ae_j, y_k\rangle_{H_2}^2 = \sum_{k=1}^\infty \langle e_j, A^* y_k\rangle_{H_2}^2$$

so that

$$\sum_{j=1}^\infty \|Ae_j\|_{H_2}^2 = \sum_{j=1}^\infty \sum_{k=1}^\infty \langle e_j, A^* y_k\rangle_{H_2}^2 = \sum_{k=1}^\infty \sum_{j=1}^\infty \langle e_j, A^* y_k\rangle_{H_2}^2 = \sum_{k=1}^\infty \|A^* y_k\|_{H_1}^2.$$

So, the convergence of the series (10.2.1) and the value of its sum are independent of the basis of H_1. We denote by $\mathcal{H}(H_1,H_2)$ the space of the Hilbert–Schmidt operators from H_1 to H_2, and we set

$$\|A\|_{\mathcal{H}(H_1,H_2)} = \left(\sum_{j=1}^\infty \|Ah_j\|_{H_2}^2\right)^{1/2},$$

for every orthonormal basis $\{h_j : j \in \mathbb{N}\}$ of H_1. Notice that if $H_1 = \mathbb{R}^n$, $H_2 = \mathbb{R}^m$, the Hilbert–Schmidt norm of any linear operator coincides with the Euclidean norm of the associated matrix.

The norm (10.2.1) comes from the inner product

$$\langle A, B\rangle_{\mathcal{H}(H_1,H_2)} = \sum_{j=1}^\infty \langle Ah_j, Bh_j\rangle_{H_2},$$

where for every couple of Hilbert–Schmidt operators A, B, the series on the right-hand side converges for every orthonormal basis $\{h_j : j \in \mathbb{N}\}$ of H_1, and its value is independent of the basis. The space $\mathcal{H}(H_1, H_2)$ is a separable Hilbert space with the above inner product.

If $H_1 = H_2 = H$, where H is the Cameron–Martin space of (X, γ), we set $\mathcal{H} := \mathcal{H}(H, H)$.

It is useful to generalise the notion of Sobolev space to H-valued functions. To this aim, we define the cylindrical E-valued functions as follows, where E is any normed space.

Definition 10.2.2. *For $k \in \mathbb{N}$ we define $\mathcal{F}C_b^k(X, E)$ (respectively, $\mathcal{F}C_b^\infty(X, E)$) as the linear span of the functions $x \mapsto v(x)y$, with $v \in \mathcal{F}C_b^k(X)$ (respectively, $v \in \mathcal{F}C_b^\infty(X)$) and $y \in E$.*

Therefore, every element of $\mathcal{F}C_b^k(X, E)$ may be written as

$$v(x) = \sum_{j=1}^{n} v_j(x) y_j \tag{10.2.2}$$

for some $n \in \mathbb{N}$, and $v_j \in \mathcal{F}C_b^k(X)$, $y_j \in E$. Such functions are Fréchet differentiable at every $x \in X$, with $v'(x) \in \mathcal{L}(X, E)$ given by $v'(x)(h) = \sum_{j=1}^n v_k j'(x)(h) y_j$ for every $h \in X$.

Similarly to the scalar case, we introduce the notion of H-differentiable function.

Definition 10.2.3. *A function $v : X \to E$ is called H-differentiable at $\overline{x} \in X$ if there exists $L \in \mathcal{L}(H, E)$ such that*

$$\|v(\overline{x} + h) - v(\overline{x}) - L(h)\|_E = o(|h|_H) \quad \textit{as } h \to 0 \textit{ in } H.$$

In this case we set $L =: D_H v(\overline{x})$.

If $v \in \mathcal{F}C_b^1(X, E)$ is given by $v(\cdot) = \psi(\cdot) y$ with $\psi \in \mathcal{F}C_b^1(X)$ and $y \in E$, then v is H-differentiable at every $\overline{x} \in X$, and

$$D_H v(\overline{x})(h) = [\nabla_H \psi(\overline{x}), h]_H \, y.$$

In particular, if $E = H$ and $\{h_j : j \in \mathbb{N}\}$ is any orthonormal basis of H we have

$$|D_H v(\overline{x})(h_j)|_H^2 \leq |[\nabla_H \psi(\overline{x}), h_j]_H^2 \, |y|_H^2$$

so that $D_H v(\overline{x})$ is a Hilbert–Schmidt operator, and we have

$$\begin{aligned} |D_H v(\overline{x})|_{\mathcal{H}}^2 &= \sum_{j=1}^{\infty} |D_H v(\overline{x})(h_j)|^2 = \sum_{j=1}^{\infty} [\nabla_H \psi(\overline{x}), h_j]_H^2 |y|_H^2 \\ &= |\nabla_H \psi(\overline{x})|_H^2 |y|_H^2. \end{aligned}$$

Moreover, $x \mapsto \nabla_H \psi(x)$ is continuous and bounded. In addition, the operator $J : H \to \mathcal{H}$,

$$(Jk)(h) := [k, h]_H y, \qquad k, h \in H$$

is bounded since

$$|Jk|^2_{\mathcal{H}} = \sum_{j=1}^{\infty} |[k, h_j]_H y|^2_H = |k|^2_H |y|^2_H.$$

Then $x \mapsto D_H v(x) = J(\nabla_H \psi(x))$ is continuous and bounded from X to $\mathcal{H}$. In particular, it belongs to $L^p(X, \gamma; \mathcal{H})$ for every $1 \le p < \infty$.

The procedure to define Sobolev spaces of H-valued functions is similar to the procedure for scalar functions. Namely, we show that the operator D_H, seen as an unbounded operator from $L^p(X, \gamma; H)$ to $L^p(X, \gamma; \mathcal{H})$ with domain $\mathcal{F}C^1_b(X, H)$, is closable.

Lemma 10.2.4. *For every $1 \le p < \infty$, the operator $D_H : \mathcal{F}C^1_b(X, H) \to L^p(X, \gamma; \mathcal{H})$ is closable in $L^p(X, \gamma; H)$.*

Proof. Let $v(\cdot) = \psi(\cdot) y$, with $\psi \in \mathcal{F}C^1_b(X)$, $y \in H$, and let $h, k \in H$. Then

$$\nabla_H [v(\cdot), h]_H = [y, h]_H \nabla_H \psi(\cdot),$$

and

$$\begin{aligned} [\nabla_H [v(\cdot), h]_H k]_H &= [y, h]_H [\nabla_H \psi(\cdot), k]_H = [[\nabla_H \psi(\cdot), k]_H y, h]_H \\ &= [D_H(\psi(\cdot) y)(k), h]_H = [(D_H v(\cdot))^\star h, k]_k. \end{aligned}$$

As this holds for all $k \in H$ one obtains

$$\nabla_H [v(\cdot), h]_H = (D_H \psi(\cdot))^\star h, \qquad h \in H. \tag{10.2.3}$$

Taking linear combinations one obtains (10.2.3) for all $v \in \mathcal{F}C^1_b(X; H)$.

Let (v_n) be a sequence in $\mathcal{F}C^1_b(X; H)$, $v_n \to 0$ in $L^p(X, \gamma; H)$, $D_H v_n \to \Phi$ in $L^p(X, \gamma; \mathcal{H})$, and let $h \in H$. Then $[v_n(\cdot), h]_H \to 0$ in $L^p(X, \gamma)$, and (10.2.3) implies

$$\nabla_H [v_n(\cdot), h]_H = (D_H v_n(\cdot))^\star (h) \to \Phi(\cdot)^\star (h)$$

in $L^p(X, \gamma; H)$. Indeed

$$\begin{aligned} \int_X |\nabla_H [v_n(x), h]_H - \Phi(x)^\star (h)|^p_H \gamma(dx) &= \int_X |(D_H v_n(x))^\star (h) - \Phi(x)^\star (h)|^p_H \gamma(dx) \\ &\le |h|^p_H \int_X \|(D_H v_n(x))^\star - \Phi(x)^\star\|^p_{\mathcal{L}(H)} \gamma(dx) \\ &\le |h|^p_H \int_X \|(D_H v_n(x))^\star - \Phi(x)^\star\|^p_{\mathcal{H}(H)} \gamma(dx), \end{aligned}$$

where we have used the relations

$$\|A\|_{\mathcal{L}(H)} \le \|A\|_{\mathcal{H}(H)}, \quad \|A^\star\|_{\mathcal{H}(H)} = \|A\|_{\mathcal{H}(H)}.$$

Since ∇_H is closable as an operator from $L^p(X, \gamma)$ to $L^p(X, \gamma; H)$, by Proposition 9.3.7 one obtains $\Phi(\cdot)^\star h = 0$. As this holds for all $h \in H$, and H is separable, one concludes that $\Phi(x)^\star = 0$ for a.e. $x \in X$, and therefore $\Phi = 0$. □

Definition 10.2.5. *For every $1 \le p < \infty$ we define $W^{1,p}(X,\gamma;H)$ as the domain of the closure of the operator $D_H : \mathcal{F}C_b^1(X,H) \to L^p(X,\gamma;\mathcal{H})$ (still denoted by D_H) in $L^p(X,\gamma;H)$.*

Then, $W^{1,p}(X,\gamma;H)$ is a Banach space with the graph norm

$$\begin{aligned}\|V\|_{W^{1,p}(X,\gamma;H)} &= \Big(\int_X |V(x)|_H^p d\gamma\Big)^{1/p} + \Big(\int_X |D_H V(x)|_{\mathcal{H}}^p d\gamma\Big)^{1/p} \\ &= \Big(\int_X\Big(\sum_{j=1}^\infty [V(x),h_j]_H^2\Big)^{p/2} d\gamma\Big)^{1/p} + \Big(\int_X\Big(\sum_{i,j=1}^\infty [D_H V(x)(h_i),h_j]_H^2\Big)^{p/2} d\gamma\Big)^{1/p}.\end{aligned}$$

Let $v \in \mathcal{F}C_b^1(X,H)$,

$$v(x) = \sum_{k=1}^n \varphi_k(x) y_k,$$

with $\varphi_k \in \mathcal{F}C_b^1(X)$ and $y_k \in H$. Then v may be written in the form

$$v(x) = \sum_{j=1}^\infty v_j(x) h_j,$$

where the series converges in $W^{1,p}(X,\gamma;H)$. Indeed, setting

$$v_j(x) = [v(x),h_j]_H = \sum_{k=1}^n \varphi_k(x)[y_k,h_j]_H, \quad j \in \mathbb{N}$$

the sequence $s_m(x) = \sum_{j=1}^m v_j(x)h_j$ converges to v in $W^{1,p}(X,\gamma;H)$, because for each $k = 1,\dots,n$, the sequence $\sum_{j=1}^m \varphi_k(x)[y_k,h_j]_H$ converges to $\varphi_k(x)y$ in $W^{1,p}(X,\gamma;H)$. Moreover,

$$D_H v(x)(h) = \sum_{j=1}^\infty [\nabla_H v_j(x),h]_H h_j$$

so that, as in finite dimension,

$$[D_H v(x)(h_i),h_j]_H = [\nabla_H v_j(x),h_i]_H = \partial_i v_j(x).$$

10.2.1 The divergence operator

Let us recall the definition of adjoint operators. If X_1, X_2 are Hilbert spaces and $T : D(T) \subset X_1 \to X_2$ is a densely defined linear operator, an element $v \in X_2$ belongs to $D(T^*)$ iff the function $D(T) \to \mathbb{R}$, $f \mapsto \langle Tf, v\rangle_{X_2}$ has a linear continuous extension to the whole X_1, namely there exists $g \in X_1$ such that

$$\langle Tf, v\rangle_{X_2} = \langle f, g\rangle_{X_1}, \quad f \in D(T).$$

In this case g is unique (because $D(T)$ is dense in X_1) and we set

$$g = T^* v.$$

We are interested now in the case $X_1 = L^2(X,\gamma)$, $X_2 = L^2(X,\gamma;H)$ and $T = \nabla_H$. For $f \in W^{1,2}(X,\gamma)$, $v \in L^2(X,\gamma;H)$ we have

$$\langle Tf, v\rangle_{L^2(X,\gamma;H)} = \int_X [\nabla_H f(x), v(x)]_H \,\gamma(dx)$$

so that $v \in D(T^*)$ if and only if there exists $g \in L^2(X,\gamma)$ such that

$$\int_X [\nabla_H f(x), v(x)]_H \,\gamma(dx) = \int_X f(x)g(x)\,\gamma(dx), \quad f \in W^{1,2}(X,\gamma). \tag{10.2.4}$$

In this case, in analogy to the finite dimensional case, we set

$$\mathrm{div}_\gamma v := -g$$

and we call $-g$ *divergence* or *Gaussian divergence* of v. As $\mathcal{F}C_b^1(X)$ is dense in $W^{1,2}(X,\gamma)$, (10.2.4) is equivalent to

$$\int_X [\nabla_H f(x), v(x)]_H \,\gamma(dx) = \int_X f(x)g(x)\,\gamma(dx), \quad f \in \mathcal{F}C_b^1(X).$$

The main achievement of this section is the embedding $W^{1,2}(X,\gamma;H) \subset D(T^*)$. For its proof, we use the following lemma.

Lemma 10.2.6. *For every $f \in W^{1,2}(X,\gamma)$ and $h \in H$, $f\hat{h} \in L^2(X,\gamma)$ and*

$$\int_X (f\hat{h})^2 d\gamma \le 4\int_X (\partial_h f)^2 d\gamma + 2|h|_H^2 \int_X f^2\,d\gamma. \tag{10.2.5}$$

Proof. We already know that $\hat{h} \in W^{1,2}(X,\gamma)$. Then, for every $f \in \mathcal{F}C_b^1(X)$ we have $f^2\hat{h} \in W^{1,2}(X,\gamma)$ and

$$\begin{aligned}
\int_X (f\hat{h})^2 d\gamma &= \int_X (f^2\hat{h})\,\hat{h}\,d\gamma = \int_X \partial_h(f^2\hat{h})\,d\gamma \quad \text{(by Proposition 9.3.10 (i))}\\
&= \int_X (2f\,\partial_h f\,\hat{h} + f^2\partial_h(\hat{h}))d\gamma\\
&= 2\int_X f\,\hat{h}\,\partial_h f\,d\gamma + |h|_H^2\int_X f^2\,d\gamma\\
&\le 2\Big(\int_X (f\hat{h})^2 d\gamma\Big)^{1/2}\Big(\int_X (\partial_h f)^2 d\gamma\Big)^{1/2} + |h|_H^2\int_X f^2\,d\gamma.
\end{aligned}$$

Using the inequality $ab \le a^2/4 + b^2$, we get

$$\int_X (f\hat{h})^2 d\gamma \le \frac{1}{2}\int_X (f\hat{h})^2 d\gamma + 2\int_X (\partial_h f)^2 d\gamma + |h|_H^2\int_X f^2\,d\gamma$$

so that f satisfies (10.2.5). Since $\mathcal{F}C_b^1(X)$ is dense in $W^{1,2}(X,\gamma)$, (10.2.5) holds for every $f \in W^{1,2}(X,\gamma)$. $\square$

Theorem 10.2.7. *The Sobolev space $W^{1,2}(X,\gamma;H)$ is continuously embedded in $D(\mathrm{div}_\gamma)$ and the estimate*

$$\|\mathrm{div}_\gamma v\|_{L^2(X,\gamma)} \leq \|v\|_{W^{1,2}(X,\gamma;H)}$$

holds. Moreover, fixing an orthonormal basis $\{h_n:\ n\in\mathbb{N}\}$ of H contained in $R_\gamma(X^*)$, and setting $v_n(x)=[v(x),h_n]_H$ for every $v\in W^{1,2}(X,\gamma;H)$ and $n\in\mathbb{N}$, we have

$$\mathrm{div}_\gamma v(x) = \sum_{n=1}^\infty \big(\partial_n v_n(x) - v_n(x)\hat{h}_n(x)\big),$$

where the series converges in $L^2(X,\gamma)$.

Proof. Consider a function $v\in W^{1,2}(X,\gamma;H)$ of the type

$$v(x)=\sum_{i=1}^n v_i(x)h_i,\quad x\in X. \tag{10.2.6}$$

with $v_i\in W^{1,2}(X,\gamma)$.

For every $f\in W^{1,2}(X,\gamma)$ we have $[\nabla_H f(x),v(x)]_H=\sum_{i=1}^n \partial_i f(x)v_i(x)$, so that

$$\int_X [\nabla_H f,v]_H d\gamma = \int_X\bigg(\sum_{i=1}^n \partial_i f\, v_i\bigg)d\gamma = \int_X \sum_{i=1}^n(-\partial_i v_i + v_i\hat{h}_i)f\,d\gamma$$

which yields

$$\mathrm{div}_\gamma v = \sum_{i=1}^n(\partial_i v_i - \hat{h}_i v_i).$$

Now we prove that

$$\int_X(\mathrm{div}_\gamma v)^2 d\gamma = \int_X |v|_H^2 d\gamma + \int_X\sum_{i,j=1}^n \partial_i v_j\,\partial_j v_i\,d\gamma, \tag{10.2.7}$$

showing, more generally, that if $u(x)=\sum_{i=1}^n u_i(x)h_i$ is another function of this type, then

$$\int_X(\mathrm{div}_\gamma v\,\mathrm{div}_\gamma u)d\gamma = \int_X [u,v]_H d\gamma + \int_X\sum_{i,j=1}^n \partial_i u_j\,\partial_j v_i\,d\gamma. \tag{10.2.8}$$

By linearity, it is sufficient to prove that (10.2.8) holds if the sums in u and v consist of a single summand, $u(x)=f(x)h_i$, $v(x)=g(x)h_j$ for some f, $g\in W^{1,2}(X,\gamma)$ and i, $j\in\mathbb{N}$. In this case, (10.2.8) reads

$$\int_X(\partial_i f-\hat{h}_i f)(\partial_j g-\hat{h}_j g)d\gamma = \int_X fg\delta_{ij}\,d\gamma + \int_X \partial_j f\,\partial_i g\,d\gamma. \tag{10.2.9}$$

First, let $f,\ g \in \mathcal{F}C_b^2(X)$. Then,

$$\begin{aligned}\int_X (\partial_i f - \hat{h}_i f)(\partial_j g - \hat{h}_j g) d\gamma &= -\int_X f \partial_i(\partial_j g - \hat{h}_j g) d\gamma \\ &= -\int_X f \partial_{ij} g\, d\gamma + \int_X f g \delta_{ij}\, d\gamma + \int_X f \hat{h}_j \partial_i g\, d\gamma \\ &= \int_X (\partial_j f - \hat{h}_j f) \partial_i g\, d\gamma + \int_X f g \delta_{ij}\, d\gamma + \int_X f \hat{h}_j \partial_i g\, d\gamma\end{aligned}$$

so that (10.2.9) holds. Since $\mathcal{F}C_b^2(X)$ is dense in $W^{1,2}(X,\gamma)$, see Exercise 10.6, (10.2.9) holds for $f,\ g \in W^{1,2}(X,\gamma)$. Summing up, (10.2.8) follows, and taking $u = v$, (10.2.7) follows as well. Since the linear span of functions in (10.2.6) is dense in $W^{1,2}(X,\gamma;H)$ both equalities hold in the whole $W^{1,2}(X,\gamma;H)$. Notice also that (10.2.7) implies

$$\int_X (\mathrm{div}_\gamma v)^2 d\gamma \le \int_X |v|_H^2 d\gamma + \int_X \|D_H v\|_{\mathcal{H}}^2\, d\gamma. \tag{10.2.10}$$

If $v \in W^{1,2}(X,\gamma;H)$ we approximate it by the sequence

$$v_n(x) = \sum_{i=1}^n [v(x), h_i]_H h_i.$$

For every $f \in W^{1,2}(X,\gamma)$ we have

$$\int_X [\nabla_H f, v_n]_H d\gamma = -\int_X f \mathrm{div}_\gamma v_n\, d\gamma. \tag{10.2.11}$$

By estimate (10.2.10), $(\mathrm{div}_\gamma v_n)$ is a Cauchy sequence in $L^2(X,\gamma)$, so that it converges in $L^2(X,\gamma)$ to $g(x) := \sum_{j=1}^\infty (\partial_j v_j(x) - v_j(x)\hat{h}_j(x))$. Letting $n \to \infty$ in (10.2.11), we get

$$\int_X [\nabla_H f, v]_H d\gamma = -\int_X f\, g\, d\gamma,$$

so that $v \in D(T^*)$ and $\mathrm{div}_\gamma v = g$. □

Note that the domain of the divergence is larger than $W^{1,2}(X,\gamma;H)$, even in finite dimension. For instance, if $X = \mathbb{R}^2$ is endowed with the standard Gaussian measure, any vector field $v(x,y) = (\alpha_1(x) + \beta_1(y), \alpha_2(x) + \beta_2(y))$ with $\alpha_1,\ \beta_2 \in W^{1,2}(\mathbb{R},\gamma_1)$, $\beta_1,\ \alpha_2 \in L^2(\mathbb{R},\gamma_1)$ belongs to the domain of the divergence, but it does not belong to $W^{1,2}(\mathbb{R}^2,\gamma_2;\mathbb{R}^2)$ unless also $\beta_1,\ \alpha_2 \in W^{1,2}(\mathbb{R},\gamma_1)$.

The divergence may be defined, still as a dual operator, also in a L^p context with $p \neq 2$. We recall that if X_1, X_2 are Banach spaces and : $D(T) \subset X_1 \to X_2$ is a densely defined linear operator, an element $v \in X_2^*$ belongs to $D(T^*)$ iff the function $D(T) \to \mathbb{R}$, $f \mapsto v(Tf)$ has a continous linear extension to the whole X_1. Such extension is an element of X_1^*; denoting it by ℓ we have $\ell(f) = v(Tf)$ for every $f \in D(T)$.

We are interested in the case $X_1 = L^p(X,\gamma)$, $X_2 = L^p(X,\gamma;H)$, with $1 < p < \infty$, and $T : D(T) = W^{1,p}(X,\gamma)$, $Tf = \nabla_H f$. The dual space X_2^* consists of all the functions of the type

$$w \mapsto \int_X [w, v]_H d\gamma,$$

$v \in L^{p'}(X,\gamma;H)$, $p' = p/(p-1)$, see [8, §IV.1], so we canonically identify $L^{p'}(X,\gamma;H)$ as $L^p(X,\gamma;H)^*$. We also identify $(L^p(X,\gamma))^*$ with $L^{p'}(X,\gamma)$. After these identifications, a function $v \in L^{p'}(X,\gamma;H)$ belongs to $D(T^*)$ iff there exists $g \in L^{p'}(X,\gamma)$ such that

$$\int_X [\nabla_H f(x), v(x)]_H \gamma(dx) = \int_X f(x)g(x)\gamma(dx), \quad \forall\ f \in W^{1,p}(X,\gamma),$$

which is equivalent to

$$\int_X [\nabla_H f(x), v(x)]_H \gamma(dx) = \int_X f(x)g(x)\gamma(dx), \quad \forall\ f \in \mathcal{F}C_b^1(X),$$

since $\mathcal{F}C_b^1(X)$ is dense in $W^{1,p}(X,\gamma)$. So, this is similar to the case $p = 2$, see (10.2.4).

Theorem 10.2.8. *Let $1 < p < \infty$, and let $T : D(T) = W^{1,p}(X,\gamma) \to L^p(X,\gamma;H)$, $Tf = \nabla_H f$. Then $W^{1,p}(X,\gamma;H) \subset D(T^*)$, and for every orthonormal basis $\{h_n : n \in \mathbb{N}\}$ of H we have*

$$T^*v(x) = -\sum_{n=1}^{\infty}(\partial_n v_n(x) - v_n(x)\hat{h}_n(x)), \quad v \in W^{1,p}(X,\gamma;H)$$

where $v_n(x) = [v(x), h_n]_H$, and the series converges in $L^p(X,\gamma)$.

The proof of Theorem 10.2.8 for $p \neq 2$ is not as easy as in the case $p = 2$. See [3, Prop. 5.8.8]. The difficult part is the estimate

$$\|T^*v\|_{L^p(X,\gamma)} \leq C\|v\|_{W^{1,p}(X,\gamma;H)},$$

even for good vector fields $v = \sum_{i=1}^n v_i(x)h_i$, with $v_i \in \mathcal{F}C_b^1(X)$.

We may still call "Gaussian divergence" the operator T^*.

10.3 The Sobolev spaces $W^{2,p}(X,\gamma)$

Let us start with regular functions, recalling the definition of the second order derivative $f''(x)$ given in Lecture 9. If $f : X \to \mathbb{R}$ is differentiable at any $x \in X$, we consider the function $X \to X^*$, $x \mapsto f'(x)$. If this function is differentiable at $\overline{x}$, we say that f is twice (Fréchet) differentiable at $\overline{x}$. In this case there exists $L \in \mathcal{L}(X, X^*)$ such that

$$\|f'(\overline{x}+h) - f'(\overline{x}) - Lh\|_{X^*} = o(\|h\|) \quad \text{as } h \to 0 \ \text{ in } X,$$

and we set $L =: f''(\overline{x})$.

In our setting we are interested in increments $h \in H$, and in H-differentiable functions. If $f : X \to \mathbb{R}$ is H-differentiable at any $x \in X$, we say that f is twice H-differentiable at $\overline{x}$ if there exists a linear operator $L_H \in \mathcal{L}(H)$ such that

$$|\nabla_H f(\overline{x}+h) - \nabla_H f(\overline{x}) - L_H h|_H = o(|h|_H) \quad \text{as } h \to 0 \text{ in } H.$$

The operator L_H is denoted by $D^2_H f(\overline{x})$, and by Definition 10.2.3, we have that $D^2_H f(\overline{x}) = D_H \nabla_H f(\overline{x})$.

We recall that if f is differentiable at x, it is also H-differentiable and we have $\nabla_H f(x) = R_\gamma f'(x)$. So, if f is twice differentiable at $\overline{x}$, with $f''(\overline{x}) = L$, then $D^2_H f(\overline{x})h = R_\gamma Lh$. Indeed,

$$|R_\gamma f'(\overline{x}+h) - R_\gamma f'(\overline{x}) - R_\gamma Lh|_H \le \|R_\gamma\|_{\mathcal{L}(X^*,H)} \|f'(\overline{x}+h) - f'(\overline{x}) - Lh\|_{X^*} = o(\|h\|)$$

as $h \to 0$ in X, and therefore,

$$|R_\gamma f'(\overline{x}+h) - R_\gamma f'(\overline{x}) - R_\gamma Lh|_H = o(|h|_H) \quad \text{as } h \to 0 \text{ in } H.$$

If $f \in \mathcal{F}C^2_b(X)$, $f(x) = \varphi(\ell_1(x), \dots, \ell_n(x))$ with $\varphi \in C^2_b(\mathbb{R}^n)$, $\ell_k \in X^*$, then f is twice differentiable at any $\overline{x} \in X$ and

$$(f''(\overline{x})v)(w) = \sum_{i,j=1}^n \partial_i \partial_j \varphi(\ell_1(\overline{x}), \dots, \ell_n(\overline{x}) \ell_i(v) \ell_j(w), \quad v,\ w \in X$$

so that

$$[D^2_H f(\overline{x})h, k]_H = \sum_{i,j=1}^n \partial_i \partial_j \varphi(\ell_1(\overline{x}), \dots, \ell_n(\overline{x}) [R_\gamma \ell_i, h]_H \, [R_\gamma \ell_j, k]_H, \quad h,\ k \in H.$$

$D^2_H f(\overline{x})$ is a Hilbert–Schmidt operator, since for any orthonormal basis $\{h_j :\ j \in \mathbb{N}\}$ of H we have

$$\begin{aligned}
\sum_{m,k=1}^\infty [D^2_H f(\overline{x}) h_m, h_k]_H^2 &\le \sum_{m,k=1}^\infty \Big(\sum_{i,j=1}^n (\partial_i \partial_j \varphi)^2 \Big) \Big(\sum_{i=1}^n [R_\gamma \ell_i, h_m]_H^2 \Big) \Big(\sum_{j=1}^n [R_\gamma \ell_j, h_k]_H^2 \Big) \\
&= \|D^2_H \varphi\|^2_{\mathcal{H}(\mathbb{R}^n,\mathbb{R}^n)} \sum_{m,k=1}^\infty \sum_{i=1}^n [R_\gamma \ell_i, h_m]_H^2 \sum_{j=1}^n [R_\gamma \ell_j, h_k]_H^2 \\
&= \|D^2 \varphi\|^2_{\mathcal{H}(\mathbb{R}^n,\mathbb{R}^n)} \sum_{i=1}^n \sum_{m=1}^\infty [R_\gamma \ell_i, h_m]_H^2 \sum_{j=1}^n \sum_{k=1}^\infty [R_\gamma \ell_j, h_k]_H^2 \\
&= \|D^2 \varphi\|^2_{\mathcal{H}(\mathbb{R}^n,\mathbb{R}^n)} \sum_{i=1}^n |R_\gamma \ell_i|_H^2 \sum_{j=1}^n |R_\gamma \ell_j|_H^2
\end{aligned}$$

where the derivatives of φ are evaluated at $(\ell_1(\overline{x}), \dots, \ell_n(\overline{x}))$. Since $\|D^2 \varphi\|_{\mathcal{H}(\mathbb{R}^n,\mathbb{R}^n)}$ is bounded, $x \to \|D^2_H f(x)\|_{\mathcal{H}}$ is bounded in X.

The next lemma is an immediate consequence of Proposition 9.3.7 and Lemma 10.2.4.

Lemma 10.3.1. *For every $1 \le p < \infty$, the operator*

$$(\nabla_H, D^2_H) : \mathcal{F}C^2_b(X) \to L^p(X,\gamma;H) \times L^p(X,\gamma;\mathcal{H})$$

is closable in $L^p(X,\gamma)$.

Proof. Let (f_n) be a sequence in $\mathcal{F}C^2_b(X)$ such that $f_n \to 0$ in $L^p(X,\gamma)$, $\nabla_H f_n \to G$ in $L^p(X,\gamma;H)$ and $D^2_H f_n = D_H \nabla_H f_n \to \Phi$ in $L^p(X,\gamma;\mathcal{H})$. Then Proposition 9.3.7 implies $G = 0$ and Lemma 10.2.4 implies $\Phi = 0$. □

Definition 10.3.2. *For every $1 \le p < \infty$, $W^{2,p}(X,\gamma)$ is the domain of the closure of*

$$(\nabla_H, D^2_H) : \mathcal{F}C^2_b(X) \to L^p(X,\gamma;H) \times L^p(X,\gamma;\mathcal{H})$$

in $L^p(X,\gamma)$. Therefore, $f \in L^p(X,\gamma)$ belongs to $W^{2,p}(X,\gamma)$ iff there exists a sequence $(f_n) \subset \mathcal{F}C^2_b(X)$ such that $f_n \to f$ in $L^p(X,\gamma)$, $\nabla_H f_n$ converges in $L^p(X,\gamma;H)$ and $D^2_H f_n$ converges in $L^p(X,\gamma;\mathcal{H})$. In this case we set $D^2_H f := \lim_{n\to\infty} D^2_H f_n$.

$W^{2,p}(X,\gamma)$ is a Banach space with the graph norm

$$\begin{aligned} \|f\|_{W^{2,p}} &:= \|f\|_{L^p(X,\gamma)} + \|\nabla_H f\|_{L^p(X,\gamma;H)} + \|D^2_H f\|_{L^p(X,\gamma;\mathcal{H})} \\ &= \Big(\int_X |f|^p d\gamma\Big)^{1/p} + \Big(\int_X |\nabla_H f|^p_H d\gamma\Big)^{1/p} + \Big(\int_X |D^2_H f|^p_{\mathcal{H}} d\gamma\Big)^{1/p}. \end{aligned} \tag{10.3.1}$$

Fixed any orthonormal basis $\{h_j : j \in \mathbb{N}\}$ of H, for every $f \in W^{2,p}(X,\gamma)$ we set

$$\partial_{ij} f(x) = [D^2_H f(x) h_j, h_i]_H.$$

For every sequence of approximating functions f_n we have

$$[D^2_H f_n(x) h_j, h_i]_H = [D^2_H f_n(x) h_i, h_j]_H, \quad x \in X,\ i,j \in \mathbb{N},$$

then the equality

$$\partial_{ij} f(x) = \partial_{ji} f(x), \quad \text{a.e.}$$

holds. Therefore, the $W^{2,p}$ norm may be rewritten as

$$\Big(\int_X |f|^p d\gamma\Big)^{1/p} + \Big(\int_X \Big(\sum_{j=1}^\infty (\partial_j f)^2\Big)^{p/2} d\gamma\Big)^{1/p} + \Big(\int_X \Big(\sum_{i,j=1}^\infty (\partial_{ij} f)^2\Big)^{p/2} d\gamma\Big)^{1/p}.$$

Let X be a Hilbert space and assume that γ is nondegenerate. Then, another class of $W^{2,p}$ spaces looks more natural. As in Remark 9.3.11, we may replace $(\nabla_H f, D^2_H f)$ in Definition 10.3.2 by $(\nabla f, f'')$. The proof of Lemma 10.3.1 works as well with this choice. So, we define $\widetilde{W}^{2,p}(X,\gamma)$ as the domain of the closure of the operator $T : \mathcal{F}C^2_b(X) \to L^p(X,\gamma;X) \times L^p(X,\gamma;\mathcal{H}(X,X))$, $f \mapsto (\nabla f, f'')$ in $L^p(X,\gamma)$ (still denoted by T), and we endow it with the graph norm of T. This space is much smaller than $W^{2,p}(X,\gamma)$ if X is infinite dimensional. Indeed, fix as usual any orthonormal basis $\{e_j : j \in \mathbb{N}\}$ of X

consisting of eigenvectors of Q, $Qe_j = \lambda_j e_j$, and set $h_j = \sqrt{\lambda_j} e_j$. Then $\{h_j :\ j \in \mathbb{N}\}$ is a orthonormal basis of H, $\partial_j f(x) = \sqrt{\lambda_j}\partial f/\partial e_j$, $\partial_{ij} f(x) = \sqrt{\lambda_i\lambda_j}\partial^2 f/\partial e_i \partial e_j$, and

$$\begin{aligned}\|f\|_{W^{2,p}(X,\gamma)} =& \|f\|_{L^p(X,\gamma)} + \left(\int_X \left(\sum_{j=1}^\infty \lambda_j \left(\frac{\partial f}{\partial e_j}\right)^2\right)^{p/2} d\gamma\right)^{1/p} \\ &+ \left(\int_X \left(\sum_{i,j=1}^\infty \lambda_i\lambda_j \left(\frac{\partial^2 f}{\partial e_i \partial e_j}\right)^2\right)^{p/2} d\gamma\right)^{1/p},\end{aligned}$$

while

$$\begin{aligned}\|f\|_{\widetilde{W}^{2,p}(X,\gamma)} =& \|f\|_{L^p(X,\gamma)} + \left(\int_X \left(\sum_{j=1}^\infty \left(\frac{\partial f}{\partial e_k}\right)^2\right)^{p/2} d\gamma\right)^{1/p} \\ &+ \left(\int_X \left(\sum_{i,j=1}^\infty \left(\frac{\partial^2 f}{\partial e_i \partial e_j}\right)^2\right)^{p/2} d\gamma\right)^{1/p}.\end{aligned}$$

Since $\lim_{j\to\infty} \lambda_j = 0$, the $\widetilde{W}^{2,p}(X,\gamma)$ norm is stronger than the $W^{2,p}(X,\gamma)$ norm. In particular, the function $f(x) = \|x\|^2$ belongs to $W^{2,p}(X,\gamma)$ for every $1 \le p < \infty$ but it does not belong to $\widetilde{W}^{2,p}(X,\gamma)$ for any $1 \le p < \infty$, because $f''(x) = 2I$ for every $x \in X$ and $\partial^2 f/\partial e_i \partial e_j = 2\delta_{ij}$.

10.4 Exercises

Exercise 10.1. Prove that (10.1.2) holds.

Exercise 10.2. Prove that if $f \in \mathcal{F}C^1(X) \cap L^p(X,\gamma)$, $1 \le p < \infty$ and $\nabla_H f \in L^p(X,\gamma)$ then $f \in W^{1,p}(X,\gamma)$.

Exercise 10.3. Prove that if $f \in W^{1,p}(X,\gamma)$ then $f^+, f^-, |f| \in W^{1,p}(X,\gamma)$ as well. Compute $\nabla_H f^+, \nabla_H f^-, \nabla_H |f|$ and deduce that $\nabla_H f = 0$ a.e. on $\{f = c\}$ for every $c \in \mathbb{R}$.

Exercise 10.4. Let $\varphi \in W^{1,p}(\mathbb{R}^n, \gamma_n)$ and let $\ell_1, \dots, \ell_n \in X^*$, with $\langle \ell_i, \ell_j\rangle_{L^2(X,\gamma)} = \delta_{ij}$. Prove that the function $f : X \to \mathbb{R}$ defined by $f(x) = \varphi(\hat{h}_1(x), \dots, \hat{h}_n(x))$ belongs to $W^{1,p}(X,\gamma)$.

Exercise 10.5. Let $f \in L^p(X,\gamma)$, $p > 1$, be such that $\mathbb{E}_n f \in W^{1,p}(X,\gamma)$ for every $n \in \mathbb{N}$, with $\sup_n \|\nabla_H \mathbb{E}_n f\|_{L^p(X,\gamma;H)} < \infty$. Prove that $f \in W^{1,p}(X,\gamma)$.

Exercise 10.6. Prove that $\mathcal{F}C^2_b(X)$ is dense in $W^{1,2}(X,\gamma)$

Lecture 11

Semigroups of Operators

In this Lecture we gather a few notions on one-parameter semigroups of linear operators, confining to the essential tools that are needed in the sequel. As usual, X is a real or complex Banach space, with norm $\|\cdot\|$. In this lecture Gaussian measures play no role.

11.1 Strongly continuous semigroups

Definition 11.1.1. *Let $\{T(t) : t \geq 0\}$ be a family of operators in $\mathcal{L}(X)$. We say that it is a* semigroup *if*

$$T(0) = I, \quad T(t+s) = T(t)T(s) \ \ \forall\, t, s \geq 0.$$

A semigroup is called strongly continuous *(or C_0-semigroup) if for every $x \in X$ the function $T(\cdot)x : [0,\infty) \to X$ is continuous.*

Let us present the most elementary properties of strongly continuous semigroups.

Lemma 11.1.2. *Let $\{T(t) : t \geq 0\} \subset \mathcal{L}(X)$ be a semigroup. The following properties hold:*

(a) if there exist $\delta > 0$, $M \geq 1$ such that

$$\|T(t)\| \leq M, \ \ 0 \leq t \leq \delta,$$

then, setting $\omega = (\log M)/\delta$ we have

$$\|T(t)\| \leq M e^{\omega t}, \ \ t \geq 0. \tag{11.1.1}$$

Moreover, for every $x \in X$ the function $t \mapsto T(t)x$ is continuous in $[0,\infty)$ iff it is continuous at 0.

(b) If $\{T(t) : t \geq 0\}$ is strongly continuous, then for any $\delta > 0$ there is $M_\delta > 0$ such that

$$\|T(t)\| \leq M_\delta, \ \ \forall\, t \in [0, \delta].$$

Proof. (a) Using repeatedly the semigroup property in Definition 11.1.1 we get $T(t) = T(\delta)^{n-1}T(t-(n-1)\delta)$ for $(n-1)\delta \le t \le n\delta$, whence $\|T(t)\| \le M^n \le Me^{\omega t}$. Let $x \in X$ be such that $t \mapsto T(t)x$ is continuous at 0, i.e., $\lim_{h\to 0^+} T(h)x = x$. Using again the semigroup property in Definition 11.1.1 it is easily seen that for every $t > 0$ the equality $\lim_{h\to 0^+} T(t+h)x = T(t)x$ holds. Moreover,

$$\|T(t-h)x - T(t)x\| = \|T(t-h)(x - T(h)x)\| \le Me^{\omega(t-h)}\|(x - T(h)x)\|, \qquad 0 < h < t,$$

whence $\lim_{h\to 0^+} T(t-h)x = T(t)x$. It follows that $t \mapsto T(t)x$ is continuous in $[0,\infty)$.

(b) Let $x \in X$. As $T(\cdot)x$ is continuous, for every $\delta > 0$ there is $M_{\delta,x} > 0$ such that

$$\|T(t)x\| \le M_{\delta,x}, \quad \forall\, t \in [0,\delta].$$

The statement follows from the Uniform Boundedness Principle, see e.g. [4, Chapter 2] or [10, §II.1]. □

If (11.1.1) holds with $M = 1$ and $\omega = 0$ then the semigroups is said *semigroup of contractions* or *contractive semigroup*. From now on, $\{T(t) : t \ge 0\}$ is a fixed strongly continuous semigroup.

Definition 11.1.3. *The* infinitesimal generator *(or, shortly, the* generator*) of the semigroup $\{T(t) : t \ge 0\}$ is the operator defined by*

$$D(L) = \Big\{x \in X : \exists \lim_{h\to 0^+} \frac{T(h)-I}{h}x\Big\}, \quad Lx = \lim_{h\to 0^+} \frac{T(h)-I}{h}x.$$

By definition, the vector Lx is the right derivative of the function $t \mapsto T(t)x$ at $t = 0$ and $D(L)$ is the subspace where this derivative exists. In general, $D(L)$ is not the whole X, but it is dense, as the next proposition shows.

Proposition 11.1.4. *The domain $D(L)$ of the generator is dense in X.*

Proof. Set

$$M_{a,t}x = \frac{1}{t}\int_a^{a+t} T(s)x\, ds, \quad a \ge 0,\ t > 0,\ x \in X$$

(this is an X-valued Bochner integral). As the function $s \mapsto T(s)x$ is continuous, we have (see Exercise 11.1)

$$\lim_{t\to 0} M_{a,t}x = T(a)x.$$

In particular, $\lim_{t\to 0^+} M_{0,t}x = x$ for every $x \in X$. Let us show that for every $t > 0$, $M_{0,t}x \in D(L)$, which implies that the statement holds. We have

$$\begin{aligned}
\frac{T(h)-I}{h}M_{0,t}x &= \frac{1}{ht}\Big(\int_0^t T(h+s)x\, ds - \int_0^t T(s)x\, ds\Big)\\
&= \frac{1}{ht}\Big(\int_h^{h+t} T(s)x\, ds - \int_0^t T(s)x\, ds\Big)\\
&= \frac{1}{ht}\Big(\int_t^{h+t} T(s)x\, ds - \int_0^h T(s)x\, ds\Big)\\
&= \frac{M_{t,h}x - M_{0,h}x}{t}.
\end{aligned}$$

Therefore, for every $x \in X$ we have $M_{0,t}x \in D(L)$ and

$$LM_{0,t}x = \frac{T(t)x - x}{t}. \tag{11.1.2}$$

□

Proposition 11.1.5. *For every $t > 0$, $T(t)$ maps $D(L)$ into itself, and L and $T(t)$ commute on $D(L)$.*

If $x \in D(L)$, then the function $T(\cdot)x$ is differentiable at every $t \geq 0$ and

$$\frac{d}{dt}T(t)x = LT(t)x = T(t)Lx, \quad t \geq 0.$$

Proof. For every $x \in X$ and for every $h > 0$ we have

$$\frac{T(h) - I}{h}T(t)x = T(t)\frac{T(h) - I}{h}x.$$

If $x \in D(L)$, letting $h \to 0$ we obtain $T(t)x \in D(L)$ and $LT(t)x = T(t)Lx$.

Fix $t_0 \geq 0$ and let $h > 0$. We have

$$\frac{T(t_0 + h)x - T(t_0)x}{h} = T(t_0)\frac{T(h) - I}{h}x \to T(t_0)Lx \quad \text{as } h \to 0^+.$$

This shows that $T(\cdot)x$ is right differentiable at t_0. Let us show that it is left differentiable, assuming $t_0 > 0$. If $h \in (0, t_0)$ we have

$$\frac{T(t_0 - h)x - T(t_0)x}{-h} = T(t_0 - h)\frac{T(h) - I}{h}x \to T(t_0)Lx \quad \text{as } h \to 0^+,$$

as

$$\Big\|T(t_0-h)\frac{T(h) - I}{h}x - T(t_0)Lx\Big\| \leq \Big\|T(t_0-h)\Big(\frac{T(h) - I}{h}x - Lx\Big)\Big\| + \|(T(t_0-h) - T(t_0))Lx\|$$

and $\|T(t_0 - h)\| \leq \sup_{0 \leq t \leq t_0} \|T(t)\| < \infty$ by Lemma 11.1.2. It follows that the function $t \mapsto T(t)x$ is differentiable at all $t \geq 0$ and its derivative is $T(t)Lx$, which is equal to $LT(t)x$ by the first part of the proof. □

Using Proposition 11.1.5 we prove that the generator L is a closed operator. Therefore, $D(L)$ is a Banach space with the graph norm $\|x\|_{D(L)} = \|x\| + \|Lx\|$.

Proposition 11.1.6. *The generator L of any strongly continuous semigroup is a closed operator.*

Proof. Let (x_n) be a sequence in $D(L)$, and let $x, y \in X$ be such that $x_n \to x$, $Lx_n =: y_n \to y$. By Proposition 11.1.5 the function $t \mapsto T(t)x_n$ is continuously differentiable in $[0, \infty)$. Hence for $0 < h < 1$ we have (see Exercise 11.1)

$$\frac{T(h) - I}{h}x_n = \frac{1}{h}\int_0^h LT(t)x_n dt = \frac{1}{h}\int_0^h T(t)y_n dt,$$

and then

$$\Big\|\frac{T(h)-I}{h}x-y\Big\| \le \Big\|\frac{T(h)-I}{h}(x-x_n)\Big\| + \Big\|\frac{1}{h}\int_0^h T(t)(y_n-y)dt\Big\| + \Big\|\frac{1}{h}\int_0^h T(t)ydt-y\Big\|$$
$$\le \frac{C+1}{h}\|x-x_n\| + C\|y_n-y\| + \Big\|\frac{1}{h}\int_0^h T(t)ydt-y\Big\|,$$

where $C=\sup_{0<t<1}\|T(t)\|$. Given $\varepsilon>0$, there is h_0 such that for $0<h\le h_0$ we have $\|\int_0^h T(t)ydt/h-y\|\le \varepsilon/3$. For $h\in(0,h_0]$, take n such that $\|x-x_n\|\le \varepsilon h/3(C+1)$ and $\|y_n-y\|\le\varepsilon/3C$: we get $\|\frac{T(h)-I}{h}x-y\|\le\varepsilon$ and therefore $x\in D(L)$ and $y=Lx$, i.e., the operator L is closed. $\square$

Proposition 11.1.5 implies that for any $x\in D(L)$ the function $u(t)=T(t)x$ is differentiable for $t\ge 0$ and it solves the Cauchy problem

$$\begin{cases} u'(t)=Lu(t), & t\ge 0, \\ \\ u(0)=x. \end{cases} \tag{11.1.3}$$

Lemma 11.1.7. *For every $x\in D(L)$, the function $u(t):=T(t)x$ is the* unique *solution of* (11.1.3) *belonging to $C([0,\infty);D(L))\cap C^1([0,\infty);X)$.*

Proof. From Proposition 11.1.5 we know that $u'(t)=T(t)Lx$ for every $t\ge 0$, and then $u'\in C([0,\infty);X)$. Therefore, $u\in C^1([0,\infty);X)$. Since $D(L)$ is endowed with the graph norm, a function $u:[0,\infty)\to D(L)$ is continuous iff both u and Lu are continuous. In our case, both u and $Lu=u'$ belong to $C([0,\infty);X)$, and then $u\in C([0,\infty);D(L))$.

Let us prove that (11.1.3) has a unique solution in $C([0,\infty);D(L))\cap C^1([0,\infty);X)$. If $u\in C([0,\infty);D(L))\cap C^1([0,\infty);X)$ is any solution, we fix $t>0$ and define the function

$$v(s):=T(t-s)u(s), \quad 0\le s\le t.$$

Then (Exercise 11.2) v is differentiable, and $v'(s)=-T(t-s)Lu(s)+T(t-s)u'(s)=0$ for $0\le s\le t$, whence $v(t)=v(0)$, i.e., $u(t)=T(t)x$. $\square$

Remark 11.1.8. If $\{T(t):\ t\ge 0\}$ is a C_0-semigroup with generator L, then for every $\lambda\in\mathbb{C}$ the family of operators

$$S(t)=e^{\lambda t}T(t), \quad t\ge 0,$$

is a C_0-semigroup as well, with generator $L+\lambda I:D(L)\to X$. The semigroup property is obvious. Concerning the generator, for every $x\in X$ we have

$$\frac{S(h)x-x}{h}=e^{\lambda h}\frac{T(h)-x}{h}+\frac{e^{\lambda h}x-x}{h}$$

and then

$$\lim_{h\to 0^+}\frac{S(h)x-x}{h}=\lim_{h\to 0^+}e^{\lambda h}\frac{T(h)-x}{h}+\frac{e^{\lambda h}x-x}{h}=Lx+\lambda x$$

iff $x\in D(L)$.

Let $\{T(t) : \ t \geq 0\}$ be a strongly continuous semigroup. Characterising the domain of its generator L may be difficult. However, for many proofs it is enough to know that "good" elements x are dense in $D(L)$. A subspace $D \subset D(L)$ is called a *core* of L if D is dense in $D(L)$ with respect to the graph norm. The following proposition gives an easily checkable sufficient condition in order that D is a core.

Lemma 11.1.9. *If $D \subset D(L)$ is a dense subspace of X and $T(t)(D) \subset D$ for every $t \geq 0$, then D is a core.*

Proof. Let M, ω be such that $\|T(t)\| \leq Me^{\omega t}$ for every $t > 0$. For $x \in D(L)$ we have

$$Lx = \lim_{t\to 0} \frac{1}{t}\int_0^t T(s)Lx\, ds.$$

Let $(x_n) \subset D$ be a sequence such that $\lim_{n\to\infty} x_n = x$. Set

$$y_{n,t} = \frac{1}{t}\int_0^t T(s)x_n\, ds = \frac{1}{t}\int_0^t T(s)(x_n - x)\, ds + \frac{1}{t}\int_0^t T(s)x\, ds.$$

As the $D(L)$-valued function $s \mapsto T(s)x_n$ is continuous in $[0,\infty)$, the vector $\int_0^t T(s)x_n ds$ belongs to $D(L)$. Moreover, it is the limit of the Riemann sums of elements of D (see Exercise 11.1), hence it belongs to the closure of D in $D(L)$. Therefore, $y_{n,t}$ belongs to the closure of D in $D(L)$ for every n and t. Furthermore,

$$\|y_{n,t} - x\| \leq \left\|\frac{1}{t}\int_0^t T(s)(x_n - x)\, ds\right\| + \left\|\frac{1}{t}\int_0^t T(s)x\, ds - x\right\|$$

tends to 0 as $t \to 0$, $n \to \infty$. By (11.1.2) we have

$$Ly_{n,t} - Lx = \frac{T(t)(x_n - x) - (x_n - x)}{t} + \frac{1}{t}\int_0^t T(s)Lx\, ds - Lx.$$

Given $\varepsilon > 0$, fix $\tau > 0$ such that

$$\left\|\frac{1}{\tau}\int_0^\tau T(s)Lx\, ds - Lx\right\| \leq \varepsilon,$$

and then take $n \in \mathbb{N}$ such that $(Me^{\omega\tau} + 1)\|x_n - x\|/\tau \leq \varepsilon$. Therefore, $\|Ly_{n,\tau} - Lx\| \leq 2\varepsilon$ and the statement follows. □

11.2 Generation Theorems

In this section we recall the main generation theorems for C_0-semigroups. The most general result is the classical Hille–Yosida Theorem, which gives a complete characterisation of the generators. For *contractive* semigroups, i.e., semigroups verifying the estimate $\|T(t)\| \leq 1$ for all $t \geq 0$, the characterisation of the generators provided by the Lumer-Phillips Theorem is often useful. We do not present here the proofs of these results, referring e.g. to [13, §II.3].

First, we recall the definition of *spectrum* and *resolvent*. The natural setting for spectral theory is that of complex Banach spaces, hence if X is real we replace it by its complexification $\widetilde{X} = \{x + iy : \ x, \ y \in X\}$ endowed with the norm

$$\|x + iy\|_{\tilde{X}} := \sup_{-\pi \le \theta \le \pi} \|x \cos\theta + y \sin\theta\|$$

(notice that the seemingly more natural "Euclidean norm" $(\|x\|^2 + \|y\|^2)^{1/2}$ is not a norm in general).

Definition 11.2.1. *Let $L : D(L) \subset X \to X$ be a linear operator. The* resolvent set $\rho(L)$ *and the* spectrum $\sigma(L)$ *of L are defined by*

$$\rho(L) = \{\lambda \in \mathbb{C} : \exists\ (\lambda I - L)^{-1} \in \mathcal{L}(X)\}, \quad \sigma(L) = \mathbb{C} \backslash \rho(L). \tag{11.2.1}$$

The complex numbers $\lambda \in \sigma(L)$ such that $\lambda I - L$ is not injective are the eigenvalues*, and the vectors $x \in D(L)$ such that $Lx = \lambda x$ are the* eigenvectors *(or* eigenfunctions*, when X is a function space). The set $\sigma_p(L)$ whose elements are all the eigenvalues of L is the* point spectrum*.*

For $\lambda \in \rho(L)$, we set

$$R(\lambda, L) := (\lambda I - L)^{-1}. \tag{11.2.2}$$

The operator $R(\lambda, L)$ is the *resolvent operator* or briefly *resolvent*.

We ask to check (Exercise 11.3) that if the resolvent set $\rho(L)$ is not empty, then L is a closed operator. We also ask to check (Exercise 11.4) the following equality, known as the *resolvent identity*

$$R(\lambda, L) - R(\mu, L) = (\mu - \lambda) R(\lambda, L) R(\mu, L), \ \ \forall\ \lambda, \ \mu \in \rho(L). \tag{11.2.3}$$

Theorem 11.2.2 (Hille–Yosida)**.** *The linear operator $L : D(L) \subset X \to X$ is the generator of a C_0-semigroup verifying estimate* (11.1.1) *iff the following conditions hold:*

$$\begin{cases} (i) & D(L) \ \text{ is dense in } \ X, \\ \\ (ii) & \rho(L) \supset \{\lambda \in \mathbb{R} : \ \lambda > \omega\}, \\ \\ (iii) & \|(R(\lambda, L))^k\|_{\mathcal{L}(X)} \le \dfrac{M}{(\lambda - \omega)^k} \ \ \forall\ k \in \mathbb{N}, \ \forall\ \lambda > \omega. \end{cases} \tag{11.2.4}$$

Before stating the Lumer–Phillips Theorem, we define the *dissipative* operators.

Definition 11.2.3. *A linear operator $L : D(L) \subset X \to X$ is called* dissipative *if*

$$\|(\lambda I - L)x\| \ge \lambda \|x\|$$

for all $\lambda > 0$, $x \in D(L)$.

Theorem 11.2.4 (Lumer–Phillips)**.** *A densely defined and dissipative operator L on X is closable and its closure is dissipative. Moreover, the following statements are equivalent.*

(i) The closure of L generates a contraction C_0-semigroup.

(ii) The range of $\lambda I - L$ is dense in X for some (hence all) $\lambda > 0$.

11.3 Invariant measures

In our lectures we shall encounter semigroups defined in L^p spaces, i.e., $X = L^p(\Omega)$ where $(\Omega, \mathscr{F}, \mu)$ is a measure space, with $\mu(\Omega) < \infty$. A property that will play an important role is the conservation of the mean value, namely

$$\int_\Omega T(t)f\,d\mu = \int_\Omega f\,d\mu \qquad \forall\, t > 0,\ \forall f \in L^p(\Omega).$$

In this case μ is called *invariant* for $T(t)$. The following proposition gives an equivalent condition for invariance, in terms of the generator of the semigroup rather than the semigroup itself.

Proposition 11.3.1. *Let $\{T(t) : t \ge 0\}$ be a strongly continuous semigroup with generator L in $L^p(\Omega, \mu)$, where (Ω, μ) is a measure space, $p \in [1, \infty)$, and $\mu(\Omega) < \infty$. Then*

$$\int_\Omega T(t)f\,d\mu = \int_\Omega f\,d\mu\ \ \forall t > 0,\ \forall f \in L^p(\Omega, \mu) \quad \Longleftrightarrow \int_\Omega Lf\,d\mu = 0\ \ \forall f \in D(L).$$

Proof. "$\Rightarrow$" Let $f \in D(L)$. Then $\lim_{t\to 0}(T(t)f - f)/t = Lf$ in $L^p(\Omega, \mu)$ and consequently in $L^1(\Omega, \mu)$. Integrating we obtain

$$\int_\Omega Lf\,d\mu = \lim_{t\to 0} \frac{1}{t} \int_\Omega (T(t)f - f)d\mu = 0.$$

"$\Leftarrow$" Let $f \in D(L)$. Then the function $t \mapsto T(t)f$ belongs to $C^1([0,\infty); L^p(\Omega, \mu))$ and $d/dt\, T(t)f = LT(t)f$, so that for every $t \ge 0$,

$$\frac{d}{dt} \int_X T(t)f\,d\mu = \int_\Omega LT(t)f\,d\mu = 0.$$

Therefore the function $t \mapsto \int_X T(t)f\,d\mu$ is constant, and equal to $\int_X f\,d\mu$. The operator $L^p(\Omega, \mu) \to \mathbb{R}$, $f \mapsto \int_\Omega (T(t)f - f)d\mu$, is bounded and vanishes on the dense subset $D(L)$; hence it vanishes in the whole $L^p(\Omega, \mu)$. $\square$

11.4 Analytic semigroups

We recall now an important class of semigroups, the analytic semigroups generated by sectorial operators. For the definition of sectorial operators we need that X is a complex Banach space.

Definition 11.4.1. *A linear operator $L : D(L) \subset X \to X$ is called* sectorial *if there are $\omega \in \mathbb{R}$, $\theta \in (\pi/2, \pi)$, $M > 0$ such that*

$$\begin{cases} (i) \quad \rho(L) \supset S_{\theta,\omega} := \{\lambda \in \mathbb{C} : \lambda \ne \omega, |\arg(\lambda - \omega)| < \theta\}, \\ \\ (ii) \quad \|R(\lambda, L)\|_{\mathcal{L}(X)} \le \dfrac{M}{|\lambda - \omega|}\ \ \forall \lambda \in S_{\theta,\omega}. \end{cases} \tag{11.4.1}$$

In the literature there are also other notions of sectorial operators, but this is the most popular and the one usueful for us.

Sectorial operators with dense domains are infinitesimal generators of semigroups with noteworthy smoothing properties. The proof of the following theorem may be found in [13, Chapter 2], [21, Chapter 2].

Theorem 11.4.2. *Let L be a sectorial operator with dense domain. Then it is the infinitesimal generator of a semigroup $\{T(t):\ t\geq 0\}$ that enjoys the following properties.*

(i) $T(t)x\in D(L^k)$ *for every* $t>0$, $x\in X$, $k\in\mathbb{N}$.

(ii) There are $M_0,\ M_1,\ M_2,\ \ldots,$ *such that*

$$\begin{cases} (a) \quad \|T(t)\|_{\mathcal{L}(X)}\leq M_0e^{\omega t}, \ \ t>0, \\ \\ (b) \quad \|t^k(L-\omega I)^kT(t)\|_{\mathcal{L}(X)}\leq M_ke^{\omega t}, \ \ t>0, \end{cases} \tag{11.4.2}$$

where ω *is the constant in* (11.4.1).

(iii) The function $t\mapsto T(t)$ *belongs to* $C^\infty((0,+\infty);\mathcal{L}(X))$, *and the equality*

$$\frac{d^k}{dt^k}T(t)=L^kT(t), \ \ t>0, \tag{11.4.3}$$

holds.

(iv) The function $t\mapsto T(t)$ *has a* $\mathcal{L}(X)$*-valued holomorphic extension in a sector* $S_{\beta,0}$ *with* $\beta>0$.

The name "analytic semigroup" comes from property (iv). If $\mathcal{O}$ is an open set in $\mathbb{C}$, and Y is a complex Banach space, a function $f:\mathcal{O}\to Y$ is called holomorphic if it is differentiable at every $z_0\in\mathcal{O}$ in the usual complex sense, i.e. there exists the limit

$$\lim_{z\to z_0}\frac{f(z)-f(z_0)}{z-z_0}=:f'(z_0).$$

As in the scalar case, such functions are infinitely many times differentiable at every $z_0\in\mathcal{O}$, and the Taylor series $\sum_{k=0}^\infty f^{(k)}(z_0)(z-z_0)^k/k!$ converges to $f(z)$ for every z in a neighborhood of z_0.

We do not present the proof of this theorem, because in the case of the Ornstein-Uhlenbeck semigroup that will be discussed in the next lectures we shall provide direct proofs of the relevant properties without relying on the above general results. A more general theory of analytic semigroups, not necessarily strongly continuous at $t=0$, is available, see [21].

11.4.1 Self-adjoint operators in Hilbert spaces

If X is a Hilbert space (inner product $\langle \cdot, \cdot \rangle$, norm $\| \cdot \|$) then we can say more on semigroups and generators in connection to self-adjointness. Notice also that the dissipativity condition can be rephrased in the Hilbert space as follows. An operator $L : D(L) \to X$ is dissipative iff (see Exercise 11.5)

$$\mathrm{Re}\, \langle Lx, x \rangle \le 0, \quad \forall x \in D(L). \tag{11.4.4}$$

Let us prove that any self-adjoint dissipative operator is sectorial.

Proposition 11.4.3. *Let $L : D(L) \subset X \to X$ be a self-adjoint dissipative operator. Then L is sectorial with $\theta < \pi$ arbitrary and $\omega = 0$.*

Proof. Let us first show that the spectrum of L is real. If $\lambda = a + ib \in \mathbb{C}$, for every $x \in D(L)$ we have

$$\|(\lambda I - L)x\|^2 = (a^2 + b^2)\|x\|^2 - 2a\langle x, Lx \rangle + \|Lx\|^2 \ge b^2\|x\|^2, \tag{11.4.5}$$

hence if $b \neq 0$ then $\lambda I - L$ is injective. Let us check that in this case it is also surjective, showing that its range is closed and dense in X. Let $(x_n) \subset D(L)$ be a sequence such that the sequence $(\lambda x_n - Lx_n)$ is convergent. From the inequality

$$\|(\lambda I - L)(x_n - x_m)\|^2 \ge b^2\|x_n - x_m\|^2, \quad n, m \in \mathbb{N},$$

it follows that the sequence (x_n) is a Cauchy sequence, hence (Lx_n) as well. Therefore, there are $x, y \in X$ such that $x_n \to x$ and $Lx_n \to y$. Since L is closed, $x \in D(L)$ and $Lx = y$, hence $\lambda x_n - Lx_n$ converges to $\lambda x - Lx \in \mathrm{rg}\,(\lambda I - L)$ and the range of $\lambda I - L$ is closed.

Let now y be orthogonal to the range of $(\lambda I - L)$. Then, for every $x \in D(L)$ we have $\langle y, \lambda x - Lx \rangle = 0$, whence $y \in D(L^\star) = D(L)$ and $\overline{\lambda} y - L^\star y = \overline{\lambda} y - Ly = 0$. As $\overline{\lambda} I - L$ injective, $y = 0$ follows. Therefore the range of $(\lambda I - L)$ is dense in X.

From the dissipativity of L it follows that the spectrum of L is contained in $(-\infty, 0]$. Indeed, if $\lambda > 0$ then for every $x \in D(L)$ we have, instead of (11.4.5),

$$\|(\lambda I - L)x\|^2 = \lambda^2\|x\|^2 - 2\lambda\langle x, Lx \rangle + \|Lx\|^2 \ge \lambda^2\|x\|^2, \tag{11.4.6}$$

and arguing as above we deeduce $\lambda \in \rho(L)$.

Let us now estimate $\|R(\lambda, L)\|$, for $\lambda = \rho e^{i\theta}$, with $\rho > 0$, $-\pi < \theta < \pi$. For $x \in X$, set $u = R(\lambda, L)x$. Multiplying the equality $\lambda u - Lu = x$ by $e^{-i\theta/2}$ and then taking the inner product with u, we get

$$\rho e^{i\theta/2}\|u\|^2 - e^{-i\theta/2}\langle Lu, u \rangle = e^{-i\theta/2}\langle x, u \rangle,$$

whence, taking the real part,

$$\rho \cos(\theta/2)\|u\|^2 - \cos(\theta/2)\langle Lu, u \rangle = \mathrm{Re}(e^{-i\theta/2}\langle x, u \rangle) \le \|x\|\,\|u\|$$

and then, as $\cos(\theta/2) > 0$, also

$$\|u\| \le \frac{\|x\|}{|\lambda|\cos(\theta/2)},$$

with $\theta = \arg\lambda$. $\square$

Proposition 11.4.4. *Let $\{T(t) :\ t \ge 0\}$ be a C_0-semigroup. The family of operators $\{T(t)^\star :\ t \ge 0\}$ is a C_0-semigroup whose generator is $L^\star$.*

Proof. The semigroup law is immediately checked. Let us prove the strong continuity. Recall that by (11.1.1) we have $\|T(t)\| = \|T(t)^\star\| \le Me^{\omega t}$, where we may assume $\omega > 0$. First, notice that $T(t)^\star x \to x$ weakly for every $x \in X$ as $t \to 0$. Indeed, by the strong continuity of $T(t)$ we have $\langle T(t)^\star x, y\rangle = \langle x, T(t)y\rangle \to \langle x, y\rangle$ as $t \to 0$ for every $y \in X$. From the estimate

$$\Big|\int_0^t \langle T(s)^\star x, y\rangle\, ds\Big| \le \frac{M}{\omega}(e^{\omega t} - 1)\|x\|\,\|y\|$$

and the Riesz Theorem we get the existence of $x_t \in X$ such that

$$\frac{1}{t}\int_0^t \langle T(s)^\star x, y\rangle\, ds = \langle x_t, y\rangle \qquad \forall\ y \in X.$$

Therefore, for $t > 0$ and $0 < h < t$ we infer

$$\begin{aligned}
|\langle T(h)^\star x_t - x_t, y\rangle| &= |\langle x_t, T(h)y\rangle - \langle x_t, y\rangle| \\
&= \Big|\frac{1}{t}\int_0^t \langle T(s)^\star x, T(h)y\rangle\, ds - \frac{1}{t}\int_0^t \langle T(s)^\star x, y\rangle\, ds\Big| \\
&= \Big|\frac{1}{t}\int_0^t \langle T(s+h)^\star x, y\rangle\, ds - \frac{1}{t}\int_0^t \langle T(s)^\star x, y\rangle\, ds\Big| \\
&= \Big|\frac{1}{t}\int_t^{t+h} \langle T(s)^\star x, y\rangle\, ds - \frac{1}{t}\int_0^h \langle T(s)^\star x, y\rangle\, ds\Big| \\
&\le \frac{1}{t}\|x\|\,\|y\|\,\frac{M}{\omega}\big[(e^{\omega(t+h)} - e^{\omega t}) + (e^{\omega h} - 1)\big].
\end{aligned}$$

Taking the supremum on $\|y\| = 1$, we deduce

$$\lim_{h\to 0}\|T(h)^\star x_t - x_t\| = 0. \tag{11.4.7}$$

Set $Y := \{x \in X :\ \lim_{h\to 0}\|T(h)^\star x - x\| = 0$. By an $\varepsilon/3$ argument, it is easily seen that Y is closed. Moreover, it is a subspace of X. Therefore (e.g, [4, Thm. 3.7]), Y is weakly closed. Since for any $x \in X$, $x_t \in Y$ and $x_t \to x$ weakly as $t \to 0$, we conclude that $Y = X$, and consequently $\{T(t)^\star : t \ge 0\}$ is strongly continuous. Denoting by A its generator, for $x \in D(L)$ and $y \in D(A)$ we have

$$\langle Lx, y\rangle = \lim_{t\to 0}\langle t^{-1}(T(t) - I)x, y\rangle = \lim_{t\to 0}\langle x, t^{-1}(T(t)^\star - I)y\rangle = \langle x, Ay\rangle,$$

so that $A \subset L^\star$. Conversely, for $y \in D(L^\star),\ x \in D(L)$ we have

$$\begin{aligned}\langle x, T(t)^\star y - y\rangle = \langle T(t)x - x, y\rangle &= \int_0^t \langle LT(s)x, y\rangle\, ds \\ &= \int_0^t \langle T(s)x, L^\star y\rangle\, ds = \int_0^t \langle x, T(s)^\star L^\star y\rangle\, ds.\end{aligned}$$

We deduce

$$T(t)^\star y - y = \int_0^t T(s)^\star L^\star y\, ds,$$

whence, dividing by t and letting $t \to 0$ we get $Ay = L^\star y$ for every $y \in D(L^\star)$ and consequently $L^\star \subset A$. □

The following result is an immediate consequence of Proposition 11.4.4.

Corollary 11.4.5. *The generator L is self-adjoint if and only if $T(t)$ is self-adjoint for every $t > 0$.*

11.5 Exercises

Exercise 11.1. Let $\mathbb{R}$ be endowed with the Lebesgue measure λ_1, and let $f : [a,b] \to X$ be a continuous function. Prove that it is Bochner integrable, that

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} \sum_{i=1}^n f(\tau_i)\frac{b-a}{n}$$

for any choice of $\tau_i \in \left[a + \frac{(b-a)(i-1)}{n}, a + \frac{(b-a)i}{n}\right]$, $i = 1, \dots, n$ (the sums in this approximation are the usual Riemann sums in the real-valued case) and that, setting

$$F(t) = \int_a^t f(s)ds, \quad a \le t \le b,$$

the function F is continuously differentiable, with

$$F'(t) = f(t), \quad a \le t \le b.$$

Exercise 11.2. Prove that if $u \in C([0,\infty); D(L)) \cap C^1([0,\infty); X)$ is a solution of problem (11.1.3), then for $t > 0$ the function $v(s) = T(t-s)u(s)$ is continuously differentiable in $[0,t]$ and it verifies $v'(s) = -T(t-s)Lu(s) + T(t-s)u'(s) = 0$ for $0 \le s \le t$.

Exercise 11.3. Let $L : D(L) \subset X \to X$ be a linear operator. Prove that if $\rho(L) \neq \emptyset$ then L is closed.

Exercise 11.4. Prove the resolvent identity (11.2.3).

Exercise 11.5. Prove that in Hilbert spaces the dissipativity condition in Definition 11.2.3 is equivalent to (11.4.4).

Exercise 11.6. Let $\{T(t) : \ t \ge 0\}$ be a bounded strongly continuous semigroup. Prove that the norm

$$|x| := \sup_{t\ge 0} \|T(t)x\|$$

is equivalent to $\|\cdot\|$ and that $T(t)$ is contractive on $(X, |\cdot|)$.

Lecture 12

The Ornstein-Uhlenbeck semigroup

All of us know the importance of the Laplacian operator Δ and of the heat semigroup,

$$\Delta f(x) = \sum_{i=1}^{d} \frac{\partial^2 f}{\partial x_i^2}(x), \qquad T(t)f(x) = \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-|x-y|^2/4t} f(y)dy, \quad t > 0,$$

that serve as prototypes for elliptic differential operators and semigroups of operators, respectively. Choosing $X = L^2(\mathbb{R}^d, \lambda_d)$, the Laplacian $\Delta : D(\Delta) = W^{2,2}(\mathbb{R}^d, \lambda_d) \to X$ is the infinitesimal generator of $T(t)$, namely given any $f \in X$, there exists the limit $\lim_{t\to 0^+}(T(t)f - f)/t$ if and only if $f \in W^{2,2}(\mathbb{R}^d, \lambda_d)$, and in this case the limit is Δf. The $W^{2,2}$ norm is equivalent to the graph norm. Moreover, the realization of the Laplacian in X is the operator associated with the quadratic form

$$\mathcal{Q}(u, v) = \int_{\mathbb{R}^d} \nabla u \cdot \nabla v \ dx, \quad u,\ v \in W^{1,2}(\mathbb{R}^d, \lambda_d).$$

This means that $D(\Delta)$ consists precisely of the elements $u \in W^{1,2}(\mathbb{R}^d, \lambda_d)$ such that the function $W^{1,2}(\mathbb{R}^d, \lambda_d) \to \mathbb{R}$, $\varphi \mapsto \int_{\mathbb{R}^d} \nabla u \cdot \nabla \varphi \, dx$ has a linear bounded extension to the whole X, namely there exists $g \in L^2(\mathbb{R}^d, \lambda_d)$ such that

$$\int_{\mathbb{R}^d} \nabla u \cdot \nabla \varphi \ dx = \int_{\mathbb{R}^d} g \, \varphi \, dx, \quad \varphi \in W^{1,2}(\mathbb{R}^d, \lambda_d),$$

and in this case $g = -\Delta u$. If $u \in W^{2,2}(\mathbb{R}^d, \lambda_d)$, the above formula follows just integrating by parts, and it is the basic formula that relates the Laplacian and the gradient. Moreover, for $u \in W^{2,2}(\mathbb{R}^d, \lambda_d)$ we have

$$\Delta u = \operatorname{div} \nabla u,$$

where the divergence div is (minus) the adjoint of the gradient $\nabla : W^{1,2}(\mathbb{R}^d, \lambda_d) \to L^2(\mathbb{R}^d, \lambda_d; \mathbb{R}^d)$, and for a vector field $v \in W^{1,2}(\mathbb{R}^d, \lambda_d; \mathbb{R}^d)$ it is given by $\sum_{i=1}^{d} \partial v_i/\partial x_i$.

In this lecture and in the next ones we introduce the Ornstein–Uhlenbeck operator and the Ornstein–Uhlenbeck semigroup, that play the role of the Laplacian and of the heat semigroup if the Lebesgue measure is replaced by the standard Gaussian measure γ_d, and that have natural extensions to our infinite dimensional setting (X, γ, H). As before, X is a separable Banach space endowed with a centred nondegenerate Gaussian measure γ, and H is the relevant Cameron–Martin space.

12.1 The Ornstein–Uhlenbeck semigroup

In this section we define the Ornstein–Uhlenbeck semigroup; we start by defining it in the space of the bounded continuous functions and then we extend it to $L^p(X, \gamma)$, for every $p \in [1, \infty)$.

The Ornstein–Uhlenbeck semigroup in $C_b(X)$ is defined as follows: $T(0) = I$, and for $t > 0$, $T(t)f$ is defined by the Mehler formula

$$T(t)f(x) := \int_X f(e^{-t}x + \sqrt{1 - e^{-2t}}y)\gamma(dy). \tag{12.1.1}$$

We list some properties of the family of operators $\{T(t) : t \geq 0\}$.

Proposition 12.1.1. *$\{T(t) : t \geq 0\}$ is a contraction semigroup in $C_b(X)$. Moreover, for every $f \in C_b(X)$ we have*

$$\int_X T(t)f \, d\gamma = \int_X f \, d\gamma, \quad t > 0. \tag{12.1.2}$$

Proof. First of all we notice that $|T(t)f(x)| \leq \|f\|_\infty$ for every $x \in X$ and $t \geq 0$. The fact that $T(t)f \in C_b(X)$ follows by Dominated Convergence Theorem. So, $T(t) \in \mathcal{L}(C_b(X))$ and $\|T(t)\|_{\mathcal{L}(C_b(X))} \leq 1$. Taking $f \equiv 1$, we have $T(t)f \equiv 1$ so that $\|T(t)\|_{\mathcal{L}(C_b(X))} = 1$ for every $t \geq 0$.

Let us prove that $\{T(t) : t \geq 0\}$ is a semigroup. For every $f \in C_b(X)$ and $t, s > 0$ we have

$$\begin{aligned}(T(t)(T(s)f))(x) &= \int_X (T(s)f)(e^{-t}x + \sqrt{1 - e^{-2t}}y)\gamma(dy) \\ &= \int_X \int_X f(e^{-s}(e^{-t}x + \sqrt{1 - e^{-2t}}y) + \sqrt{1 - e^{-2s}}z)\gamma(dy)\gamma(dz).\end{aligned}$$

Setting now $\Phi(y, z) = e^{-s}\frac{\sqrt{1-e^{-2t}}}{\sqrt{1-e^{-2t-2s}}}y + \frac{\sqrt{1-e^{-2s}}}{\sqrt{1-e^{-2t-2s}}}z$, and using Proposition 2.2.7(iv), we get

$$\begin{aligned}(T(t)(T(s)f))(x) &= \int_X \int_X f(e^{-s-t}x + \sqrt{1 - e^{-2t-2s}}\Phi(y, z))\gamma(dy)\gamma(dz) \\ &= \int_X f(e^{-t-s}x + \sqrt{1 - e^{-2t-2s}}\xi)((\gamma \otimes \gamma) \circ \Phi^{-1})(d\xi) \\ &= \int_X f(e^{-t-s}x + \sqrt{1 - e^{-2t-2s}}\xi)\gamma(d\xi) \\ &= T(t+s)f(x).\end{aligned}$$

Let us prove that (12.1.2) holds. For any $f \in C_b(X)$ we have

$$\int_X T(t)f\,\gamma(dx) = \int_X \int_X f(e^{-t}x + \sqrt{1-e^{-2t}}y)\gamma(dy)\gamma(dx).$$

Setting $\phi(x,y) = e^{-t}x + \sqrt{1-e^{-2t}}y$, we apply Proposition 2.2.7(iv) with any $\theta \in \mathbb{R}$ such that $e^{-t} = \cos\theta$, $\sqrt{1-e^{-2t}} = \sin\theta$, and we get

$$\int_X T(t)f\,d\gamma = \int_X f(\xi)(\gamma\otimes\gamma)\circ\phi^{-1}(d\xi) = \int_X f(\xi)\gamma(d\xi).$$

□

We point out that the semigroup $\{T(t) : t \geq 0\}$ is not strongly continuous in $C_b(X)$, and not even in the subspace $BUC(X)$ of the bounded and uniformly continuous functions. In fact, we have the following characterisation.

Lemma 12.1.2. *Let $f \in BUC(X)$. Then*

$$\lim_{t\to 0^+} \|T(t)f - f\|_\infty = 0 \Longleftrightarrow \lim_{t\to 0^+} \|f(e^{-t}\cdot) - f\|_\infty = 0.$$

Proof. For every $t > 0$ and $x \in X$ we have

$$(T(t)f - f(e^{-t}\cdot))(x) = \int_X (f(e^{-t}x + \sqrt{1-e^{-2t}}y) - f(e^{-t}x))\gamma(dy)$$

and the right hand side goes to 0, uniformly in X, as $t \to 0^+$. Indeed, given $\varepsilon > 0$ fix $R > 0$ such that $\gamma(X \setminus B(0,R)) \leq \varepsilon$, and fix $\delta > 0$ such that $|f(u) - f(v)| \leq \varepsilon$ for $\|u - v\| \leq \delta$. Then, for every t such that $\sqrt{1-e^{-2t}}R \leq \delta$ and for every $x \in X$ we have

$$\begin{aligned}
&\left| \int_X (f(e^{-t}x + \sqrt{1-e^{-2t}}y) - f(e^{-t}x))\gamma(dy) \right| \\
&= \left| \left(\int_{B(0,R)} + \int_{X\setminus B(0,R)} \right)(f(e^{-t}x + \sqrt{1-e^{-2t}}y) - f(e^{-t}x))\gamma(dy) \right| \\
&\leq \varepsilon + 2\|f\|_\infty\varepsilon.
\end{aligned}$$

□

The simplest counterexample to strong continuity in $BUC(X)$ is one dimensional: see Exercise 12.1.

However, for every $f \in C_b(X)$ the function $(t,x) \mapsto T(t)f(x)$ is continuous in $[0,\infty)\times X$ by the Dominated Convergence Theorem. In particular,

$$\lim_{t\to 0^+} T(t)f(x) = f(x), \quad \forall x \in X,$$

which is enough for many purposes.

The semigroup $\{T(t) : t \geq 0\}$ enjoys important smoothing and summability improving properties. The first smoothing property is in the next proposition.

Proposition 12.1.3. *For every $f \in C_b(X)$ and $t > 0$, $T(t)f$ is H-differentiable at every $x \in X$, and we have*

$$[\nabla_H T(t)f(x), h]_H = \frac{e^{-t}}{\sqrt{1-e^{-2t}}} \int_X f(e^{-t}x + \sqrt{1-e^{-2t}}y)\hat{h}(y)\gamma(dy). \tag{12.1.3}$$

Therefore,

$$|\nabla_H T(t)f(x)|_H \leq \frac{e^{-t}}{\sqrt{1-e^{-2t}}}\|f\|_\infty, \quad x \in X. \tag{12.1.4}$$

Proof. Set

$$c(t) := \frac{e^{-t}}{\sqrt{1-e^{-2t}}}.$$

For every $h \in H$ we have

$$\begin{aligned} T(t)f(x+h) &= \int_X f(e^{-t}x + \sqrt{1-e^{-2t}}(c(t)h + y))\gamma(dy) \\ &= \int_X f(e^{-t}x + \sqrt{1-e^{-2t}}z) \exp\Big\{c(t)\hat{h}(z) - c(t)^2\frac{|h|_H^2}{2}\Big\}\, \gamma(dz), \end{aligned}$$

by the Cameron–Martin formula. Therefore, denoting by $l_{t,x}(h)$ the right hand side of (12.1.3),

$$\begin{aligned} &\frac{|T(t)f(x+h) - T(t)f(x) - l_{t,x}(h)|}{|h|_H} \leq \\ &\leq \frac{1}{|h|_H}\int_X \Big| f(e^{-t}x + \sqrt{1-e^{-2t}}y)\Big(\exp\Big\{c(t)\hat{h}(y) - c(t)^2\frac{|h|_H^2}{2}\Big\} - 1 - c(t)\hat{h}(y)\Big)\Big|\gamma(dy) \\ &\leq \frac{\|f\|_\infty}{|h|_H}\int_{\mathbb{R}} \Big|\exp\Big\{c(t)\xi - c(t)^2\frac{|h|_H^2}{2}\Big\} - 1 - c(t)\xi\Big|\mathscr{N}(0,|h|_H^2)(d\xi) \\ &= \|f\|_\infty \int_{\mathbb{R}} \Big|\exp\Big\{c(t)|h|_H\eta - c(t)^2\frac{|h|_H^2}{2}\Big\} - 1 - c(t)|h|_H\eta\Big|\mathscr{N}(0,1)(d\eta) \end{aligned}$$

where the right hand side vanishes as $|h|_H \to 0$, by the Dominated Convergence Theorem. This proves (12.1.3). In its turn, (12.1.3) yields

$$|[\nabla_H T(t)f(x), h]_H| \leq c(t)\|f\|_\infty\|\hat{h}\|_{L^1(X,\gamma)} \leq c(t)\|f\|_\infty\|\hat{h}\|_{L^2(X,\gamma)} = c(t)\|f\|_\infty|h|_H$$

for every $h \in H$, and (12.1.4) follows. □

Notice that in the case $X = \mathbb{R}^d$, $\gamma = \gamma_d$, we have $\nabla_H = \nabla$ and formula (12.1.3) reads as

$$D_i T(t)f(x) = \frac{e^{-t}}{\sqrt{1-e^{-2t}}}\int_{\mathbb{R}^d} f(e^{-t}x + \sqrt{1-e^{-2t}}y)y_i\, \gamma_d(dy), \quad i = 1, \ldots, d. \tag{12.1.5}$$

Let us consider now more regular functions f.

Proposition 12.1.4. *For every $f \in C_b^1(X)$, $T(t)f \in C_b^1(X)$ for any $t \geq 0$, and its derivative at x is*

$$(T(t)f)'(x)(v) = e^{-t}\int_X f'(e^{-t}x + \sqrt{1-e^{-2t}}y)(v)\,\gamma(dy). \tag{12.1.6}$$

In particular,

$$\frac{\partial}{\partial v}T(t)f(x) = e^{-t}\Big(T(t)\Big(\frac{\partial}{\partial v}f\Big)\Big)(x), \quad v \in X,\ x \in X. \tag{12.1.7}$$

For every $f \in C_b^2(X)$, $T(t)f \in C_b^2(X)$ for any $t \geq 0$, and its second order derivative at x is

$$(T(t)f)''(x)(u)(v) = e^{-2t}\int_X f''(e^{-t}x + \sqrt{1-e^{-2t}}y)(u)(v)\,\gamma(dy), \tag{12.1.8}$$

so that

$$\frac{\partial^2 T(t)f}{\partial u \partial v}(x) = e^{-2t}T(t)\bigg(\frac{\partial^2 f}{\partial u \partial v}\bigg)(x), \quad u,\ v \in X,\ x \in X. \tag{12.1.9}$$

Proof. Fix $t > 0$ and set $\Phi(x, y) = e^{-t}x + \sqrt{1-e^{-2t}}y$. For every x, $v \in X$ we have

$$\begin{aligned}
&\bigg|(T(t)f)(x+v) - (T(t)f)(x) - e^{-t}\int_X f'(e^{-t}x + \sqrt{1-e^{-2t}}y)(v)\,\gamma(dy)\bigg|\frac{1}{\|v\|} \\
&\leq \int_X |f(\Phi(x,y) + e^{-t}v) - f(\Phi(x,y)) - f'(\Phi(x,y))(e^{-t}v)|\,\gamma(dy)\frac{1}{\|v\|}.
\end{aligned}$$

On the other hand, for every $y \in X$ we have

$$\lim_{v \to 0} \frac{f(\Phi(x,y) + e^{-t}v) - f(\Phi(x,y)) - f'(\Phi(x,y))(e^{-t}v)}{\|v\|} = 0,$$

and (see Exercise 12.3)

$$\frac{|f(\Phi(x,y) + e^{-t}v) - f(\Phi(x,y)) - f'(\Phi(x,y))(e^{-t}v)|}{\|v\|} \leq 2e^{-t}\sup_{z \in X}\|f'(z)\|_{X^*},$$

and (12.1.6) follows by the Dominated Convergence Theorem. Formula (12.1.7) is an immediate consequence. The derivative $(T(t)f)'$ is continuous still by the Dominated Convergence Theorem. The verification of (12.1.8) and (12.1.9) for $f \in C_b^2(X)$ follow iterating this procedure (see Exercise 12.4). □

Let us compare Proposition 12.1.3 and Proposition 12.1.4. Proposition 12.1.3 describes a smoothing property of $T(t)$, while Proposition 12.1.4 says that $T(t)$ preserves the spaces $C_b^1(X)$ and $C_b^2(X)$. In general, $T(t)$ regularises only along H and it does not map $C_b(X)$ into $C^1(X)$. If $X = \mathbb{R}^d$ and $\gamma = \gamma_d$ we have $H = \mathbb{R}^d$, this difficulty does not arise, Proposition 12.1.3 says that $T(t)$ maps $C_b(\mathbb{R}^d)$ into $C_b^1(\mathbb{R}^d)$ and in fact one can check that $T(t)$ maps $C_b(\mathbb{R}^d)$ into $C_b^k(\mathbb{R}^d)$ for every $k \in \mathbb{N}$, as we shall see in the next lecture.

If $f \in C^1_b(X)$ we can write $\nabla_H T(t)f(x)$ in two different ways, using (12.1.3) and (12.1.6): for every $h \in H$ we have

$$\begin{aligned}[\nabla_H T(t)f(x), h]_H &= \frac{e^{-t}}{\sqrt{1-e^{-2t}}}\int_X f(e^{-t}x+\sqrt{1-e^{-2t}}y)\hat{h}(y)\gamma(dy)\\ &= e^{-t}\int_X f'(e^{-t}x+\sqrt{1-e^{-2t}}y)(h)\,\gamma(dy).\end{aligned}$$

We now extend $T(t)$ to $L^p(X,\gamma)$, $1 \le p < \infty$.

Proposition 12.1.5. *Let $t \ge 0$. For every $f \in C_b(X)$ and $p \in [1,\infty)$ we have*

$$\|T(t)f\|_{L^p(X,\gamma)} \le \|f\|_{L^p(X,\gamma)}. \tag{12.1.10}$$

Hence $\{T(t):\ t \ge 0\}$ is uniquely extendable to a contraction semigroup $\{T_p(t):\ t \ge 0\}$ in $L^p(X,\gamma)$. Moreover

(i) $\{T_p(t):\ t \ge 0\}$ is strongly continuous in $L^p(X,\gamma)$, for every $p \in [1,\infty)$;

(ii) $T_2(t)$ is self–adjoint and nonnegative in $L^2(X,\gamma)$ for every $t > 0$;

(iii) γ is an invariant measure for $\{T_p(t):\ t \ge 0\}$.

Proof. For every $f \in C_b(X)$, $t > 0$ and $x \in X$ the Hölder inequality yields

$$|T(t)f(x)|^p \le \int_X |f(e^{-t}x+\sqrt{1-e^{-2t}}y)|^p d\gamma = T(t)(|f|^p)(x).$$

Integrating over X and using (12.1.2) we obtain

$$\int_X |T(t)f|^p d\gamma \le \int_X T(t)(|f|^p)\,d\gamma = \int_X |f|^p d\gamma.$$

Since $C_b(X)$ is dense in $L^p(X,\gamma)$, $T(t)$ has a unique bounded extension $T_p(t)$ to the whole $L^p(X,\gamma)$, such that $\|T_p(t)\|_{\mathcal{L}(L^p(X,\gamma))} \le 1$. In fact, taking $f \equiv 1$, $T_p(t)f \equiv 1$ so that $\|T_p(t)\|_{\mathcal{L}(L^p(X,\gamma))} = 1$.

Let us prove that $\{T_p(t):\ t \ge 0\}$ is strongly continuous. We already know that for $f \in C_b(X)$ we have $\lim_{t\to 0^+} T(t)f(x) = f(x)$ for every $x \in X$, and moreover $|T(t)f(x)| \le \|f\|_\infty$ for every x. By the Dominated Convergence Theorem, $\lim_{t\to 0^+} T(t)f = f$ in $L^p(X,\gamma)$. Since $C_b(X)$ is dense in $L^p(X,\gamma)$ and $\|T_p(t)\|_{\mathcal{L}(L^p(X,\gamma))} = 1$ for every t, $\lim_{t\to 0^+} T_p(t)f = f$ for every $f \in L^p(X,\gamma)$.

Let us prove statement (ii). Let f, $g \in C_b(X)$, $t > 0$. Then

$$\langle T(t)f, g\rangle_{L^2(X,\gamma)} = \int_X\int_X f(e^{-t}x+\sqrt{1-e^{-2t}}y)g(x)\gamma(dy)\gamma(dx)$$

and setting $u = e^{-t}x + \sqrt{1-e^{-2t}}y$, $v = -\sqrt{1-e^{-2t}}x + e^{-t}y$, $(u,v) =: R(x,y)$, using Proposition 2.2.7(iii)

$$\begin{aligned}
\langle T(t)f, g\rangle_{L^2(X,\gamma)} &= \int_X\int_X f(e^{-t}u + \sqrt{1-e^{-2t}}v)g(u)\gamma(dv)\gamma(du) \\
&= \int_{X\times X} f(u)g(e^{-t}u - \sqrt{1-e^{-2t}}v)((\gamma\otimes\gamma)\circ R^{-1})(d(u,v)) \\
&= \int_X\int_X f(u)g(e^{-t}u - \sqrt{1-e^{-2t}}v)\gamma(dv)\gamma(du) \\
&= \int_X\int_X f(u)g(e^{-t}u + \sqrt{1-e^{-2t}}v)\gamma(dv)\gamma(du) \\
&= \langle f, T(t)g\rangle_{L^2(X,\gamma)}.,
\end{aligned}$$

where in the second to last equality we have used the fact that γ is centred. Approximating any f, $g \in L^2(X,\gamma)$ by elements of $C_b(X)$, we obtain $\langle T_2(t)f, g\rangle_{L^2(X,\gamma)} = \langle f, T_2(t)g\rangle_{L^2(X,\gamma)}$.

Still for every $f \in L^2(X,\gamma)$ and $t > 0$ we have

$$\langle T_2(t)f, f\rangle_{L^2(X,\gamma)} = \langle T_2(t/2)T_2(t/2)f, f\rangle_{L^2(X,\gamma)} = \langle T_2(t/2)f, T_2(t/2)f\rangle_{L^2(X,\gamma)} \geq 0.$$

Statement (iii) is an immediate consequence of (12.1.2). □

To simplify some statements we extend the Ornstein–Uhlenbeck semigroup $T(t)$ to H–valued functions. For $v \in C_b(X;H)$ and $t > 0$ we set

$$T(t)v(x) = \int_X v(e^{-t}x + \sqrt{1-e^{-2t}}y)\gamma(dy), \quad t > 0,\ x \in X.$$

By Remark 9.2.6, for every orthonormal basis $\{h_i : i \in \mathbb{N}\}$ of H we have

$$T(t)v(x) = \sum_{i=1}^{\infty} T(t)([v(\cdot), h_i]_H)(x)h_i.$$

Using Proposition 9.2.5(i) we get the estimate

$$|T(t)v(x)|_H \leq \int_X |v(e^{-t}x + \sqrt{1-e^{-2t}}y)|_H\,\gamma(dy) = (T(t)(|v|_H))(x), \quad x \in X. \tag{12.1.11}$$

Notice that the right hand side concerns the scalar valued function $|v|_H$. Raising to the power p, integrating over X and recalling (12.1.2), we obtain

$$\|T(t)v\|_{L^p(X,\gamma;H)} \leq \|v\|_{L^p(X,\gamma;H)}, \quad t \geq 0. \tag{12.1.12}$$

As in the case of real valued functions, since $C_b(X;H)$ is dense in $L^p(X,\gamma;H)$, estimate (12.1.12) allows to extend $T(t)$ to a bounded (contraction) operator in $L^p(X,\gamma;H)$, called $T_p(t)$. We will not develop the theory for H–valued functions, but we shall use this notion to write some formulae in a concise way, see e.g. (12.1.13).

L^p gradient estimates for $T_p(t)f$ are provided by the next proposition.

Proposition 12.1.6. *Let $1 \le p < \infty$.*

(i) For every $f \in W^{1,p}(X,\gamma)$ and $t > 0$, $T_p(t)f \in W^{1,p}(X,\gamma)$ and

$$\nabla_H T_p(t)f = e^{-t} T_p(t)(\nabla_H f), \tag{12.1.13}$$

$$\|T_p(t)f\|_{W^{1,p}(X,\gamma)} \le \|f\|_{W^{1,p}(X,\gamma)}. \tag{12.1.14}$$

(ii) If $p > 1$, for every $f \in L^p(X,\gamma)$ and $t > 0$, $T_p(t)f \in W^{1,p}(X,\gamma)$ and

$$\int_X |\nabla_H T_p(t)f(x)|_H^p d\gamma \le c(t,p)^p \int_X |f|^p d\gamma, \tag{12.1.15}$$

with $c(t,p) = \left(\int_{\mathbb{R}} |\xi|^{p'} \mathscr{N}(0,1)(d\xi)\right)^{1/p'} e^{-t}/\sqrt{1-e^{-2t}}$.

Proof. (i). Let $f \in C_b^1(X)$. By Proposition 12.1.4, $T(t)f \in C_b^1(X)$ and $(\partial_h T(t)f)(x) = e^{-t}(T(t)(\partial_h f))(x)$ for every $h \in H$, namely $[(\nabla_H T(t)f)(x), h]_H = e^{-t}T(t)([\nabla_H f, h]_H)(x)$. Therefore, $|(\nabla_H T(t)f)(x)|_H \le e^{-t}T(t)(|\nabla_H f|_H)(x)$, for every $x \in X$. Consequently,

$$|\nabla_H T(t)f(x)|_H^p \le e^{-tp}(T(t)(|\nabla_H f|_H)(x))^p \le e^{-tp}(T(t)(|\nabla_H f|_H^p)(x))$$

and integrating we obtain

$$\int_X |\nabla_H T(t)f(x)|_H^p d\gamma \le e^{-tp}\int_X T(t)(|\nabla_H f|_H)^p d\gamma = e^{-tp}\int_X |\nabla_H f|_H^p d\gamma,$$

so that

$$\begin{aligned}\|T(t)f\|_{W^{1,p}(X,\gamma)} &= \|T(t)f\|_{L^p(X,\gamma)} + \|\,|\nabla_H T(t)f|_H\,\|_{L^p(X,\gamma)} \\ &\le \|f\|_{L^p(X,\gamma)} + \|\,|\nabla_H f|_H\,\|_{L^p(X,\gamma)} = \|f\|_{W^{1,p}(X,\gamma)}.\end{aligned}$$

Since $C_b^1(X)$ is dense in $W^{1,p}(X,\gamma)$, (12.1.14) follows.

(ii) Let $f \in C_b(X)$. By Proposition 12.1.3, $T(t)f$ is H-differentiable at every x, and we have

$$|\nabla_H T(t)f(x)|_H = \sup_{h\in H,\,|h|_H=1} |[\nabla_H T(t)f(x), h]_H|.$$

Let us estimate $|[\nabla_H T(t)f(x), h]_H| = |l_{t,x}(h)|$ (where $l_{t,x}(h)$ is as in the proof of Proposition 12.1.3), for $|h|_H = 1$, using formula (12.1.3). We have

$$\begin{aligned}|l_{t,x}(h)| &\le \frac{e^{-t}}{\sqrt{1-e^{-2t}}}\int_X |f(e^{-t}x + \sqrt{1-e^{-2t}}y)|\hat{h}(y)\gamma(dy) \\ &\le \frac{e^{-t}}{\sqrt{1-e^{-2t}}}\left(\int_X |f(e^{-t}x + \sqrt{1-e^{-2t}}y)|^p\gamma(dy)\right)^{1/p}\left(\int_X |\hat{h}|^{p'}d\gamma\right)^{1/p'} \\ &= \frac{e^{-t}}{\sqrt{1-e^{-2t}}}(T(t)(|f|^p)(x))^{1/p}\left(\int_{\mathbb{R}} |\xi|^{p'}\mathscr{N}(0,1)(d\xi)\right)^{1/p'}.\end{aligned}$$

By the invariance property (12.1.2) of γ,

$$\begin{aligned}\int_X |\nabla_H T(t)f(x)|_H^p\,\gamma(dx) &\le \left(\frac{e^{-t}}{\sqrt{1-e^{-2t}}}\right)^p \int_X (T(t)(|f|^p)\,d\gamma\left(\int_{\mathbb{R}} |\xi|^{p'}\mathscr{N}(0,1)(d\xi)\right)^{p/p'}\\ &= \left(\frac{e^{-t}}{\sqrt{1-e^{-2t}}}\right)^p \int_X |f|^p d\gamma\left(\int_{\mathbb{R}} |\xi|^{p'}\mathscr{N}(0,1)(d\xi)\right)^{p/p'}.\end{aligned}$$

Therefore, $T(t)f \in W^{1,p}(X,\gamma)$ and estimate (12.1.15) holds for every $f \in C_b(X)$. Since $C_b(X)$ is dense in $L^p(X,\gamma)$, the statement follows. □

Note that the proof of (ii) fails for $p = 1$, since the function $\hat{h}$ does not belong to L^∞ for every $h \in H$, and the constant $c(t,p)$ in estimate (12.1.15) blows up as $p \to 1$. Indeed, $T(t)$ does not map $L^1(X,\gamma)$ into $W^{1,1}(X,\gamma)$ for $t > 0$, even in the simplest case $X = \mathbb{R}$, $\gamma = \gamma_1$ (see for instance [22, Corollary 5.1]).

12.2 Exercises

Exercise 12.1. Let $X = \mathbb{R}$ and set $f(x) = \sin x$. Prove that $T(t)f$ does not converge uniformly to f in $\mathbb{R}$ as $t \to 0$.

Exercise 12.2. Show that the argument used in Proposition 12.1.5 to prove that $T(t)$ is self–adjoint in $L^2(X,\gamma)$ implies that $T_{p'}(t) = T_p(t)^*$ for $p \in (1,\infty)$ with $1/p + 1/p' = 1$.

Exercise 12.3. Prove that for every $f \in C^1(X)$ and for every x, $y \in X$ we have

$$f(y) - f(x) = \int_0^1 f'(\sigma y + (1-\sigma)x)(y-x)\,d\sigma,$$

so that, if $f \in C_b^1(X)$,

$$|f(y) - f(x)| \le \sup_{z\in X} \|f'(z)\|_{X'}\|y-x\|.$$

Exercise 12.4. Prove that for every $f \in C_b^2(X)$ and $t > 0$, $T(t)f \in C_b^2(X)$ and (12.1.8), (12.1.9) hold.

Lecture 13

The Ornstein-Uhlenbeck operator

In this lecture we study the infinitesimal generator of $T_p(t)$, for $p \in [1, \infty)$. The strongest result is the characterisation of the domain of the generator L_2 of $T_2(t)$ as the Sobolev space $W^{2,2}(X, \gamma)$. A similar result holds for $p \in (1, \infty) \setminus \{2\}$, but the proof is much more complicated and will not be given here.

13.1 The finite dimensional case

Here, $X = \mathbb{R}^d$ and $\gamma = \gamma_d$. We describe the infinitesimal generator L_p of $T_p(t)$ in $L^p(\mathbb{R}^d, \gamma_d)$, for $p \in (1, \infty)$, which is a suitable realisation of the Ornstein-Uhlenbeck differential operator

$$\mathcal{L}f(x) := \Delta f(x) - x \cdot \nabla f(x) \tag{13.1.1}$$

in $L^p(\mathbb{R}^d, \gamma_d)$.

We recall that

$$D(L_p) = \left\{ f \in L^p(\mathbb{R}^d, \gamma_d) : \exists L^p - \lim_{t \to 0^+} \frac{T(t)f - f}{t} \right\},$$
$$L_p f = \lim_{t \to 0^+} \frac{T(t)f - f}{t}.$$

If $f \in D(L_p)$, by Lemma 11.1.7 the function $t \mapsto T(t)f$ belongs to $C^1([0, \infty); L^p(\mathbb{R}^d, \gamma_d))$ $\cap C([0, \infty); D(L_p))$ and $d/dt\, T(t)f = L_p T(t)f$, for every $t \geq 0$. To find an expression of L_p, we differentiate $T(t)f$ with respect to time for good f. We recall that for $f \in C_b(\mathbb{R}^d)$, $T_p(t)f = T(t)f$ is given by formula (12.1.1).

Lemma 13.1.1. *For every $f \in C_b(\mathbb{R}^d)$, the function $(t, x) \mapsto T(t)f(x)$ is smooth in $(0, \infty) \times \mathbb{R}^d$, and we have*

$$\frac{d}{dt}(T(t)f)(x) = \Delta T(t)f(x) - x \cdot \nabla T(t)f(x), \quad t > 0,\ x \in \mathbb{R}^d. \tag{13.1.2}$$

If $f \in C_b^2(\mathbb{R}^d)$, for every $x \in \mathbb{R}^d$ the function $t \mapsto T(t)f(x)$ is differentiable also at $t = 0$, with

$$\frac{d}{dt}(T(t)f)(x)_{|t=0} = \Delta f(x) - x \cdot \nabla f(x), \quad x \in \mathbb{R}^d, \tag{13.1.3}$$

and the function $(t, x) \mapsto d/dt\,(T(t)f)(x)$ is continuous in $[0, \infty) \times \mathbb{R}^d$.

Proof. Setting $z = e^{-t}x + \sqrt{1 - e^{-2t}}y$ in (12.1.1) we see that $(t, x) \mapsto T(t)f(x)$ is smooth in $(0, \infty) \times \mathbb{R}^d$, and that

$$\begin{aligned}
\frac{d}{dt}(T(t)f)(x) &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(z) \frac{\partial}{\partial t}\bigg(\exp\Big\{-\frac{|z - e^{-t}x|^2}{2(1 - e^{-2t})}\Big\}(1 - e^{-2t})^{-d/2}\bigg) dz \\
&= \frac{(1 - e^{-2t})^{-d/2}}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(z) \exp\Big\{-\frac{|z - e^{-t}x|^2}{2(1 - e^{-2t})}\Big\} \\
&\qquad \bigg(-\frac{e^{-t}(z - e^{-t}x) \cdot x}{1 - e^{-2t}} + \frac{e^{-2t}|z - e^{-t}x|^2}{(1 - e^{-2t})^2} - \frac{de^{-2t}}{1 - e^{-2t}} \bigg) dz \\
&= \int_{\mathbb{R}^d} f(e^{-t}x + \sqrt{1 - e^{-2t}}y)(-c(t)\, y \cdot x + c(t)^2|y|^2 - dc(t)^2)\gamma_d(dy),
\end{aligned}$$

where $c(t) = e^{-t}/\sqrt{1 - e^{-2t}}$. Differentiating twice with respect to x in (12.1.1) (recall (12.1.5)), we obtain

$$D_{ij}(T(t)f)(x) = c(t)^2 \int_{\mathbb{R}^d} f(e^{-t}x + \sqrt{1 - e^{-2t}}y)(-\delta_{ij} + y_i y_j)\gamma_d(dy)$$

so that

$$\Delta T(t)f(x) = c(t)^2 \int_{\mathbb{R}^d} f(e^{-t}x + \sqrt{1 - e^{-2t}}y)(-d + |y|^2)\gamma_d(dy).$$

Therefore,

$$\begin{aligned}
\frac{d}{dt}(T(t)f)(x) - \Delta T(t)f(x) &= -c(t) \int_{\mathbb{R}^d} f(e^{-t}x + \sqrt{1 - e^{-2t}}y)\, x \cdot y\, \gamma_d(dy) \\
&= -\nabla T(t)f(x) \cdot x,
\end{aligned}$$

and (13.1.2) follows.

For $f \in C_b^2(\mathbb{R}^d)$, we rewrite formula (13.1.2) as

$$\begin{aligned}
\frac{d}{dt}(T(t)f)(x) &= (\mathcal{L}T(t)f)(x) \\
&= e^{-2t}(T(t)\Delta f)(x) - e^{-t}x \cdot (T(t)\nabla f)(x), \quad t > 0,\ x \in \mathbb{R}^d,
\end{aligned} \tag{13.1.4}$$

taking into account (12.1.7) and (12.1.9). (We recall that $(T(t)\nabla f)(x)$ is the vector whose j-th component is $T(t)D_j f(x)$). Since Δf and each $D_j f$ are continuous and bounded in $\mathbb{R}^d$, the right hand side is continuous in $[0, \infty) \times \mathbb{R}^d$. So, for every $x \in \mathbb{R}^d$ the function $\theta(t) := T(t)f(x)$ is continuous in $[0, \infty)$, it is differentiable in $(0, \infty)$ and $\lim_{t \to 0} \theta'(t) = \Delta f(x) - x \cdot \nabla f(x)$. Therefore, θ is differentiable at 0 too, (13.1.4) holds also at $t = 0$, and (13.1.3) follows. $\square$

Lemma 13.1.1 suggests that L_p is a suitable realisation of the Ornstein-Uhlenbeck differential operator $\mathcal{L}$ defined in (13.1.1). For a first characterisation of L_p we use Lemma 11.1.9.

Proposition 13.1.2. *For $1 \le p < \infty$ and $k \in \mathbb{N}$, $k \ge 2$, the operator $\mathcal{L} : D(\mathcal{L}) = C_b^k(\mathbb{R}^d) \subset L^p(\mathbb{R}^d, \gamma_d) \to L^p(\mathbb{R}^d, \gamma_d)$ is closable, and its closure is L_p. So, $D(L_p)$ consists of all $f \in L^p(\mathbb{R}^d, \gamma_d)$ for which there exists a sequence of functions $(f_n) \subset C_b^k(\mathbb{R}^d)$ such that $f_n \to f$ in $L^p(\mathbb{R}^d, \gamma_d)$ and $(\mathcal{L}f_n)$ converges in $L^p(\mathbb{R}^d, \gamma_d)$. In this case, $L_p f = L^p - \lim_{n\to\infty} \mathcal{L}f_n$.*

Proof. We check that $D = C_b^k(\mathbb{R}^d)$ satisfies the assumptions of Proposition 11.1.9, i.e., it is a core of L_p. We already know, from Proposition 12.1.4, that $T(t)$ maps $C_b^k(\mathbb{R}^d)$ into itself for $k = 1, 2$. The proof of the fact that $C_b^k(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d, \gamma_d)$ and that $T(t)$ maps $C_b^k(\mathbb{R}^d)$ into itself for $k \ge 3$ is left as Exercise 13.2.

Since $C_b^k(\mathbb{R}^d) \subset C_b^2(\mathbb{R}^d)$ for $k \ge 2$, it remains to prove that $C_b^2(\mathbb{R}^d) \subset D(L_p)$, and $L_p f = \mathcal{L}f$ for every $f \in C_b^2(\mathbb{R}^d)$.

By Lemma 13.1.1, for every $f \in C_b^2(\mathbb{R}^d)$ we have $d/dt\, T(t)f(x) = \mathcal{L}T(t)f(x)$ for every $x \in \mathbb{R}^d$ and $t \ge 0$; moreover $t \mapsto d/dt\, T(t)f(x)$ is continuous for every x. Therefore for every $t > 0$ we have

$$\frac{T(t)f(x) - f(x)}{t} = \frac{1}{t}\int_0^t \frac{d}{ds}T(s)f(x)\,ds = \frac{1}{t}\int_0^t \mathcal{L}T(s)f(x)\,ds,$$

and

$$\int_{\mathbb{R}^d} \left|\frac{T(t)f(x) - f(x)}{t} - \mathcal{L}f(x)\right|^p \gamma_d(dx) \le \int_{\mathbb{R}^d} \left(\frac{1}{t}\int_0^t |\mathcal{L}T(s)f(x) - \mathcal{L}f(x)|\,ds\right)^p \gamma_d(dx).$$

Since $s \mapsto \mathcal{L}T(s)f(x)$ is continuous, for every x we have

$$\lim_{t\to 0^+} \left(\frac{1}{t}\int_0^t |\mathcal{L}T(s)f(x) - \mathcal{L}f(x)|\,ds\right)^p = 0.$$

Moreover, by (13.1.4),

$$\left(\frac{1}{t}\int_0^t |\mathcal{L}T(s)f(x) - \mathcal{L}f(x)|\,ds\right)^p \le 2^p(\|\Delta f\|_\infty + |x|\,\|\,|\nabla f|\,\|_\infty)^p \in L^1(\mathbb{R}^d, \gamma_d).$$

By the Dominated Convergence Theorem,

$$\lim_{t\to 0^+} \int_{\mathbb{R}^d} \left|\frac{T(t)f(x) - f(x)}{t} - \mathcal{L}f(x)\right|^p \gamma_d(dx) = 0.$$

Then, $C_b^k(\mathbb{R}^d) \subset D(L_p)$. Since L_p is a closed operator and it is an extension of $\mathcal{L} : C_b^k(\mathbb{R}^d) \to L^p(\mathbb{R}^d, \gamma_d)$, $\mathcal{L} : C_b^k(\mathbb{R}^d) \to L^p(\mathbb{R}^d, \gamma_d)$ is closable.

Applying Lemma 11.1.9 with $D = C_b^k(\mathbb{R}^d)$, we obtain that $D(L_p)$ is the closure of $C_b^k(\mathbb{R}^d)$ in the graph norm of L_p, namely $f \in D(L_p)$ iff there exists a sequence $(f_n) \subset C_b^k(\mathbb{R}^d)$ such that $f_n \to f$ in $L^p(\mathbb{R}^d, \gamma_d)$ and $L_p f_n = \mathcal{L}f_n$ converges in $L^p(\mathbb{R}^d, \gamma_d)$. This shows that L_p is the closure of $\mathcal{L} : C_b^k(\mathbb{R}^d) \to L^p(\mathbb{R}^d, \gamma_d)$. □

In the case $p = 2$ we obtain other characterisations of $D(L_2)$. To start with, we point out some important properties of $\mathcal{L}$, when applied to elements of $W^{2,2}(\mathbb{R}^d, \gamma_d)$.

Lemma 13.1.3. *(a) $\mathcal{L} : W^{2,2}(\mathbb{R}^d, \gamma_d) \to L^2(\mathbb{R}^d, \gamma_d)$ is a bounded operator;*

(b) for every $f \in W^{2,2}(\mathbb{R}^d, \gamma_d)$, $g \in W^{1,2}(\mathbb{R}^d, \gamma_d)$ we have

$$\int_{\mathbb{R}^d} \mathcal{L}f\, g\, d\gamma_d = -\int_{\mathbb{R}^d} \nabla f \cdot \nabla g\, d\gamma_d. \tag{13.1.5}$$

(c) for every $f \in W^{2,2}(\mathbb{R}^d, \gamma_d)$ we have

$$\mathcal{L}f = \mathrm{div}_{\gamma_d} \nabla f. \tag{13.1.6}$$

Proof. To prove (a) it is sufficient to show that the mapping $T : W^{2,2}(\mathbb{R}^d, \gamma_d) \to L^2(\mathbb{R}^d, \gamma_d)$ defined by $(Tf)(x) := x \cdot \nabla f(x)$ is bounded. For every $i = 1, \dots, d$, set $g_i(x) = x_i D_i f(x)$. The mapping $f \mapsto g_i$ is bounded from $W^{2,2}(\mathbb{R}^d, \gamma_d)$ to $L^2(\mathbb{R}^d, \gamma_d)$ by Lemma 10.2.6, and summing up the statement follows.

To prove (b) it is sufficient to apply the integration by parts formula (9.1.3) to compute $\int_{\mathbb{R}^d} D_{ii} f\, g\, d\gamma_d$, for every $i = 1, \dots, d$, and to sum up. In fact, (9.1.3) was stated for C^1_b functions, but it is readily extended to Sobolev functions using Proposition 9.1.5.

Statement (c) follows from Theorem 10.2.7. In this case we have $H = \mathbb{R}^d$, and it is convenient to take the canonical basis of $\mathbb{R}^d$ as a basis for H. So, we have $\hat{h}_i(x) = x_i$ for $i = 1, \dots, d$ and $\mathrm{div}_{\gamma_d} v(x) = \sum_{i=1}^d D_i v_i - x_i v_i$, for every $v \in W^{1,2}(\mathbb{R}^d, \gamma_d; \mathbb{R}^d)$. Taking $v = \nabla f$, (13.1.6) follows. □

The first characterisation of $D(L_2)$ is the following.

Theorem 13.1.4. *$D(L_2) = W^{2,2}(\mathbb{R}^d, \gamma_d)$, and $L_2 f = \mathcal{L}f$ for every $f \in W^{2,2}(\mathbb{R}^d, \gamma_d)$. Moreover, for every $f \in W^{2,2}(\mathbb{R}^d, \gamma_d)$,*

$$\|f\|^2_{L^2(\mathbb{R}^d,\gamma_d)} + \|\mathcal{L}f\|_{L^2(\mathbb{R}^d,\gamma_d)} = \|f\|^2_{L^2(\mathbb{R}^d,\gamma_d)} + \|\nabla f\|^2_{L^2(\mathbb{R}^d,\gamma_d;\mathbb{R}^d)} + \|D^2 f\|^2_{L^2(\mathbb{R}^d,\gamma_d;\mathbb{R}^{d\times d})}, \tag{13.1.7}$$

where the square root of the right hand side is an equivalent norm in $W^{2,2}(\mathbb{R}^d, \gamma_d)$.

Proof. Let us prove that (13.1.7) holds for every $f \in C^3_b(\mathbb{R}^d)$. We set $\mathcal{L}f =: g$ and we differentiate with respect to x_j (this is why we consider C^3_b, instead of only C^2_b, functions) for every $j = 1, \dots, d$. We obtain

$$D_j(\Delta f) - \sum_{i=1}^d (\delta_{ij} D_i f + x_i D_{ji} f) = D_j g.$$

Multiplying by $D_j f$ and summing up we get

$$\sum_{j=1}^d D_j f \Delta(D_j f) - |\nabla f|^2 - \sum_{j=1}^d x \cdot \nabla(D_j f) D_j f = \nabla f \cdot \nabla g.$$

Note that each term in the above sum belongs to $L^p(\mathbb{R}^d, \gamma_d)$ for every $p > 1$. We integrate over $\mathbb{R}^d$ and we obtain

$$\int_{\mathbb{R}^d} \Big(\sum_{j=1}^d D_j f \, \mathcal{L}(D_j f) - |\nabla f|^2 \Big) d\gamma_d = \int_{\mathbb{R}^d} \nabla f \cdot \nabla g \, d\gamma_d.$$

Now we use the integration formula (13.1.5), both in the left hand side and in the right hand side, obtaining

$$-\int_{\mathbb{R}^d} \sum_{j=1}^d |\nabla D_j f|^2 d\gamma_d - \int_{\mathbb{R}^d} |\nabla f|^2 d\gamma_d = -\int_{\mathbb{R}^d} g \, \mathcal{L} f \, d\gamma_d$$

so that, since $g = \mathcal{L} f$

$$\int_{\mathbb{R}^d} (\mathcal{L} f)^2 d\gamma_d = \int_{\mathbb{R}^d} \sum_{i,j=1}^d (D_{ij} f)^2 d\gamma_d + \int_{\mathbb{R}^d} |\nabla f|^2 d\gamma_d.$$

Since C^3_b is dense both in $W^{2,2}(\mathbb{R}^d, \gamma_d)$ and in $D(L_2)$, the statement follows.

□

The next characterisation fits last year Isem. We recall below general results about bilinear forms in Hilbert spaces. We only need a basic result; more refined results are in last year Isem lecture notes.

Let $V \subset W$ be real Hilbert spaces, with continuous and dense embedding, and let $\mathcal{Q} : V \times V \to \mathbb{R}$ be a bounded bilinear form. "Bounded" means that there exists $M > 0$ such that $|\mathcal{Q}(u, v)| \leq M \|u\|_V \, \|v\|_V$ for every u, $v \in V$; "bilinear" means that $\mathcal{Q}$ is linear both with respect to u and with respect to v. $\mathcal{Q}$ is called "nonnegative" if $\mathcal{Q}(u, u) \geq 0$ for every $u \in V$, and "coercive" if there is $c > 0$ such that $\mathcal{Q}(u, u) \geq c\|u\|_V^2$, for every $u \in V$; it is called "symmetric" if $\mathcal{Q}(u, v) = \mathcal{Q}(v, u)$ for every u, $v \in V$. Note that the form in (13.1.8), with $V = W^{1,2}(\mathbb{R}^d, \gamma_d)$, $W = L^2(\mathbb{R}^d, \gamma_d)$ is bounded, bilinear, symmetric and nonnegative. It is not coercive, but $\mathcal{Q}(u, v) + \alpha\langle u, v\rangle_{L^2(\mathbb{R}^d, \gamma_d)}$ is coercive for every $\alpha > 0$.

For any bounded bilinear form $\mathcal{Q}$, an unbounded linear operator A in the space W is naturally associated with $\mathcal{Q}$. $D(A)$ consists of the elements $u \in V$ such that the mapping $V \to \mathbb{R}$, $v \mapsto \mathcal{Q}(u, v)$, has a linear bounded extension to the whole W. By the Riesz Theorem, this is equivalent to the existence of $g \in W$ such that $\mathcal{Q}(u, v) = \langle g, v\rangle_W$, for every $v \in V$. Note that g is unique, because V is dense in W. Then we set $Au = -g$, where g is the unique element of W such that $\mathcal{Q}(u, v) = \langle g, v\rangle_W$, for every $v \in V$.

Theorem 13.1.5. *Let $V \subset W$ be real Hilbert spaces, with continuous and dense embedding, and let $\mathcal{Q} : V \times V \to \mathbb{R}$ be a bounded bilinear symmetric form, such that $(u, v) \mapsto \mathcal{Q}(u, v) + \alpha\langle u, v\rangle_W$ is coercive for some $\alpha > 0$. Then the operator $A : D(A) \to W$ defined above is densely defined and self-adjoint. If in addition $\mathcal{Q}$ is nonnegative, A is dissipative.*

Proof. The mapping $(u, v) \mapsto \mathcal{Q}(u, v) + \alpha\langle u, v\rangle_W$ is an inner product in V, and the associated norm is equivalent to the V-norm, by the continuity of $\mathcal{Q}$ and the coercivity assumption.

It is convenient to consider the operator $\widetilde{A} : D(\widetilde{A}) = D(A) \to W$, $\widetilde{A}u := Au + \alpha u$. Of course if $\widetilde{A}$ is self-adjoint, also A is self-adjoint.

We consider the canonical isomorphism $T : V \to V^*$ defined by $(Tu)(v) = \mathcal{Q}(u, v) + \alpha\langle u, v\rangle_W$ (we are using the new inner product above defined), and the embedding $J : W \to V^*$, such that $(Ju)(v) = \langle u, v\rangle_W$. T is an isometry by the Riesz Theorem, and J is bounded since for every $u \in W$ and $v \in V$ we have $|(Ju)(v)| \leq \|u\|_W\|v\|_W \leq C\|u\|_W\|v\|_V$, where C is the norm of the embedding $V \subset W$. Moreover, J is one to one, since V is dense in W.

By definition, $u \in D(\widetilde{A})$ iff there exists $g \in W$ such that $\mathcal{Q}(u, v) + \alpha\langle u, v\rangle_W = \langle g, v\rangle_W$ for every $v \in V$, which means $Tu = Jg$, and in this case $\widetilde{A}u = -g$.

The range of J is dense in V^*. If it were not, there would exist $\Phi \in V^* \setminus \{0\}$ such that $\langle Jw, \Phi\rangle_{V^*} = 0$ for every $w \in W$. So, there would exists $\varphi \in V \setminus \{0\}$ such that $Jw(\varphi) = 0$, namely $\langle w, \varphi\rangle_W = 0$ for every $w \in W$. This implies $\varphi = 0$, a contradiction. Since T is an isomorphism, the range of $T^{-1}J$, which is nothing but the domain of $\widetilde{A}$, is dense in V. Since V is in its turn dense in W, $D(\widetilde{A})$ is dense in W.

The symmetry of $\mathcal{Q}$ implies immediately that $\widetilde{A}$ is self–adjoint. Indeed, for u, $v \in D(\widetilde{A})$ we have

$$\langle \widetilde{A}u, v\rangle_W = \mathcal{Q}(u, v) + \alpha\langle u, v\rangle_W = \mathcal{Q}(v, u) + \alpha\langle v, u\rangle_W = \langle u, \widetilde{A}v\rangle_W.$$

Since $\widetilde{A}$ is onto, it is self-adjoint.

The last statement is obvious: since $\langle Au, u\rangle = -\mathcal{Q}(u, u)$ for every $u \in D(A)$, if $\mathcal{Q}$ is nonnegative, A is dissipative. □

In our setting the bilinear form is

$$\mathcal{Q}(u, v) := \int_{\mathbb{R}^d} \nabla u \cdot \nabla v \, d\gamma_d, \quad u,\ v \in W^{1,2}(\mathbb{R}^d, \gamma_d), \tag{13.1.8}$$

so that the assumptions of Theorem 13.1.5 are satisfied with $W = L^2(\mathbb{R}^d, \gamma_d)$, $V = W^{1,2}(\mathbb{R}^d, \gamma_d)$ and every $\alpha > 0$. $D(A)$ is the set

$$\Big\{u \in W^{1,2}(\mathbb{R}^d, \gamma_d) : \ \exists g \in L^2(\mathbb{R}^d, \gamma_d) \text{ such that } \mathcal{Q}(u, v) = \int_{\mathbb{R}^d} g\, v\, d\gamma_d,\ \forall v \in W^{1,2}(\mathbb{R}^d, \gamma_d)\Big\}$$

and $Au = -g$.

Theorem 13.1.6. *Let $\mathcal{Q}$ be the bilinear form in* (13.1.8)*. Then $D(A) = W^{2,2}(\mathbb{R}^d, \gamma_d)$, and $A = L_2$.*

Proof. Let $u \in W^{2,2}(\mathbb{R}^d, \gamma_d)$. By (13.1.5) and Theorem 13.1.4, for every $v \in W^{1,2}(\mathbb{R}^d, \gamma_d)$ we have

$$\mathcal{Q}(u, v) = -\int_{\mathbb{R}^d} \mathcal{L}u\, v\, d\gamma_d$$

Therefore, the function $g = \mathcal{L}u = L_2 u$ fits the definition of Au (recall that $g \in L^2(\mathbb{R}^d, \gamma_d)$ by Lemma 13.1.3(a)). So, $W^{2,2}(\mathbb{R}^d, \gamma_d) \subset D(A)$ and $Au = L_2 u$ for every $u \in W^{2,2}(\mathbb{R}^d, \gamma_d)$ (the last equality follows from Theorem 13.1.4). In other words, A is a self-adjoint extension of L_2. L_2 itself is self-adjoint by Corollary 11.4.5, because $T_2(t)$ is self-adjoint in $L^2(\mathbb{R}^d, \gamma_d)$ by Proposition 12.1.5(ii), for every $t > 0$. Self-adjoint operators have no proper self-adjoint extensions (this is because $D(L_2) \subset D(A) \Rightarrow D(A^*) \subset D(L_2^*)$, but $D(A^*) = D(A)$ and $D(L_2^*) = D(L_2)$), hence $D(A) = D(L_2)$ and $A = L_2$. □

13.2 The infinite dimensional case

Here, as usual, X is a separable Banach space endowed with a centred nondegenerate Gaussian measure γ, and H is the relevant Cameron-Martin space.

The connection between finite dimension and infinite dimension is provided by the cylindrical functions. In the next proposition we show that suitable cylindrical functions belong to $D(L_p)$ for every $p \in (1, \infty)$, and we write down an explicit expression of $L_p f$ for such f. Precisely, we fix an orthonormal basis $\{h_j : j \in \mathbb{N}\}$ of H contained in $R_\gamma(X^*)$, and we denote by Σ the set of the cylindrical functions of the type $f(x) = \varphi(\hat{h}_1(x), \ldots, \hat{h}_d(x))$ with $\varphi \in C_b^2(\mathbb{R}^d)$, for some $d \in \mathbb{N}$. This is a dense subspace of $L^p(X, \gamma)$ for every $p \in [1, \infty)$, see Exercise 13.3. For such f, we have

$$\partial_i f(x) = \frac{\partial \varphi}{\partial \xi_i}(\hat{h}_1(x), \ldots, \hat{h}_d(x)),\ i \leq d; \quad \partial_i f(x) = 0,\ i > d. \tag{13.2.1}$$

To distinguish between the finite and the infinite dimensional case, we use the superscript (d) when dealing with the Ornstein-Uhlenbeck semigroup and the Ornstein-Uhlenbeck semigroup in $\mathbb{R}^d$. So, $L_p^{(d)}$ is the infinitesimal generator of the Ornstein-Uhlenbeck semigroup $T^{(d)}(t)$ in $L^p(\mathbb{R}^d, \gamma_d)$. We recall that $L_p^{(d)}$ is a realisation of the operator $\mathcal{L}^{(d)} = \Delta - x \cdot \nabla$, namely $L_p^{(d)} f = \mathcal{L}^{(d)} f$ for every $f \in D(L_p^{(d)})$.

Proposition 13.2.1. *Let $\{h_j : j \in \mathbb{N}\}$ be an orthonormal basis of H contained in $R_\gamma(X^*)$. Let $d \in \mathbb{N}$, $p \in [1, \infty)$ and $f(x) = \varphi(\hat{h}_1(x), \ldots, \hat{h}_d(x))$ with $\varphi \in L^p(\mathbb{R}^d, \gamma_d)$. Then for every $t > 0$ and γ-a.e. $x \in X$,*

$$T_p(t) f(x) = (T_p^{(d)}(t)\varphi)(\hat{h}_1(x), \ldots, \hat{h}_d(x)).$$

If in addition $\varphi \in D(L_p^{(d)})$, then $f \in D(L_p)$, and

$$L_p f(x) = L_p^{(d)} \varphi(\hat{h}_1(x), \ldots, \hat{h}_d(x)).$$

If $\varphi \in C_b^2(\mathbb{R}^d)$, then

$$L_p f(x) = \mathcal{L}^{(d)} \varphi(\hat{h}_1(x), \ldots, \hat{h}_d(x)) = \sum_{i=1}^{d} (\partial_{ii} f(x) - \hat{h}_i(x) \partial_i f(x)) = \operatorname{div}_\gamma \nabla_H f(x).$$

Proof. Assume first that $\varphi \in C_b(\mathbb{R}^d)$. For $t > 0$ we have

$$\begin{aligned} T(t)f(x) &= \int_X f(e^{-t}x + \sqrt{1-e^{-2t}}y)\gamma(dy) \\ &= \int_X \varphi(e^{-t}\hat{h}_1(x) + \sqrt{1-e^{-2t}}\hat{h}_1(y), \dots, e^{-t}\hat{h}_d(x) + \sqrt{1-e^{-2t}}\hat{h}_d(y))\gamma(dy) \\ &= \int_{\mathbb{R}^d} \varphi(e^{-t}\hat{h}_1(x) + \sqrt{1-e^{-2t}}\xi_1, \dots, e^{-t}\hat{h}_d(x) + \sqrt{1-e^{-2t}}\xi_d)\gamma_d(d\xi) \\ &= (T^{(d)}(t)\varphi)(\hat{h}_1(x), \dots, \hat{h}_d(x)), \end{aligned}$$

because $\gamma \circ (\hat{h}_1, \dots, \hat{h}_d)^{-1} = \gamma_d$ by Exercise 2.4. If $\varphi \in L^p(\mathbb{R}^d, \gamma_d)$ is not continuous, we approximate it in $L^p(\mathbb{R}^d, \gamma_d)$ by a sequence of continuous and bounded functions φ_n. The sequence $f_n(x) := \varphi_n(\hat{h}_1(x), \dots, \hat{h}_d(x))$ converges to f and the sequence $g_n(x) := (T^{(d)}(t)\varphi_n)(\hat{h}_1(x), \dots, \hat{h}_d(x))$ converges to $(T^{(d)}(t)\varphi)(\hat{h}_1(x), \dots, \hat{h}_d(x))$ in $L^p(X, \gamma)$, still by Exercise 2.4. Therefore, $T(t)f_n$ converges to $T(t)f$ in $L^p(X, \gamma)$ for every $t > 0$, and the first statement follows.

Let now $\varphi \in D(L_p^{(d)})$. For every $t > 0$ we have

$$\begin{aligned} &\int_X \Big| \frac{T(t)f(x) - f(x)}{t} - L_p^{(d)}\varphi(\hat{h}_1(x), \dots, \hat{h}_d(x))\Big|^p \gamma(dx) \\ &= \int_X \Big| \frac{T^{(d)}(t)\varphi(\hat{h}_1(x), \dots, \hat{h}_d(x)) - \varphi(\hat{h}_1(x), \dots, \hat{h}_d(x))}{t} - L_p^{(d)}\varphi(\hat{h}_1(x), \dots, \hat{h}_d(x))\Big|^p \gamma(dx) \\ &= \int_{\mathbb{R}^d} \Big| \frac{T^{(d)}(t)\varphi(\xi) - \varphi(\xi)}{t} - L_p^{(d)}\varphi(\xi)\Big|^p \gamma_d(d\xi) \end{aligned}$$

that vanishes as $t \to 0$. So, the second statement follows.

Let $\varphi \in C_b^2(\mathbb{R}^d)$. By Theorem 13.1.2 we have

$$L_p^{(d)}\varphi(\xi) = \sum_{i=1}^d (D_{ii}\varphi(\xi) - \xi_i D_i\varphi(\xi)) = \mathcal{L}^{(d)}\varphi(\xi), \quad \xi \in \mathbb{R}^d.$$

Therefore,

$$\begin{aligned} L_p f(x) &= (\mathcal{L}^{(d)}\varphi)(\hat{h}_1(x), \dots, \hat{h}_d(x)) = \sum_{i=1}^d (D_{ii}\varphi(\xi) - \xi_i D_i\varphi(\xi))_{|\xi = (\hat{h}_1(x), \dots, \hat{h}_d(x))} \\ &= \sum_{i=1}^d (\partial_{ii} f(x) - \hat{h}_i(x)\partial_i f(x)), \end{aligned}$$

which coincides with $\mathrm{div}_\gamma \nabla_H f(x)$. See Theorem 10.2.7. □

As a consequence of Propositions 13.2.1 and 11.1.9, we obtain a characterisation of $D(L_p)$ which is quite similar to the finite dimensional one.

Theorem 13.2.2. *Let $\{h_j : j \in \mathbb{N}\}$ be an orthonormal basis of H contained in $R_\gamma(X^*)$. Then the subspace Σ of $\mathcal{F}C_b^2(X)$ defined above is a core of L_p for every $p \in [1,\infty)$, the restriction of L_p to Σ is closable in $L^p(X,\gamma)$ and its closure is L_p. In other words, $D(L_p)$ consists of all $f \in L^p(X,\gamma)$ such that there exists a sequence (f_n) in Σ which converges to f in $L^p(X,\gamma)$ and such that $L_p f_n = \mathrm{div}_\gamma \nabla_H f_n$ converges in $L^p(X,\gamma)$.*

Proof. By Proposition 13.2.1, $\Sigma \subset D(L_p)$. For every $t > 0$, $T(t)f \in \Sigma$ if $f \in \Sigma$, by Proposition 13.2.1 and Proposition 12.1.4. By Lemma 11.1.9, Σ is a core of L_p. □

For $p = 2$ we can prove other characterisations.

Theorem 13.2.3. $D(L_2) = W^{2,2}(X,\gamma)$, *and for every $f \in W^{2,2}(X,\gamma)$ we have*

$$L_2 f = \mathrm{div}_\gamma \nabla_H f, \tag{13.2.2}$$

and

$$\|f\|_{L^2(X,\gamma)} + \|L_2 f\|_{L^2(X,\gamma)} \le \|f\|_{W^{2,2}(X,\gamma)} \le \frac{3}{2}(\|f\|_{L^2(X,\gamma)} + \|L_2 f\|_{L^2(X,\gamma)}). \tag{13.2.3}$$

Proof. Fix an orthonormal basis of H contained in $R_\gamma(X^*)$. By Exercise 13.3, Σ is dense in $W^{2,2}(X,\gamma)$, and by Theorem 13.2.2 it is dense in $D(L_2)$.

We claim that every $f \in \Sigma$ satisfies (13.2.3), so that the $W^{2,2}$ norm is equivalent to the graph norm of L_2 on Σ. For every $f \in \Sigma$, if $f(x) = \varphi(\hat{h}_1(x), \ldots, \hat{h}_d(x))$, by Proposition 13.2.1 we have $L_2 f(x) = (\mathcal{L}^{(d)}\varphi)(\hat{h}_1(x), \ldots, \hat{h}_d(x))$, where $\mathcal{L}^{(d)}$ is defined in (13.1.1). Recalling that $\gamma \circ (\hat{h}_1, \ldots, \hat{h}_d)^{-1} = \gamma_d$, we get

$$\int_X f^2 d\gamma = \int_{\mathbb{R}^d} \varphi^2 d\gamma_d, \quad \int_X (L_2 f)^2 d\gamma = \int_{\mathbb{R}^d} (\mathcal{L}^{(d)}\varphi)^2 d\gamma_d,$$

and, using (13.2.1),

$$\|f\|_{W^{2,2}(X,\gamma)} = \|\varphi\|_{W^{2,2}(\mathbb{R}^d,\gamma_d)}.$$

Therefore, estimates (13.1.7) imply that f satisfies (13.2.3), and the claim is proved.

The statement is now a standard consequence of the density of Σ in $W^{2,2}(X,\gamma)$ and in $D(L_2)$. Indeed, to prove that $W^{2,2}(X,\gamma) \subset D(L_2)$, and that $L_2 f = \mathrm{div}_\gamma \nabla_H f$ for every $f \in W^{2,2}(\mathbb{R}^d,\gamma_d)$, it is sufficient to approximate any $f \in W^{2,2}(X,\gamma)$ by a sequence (f_n) of elements of Σ; then f_n converges to f and $L_2 f_n = \mathrm{div}_\gamma \nabla_H f_n$ converges to $\mathrm{div}_\gamma \nabla_H f$ in $L^2(X,\gamma)$ by Theorem 10.2.7, as $\nabla_H f_n$ converges to $\nabla_H f$ in $L^2(X,\gamma;H)$. Since L_2 is a closed operator, $f \in D(L_2)$ and $L_2 f = \mathrm{div}_\gamma \nabla_H f$. Similarly, to prove that $D(L_2) \subset W^{2,2}(X,\gamma)$ we approximate any $f \in D(L_2)$ by a sequence (f_n) of elements of Σ that converges to f in the graph norm; then (f_n) is a Cauchy sequence in $W^{2,2}(X,\gamma)$ and therefore $f \in W^{2,2}(X,\gamma)$. □

Finally, as in finite dimension, we have a characterisation of L_2 in terms of the bilinear form

$$\mathcal{Q}(u,v) = \int_X [\nabla_H u, \nabla_H v]_H d\gamma, \quad u,\ v \in W^{1,2}(X,\gamma). \tag{13.2.4}$$

Applying Theorem 13.1.5 with $W = L^2(X,\gamma)$, $V = W^{1,2}(X,\gamma)$ we obtain

Theorem 13.2.4. *Let A be the operator associated with the bilinear form $\mathcal{Q}$ above. Then $D(A) = W^{2,2}(X,\gamma)$, and $A = L_2$.*

The proof is identical to the proof of Theorem 13.1.6 and it is omitted.

Note that Theorem 13.2.4 implies that for every $f \in D(L_2) = W^{2,2}(X,\gamma)$ and for every $g \in W^{1,2}(X,\gamma)$ we have

$$\int_X L_2 f\, g\, d\gamma = -\int_X [\nabla_H f, \nabla_H g]_H d\gamma, \tag{13.2.5}$$

which is the infinite dimensional version of (13.1.5). Proposition 11.4.3 implies that L_2 is a sectorial operator, therefore the Ornstein-Uhlenbeck semigroup is analytic in $L^2(X,\gamma)$.

We mention that, by general results about semigroups and interpolation theory (e.g. [7, Thm. 1.4.2]), $\{T_p(t) : \ t \geq 0\}$ is an analytic semigroup in $L^p(X,\gamma)$ for every $p \in (1,\infty)$. However, this fact will not be used in these lectures.

A result similar to Theorem 13.2.3 holds also for $p \neq 2$. More precisely, for every $p \in (1,\infty)$, $D(L_p) = W^{2,p}(X,\gamma)$, and the graph norm of $D(L_p)$ is equivalent to the $W^{2,p}$ norm. But the proof is not as simple. We refer to [25] and [3, Sect. 5.5] for the infinite dimensional case, and to [23] for an alternative proof in the finite dimensional case.

13.3 Exercises

Exercise 13.1. Let $\varrho : \mathbb{R}^d \to [0,\infty)$ be a mollifier, i.e. a smooth function with support in $B(0,1)$ such that

$$\int_{B(0,1)} \varrho(x)dx = 1.$$

For $\varepsilon > 0$ set

$$\varrho_\varepsilon(x) = \varepsilon^{-d}\varrho\left(\frac{x}{\varepsilon}\right), \quad x \in \mathbb{R}^d.$$

Prove that if $p \in [1,\infty)$ and $f \in L^p(\mathbb{R}^d, \gamma_d)$, then

$$f_\varepsilon(x) := f * \varrho_\varepsilon(x) = \int_{\mathbb{R}^d} f(y)\varrho_\varepsilon(x-y)dy,$$

is well defined, belongs to $L^p(\mathbb{R}^d, \gamma_d)$ and converges to f in $L^p(\mathbb{R}^d, \gamma_d)$ as $\varepsilon \to 0^+$.

Exercise 13.2. Prove that for every $k \in \mathbb{N}$, $k \geq 3$, $C^k_b(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d, \gamma_d)$ and that $T(t) \in \mathcal{L}(C^k_b(\mathbb{R}^d))$ for every $t > 0$.

Exercise 13.3. Let $\{h_j : \ j \in \mathbb{N}\}$ be an orthonormal basis of H contained in $R_\gamma(X^*)$. Prove that the set Σ of the cylindrical functions of the type $f(x) = \varphi(\hat{h}_1(x), \ldots, \hat{h}_d(x))$ with $\varphi \in C^2_b(\mathbb{R}^d)$, for some $d \in \mathbb{N}$, is dense in $L^p(X,\gamma)$ and in $W^{2,p}(X,\gamma)$ for every $p \in [1,\infty)$.

Exercise 13.4.

(i) With the help of Proposition 10.1.2, show that if $f \in W^{1,p}(X, \gamma)$ with $p \in [1, \infty)$ is such that $\nabla_H f = 0$ a.e., then f is a.e. constant.

(ii) Use point (i) to show that for every $p \in (1, \infty)$ the kernel of L_p consists of the constant functions.
(HINT: First of all, prove that $T(t)f = f$ for all $f \in D(L_p)$ such that $L_p f = 0$ and then pass to the limit as $t \to \infty$ in (12.1.3))

Lecture 14

More on Ornstein-Uhlenbeck operator and semigroup

In this lecture we go on in the study of the realisation of the Ornstein-Uhlenbeck operator and of the Ornstein-Uhlenbeck semigroup in L^p spaces. As in the last lectures, X is a separable Banach space endowed with a centred nondegenerate Gaussian measure γ, and H is the Cameron-Martin space. We use the notation of Lectures 12 and 13.

We start with the description of the spectrum of L_2. Although the domain of L_2 is not compactly embedded in $L^2(X,\gamma)$ if X is infinite dimensional, the spectrum of L_2 consists of a sequence of eigenvalues, and the corresponding eigenfunctions are the Hermite polynomials that we already encountered in Lecture 8. So, $L^2(X,\gamma)$ has an orthonormal basis made by eigenfunctions of L_2. This is used to obtain another representation formula for $T_2(t)$ and another characterisation of $D(L_2)$ in terms of Hermite polynomials.

In the second part of the lecture we present two important inequalities, the the Logarithmic Sobolev and Poincaré inequalities, that hold for C^1_b functions and are easily extended to Sobolev functions. They are used to prove summability improving properties and asymptotic behavior results for $T_p(t)$.

14.1 Spectral properties of L_2

Let $\{h_j:\ j\in\mathbb{N}\}$ be an orthonormal basis of H contained in $R_\gamma(X^*)$. We recall the definition of the Hermite polynomials, given in Lecture 8.

Λ is the set of of multi-indices $\alpha\in(\mathbb{N}\cup\{0\})^{\mathbb{N}}$, $\alpha=(\alpha_j)$, with finite length $|\alpha|=\sum_{j=1}^\infty\alpha_j<\infty$. For every $\alpha\in\Lambda$, $\alpha=(\alpha_j)$, the Hermite polynomial H_α is defined by

$$H_\alpha(x)=\prod_{j=1}^\infty H_{\alpha_j}(\hat h_j(x)),\quad x\in X.$$

where the polynomial H_{α_j} is defined in (8.1.1). By Lemma 8.1.2, for every $k \in \mathbb{N}$ we have

$$H_k''(\xi) - \xi H_k'(\xi) = -kH_k(\xi), \quad \xi \in \mathbb{R},$$

namely H_k is an eigenfunction of the one-dimensional Ornstein-Uhlenbeck operator, with eigenvalue $-k$. This property is extended to any dimension as follows.

Proposition 14.1.1. *For every $\alpha \in \Lambda$, H_α belongs to $D(L_2)$ and*

$$L_2 H_\alpha = -|\alpha| H_\alpha.$$

Proof. As a first step, we consider the finite dimensional case $X = \mathbb{R}^d$, $\gamma = \gamma_d$. Then $H = \mathbb{R}^d$ and we take the canonical basis of $\mathbb{R}^d$ as a basis for H, so that $\hat{h}_j(x) = x_j$ for $j = 1, \dots d$.

We fix a Hermite polynomial H_α in $\mathbb{R}^d$, with $\alpha = (\alpha_1, \dots, \alpha_d) \in (\mathbb{N} \cup \{0\})^d$,

$$H_\alpha(x) = \prod_{j=1}^{d} H_{\alpha_j}(x_j), \quad x \in \mathbb{R}^d.$$

H_α belongs to $W^{2,2}(\mathbb{R}^d, \gamma_d)$ (in fact, it belongs to $W^{2,p}(\mathbb{R}^d, \gamma_d)$ for every $p \in [1, \infty)$) and therefore by Theorem 13.1.4, it is in $D(L_2^{(d)})$. By (8.1.4) we know that $L_2^{(d)} H_\alpha = \mathcal{L}^{(d)} H_\alpha = -|\alpha| H_\alpha$.

Now we turn to the infinite dimensional case. Let $\alpha \in \Lambda$ and let $d \in \mathbb{N}$ be such that $\alpha_j = 0$ for each $j > d$. Then $H_\alpha(x) = \varphi(\hat{h}_1(x), \dots, \hat{h}_d(x))$, where φ is a Hermite polynomial in $\mathbb{R}^d$. Proposition 13.2.1 implies that $H_\alpha \in D(L_2)$, and

$$L_2 H_\alpha = \mathcal{L}^{(d)} \varphi(\hat{h}_1(\cdot), \dots, \hat{h}_d(\cdot)) = -|\alpha| \varphi(\hat{h}_1(\cdot), \dots, \hat{h}_d(\cdot)) = -|\alpha| H_\alpha.$$

□

As a consequence of Propositions 14.1.1 and 8.1.9, we characterise the spectrum of L_2. We recall that, for every $k \in \mathbb{N} \cup \{0\}$, I_k is the orthogonal projection on the subspace $\mathcal{X}_k = \overline{\mathrm{span}\{H_\alpha : \ \alpha \in \Lambda, \ |\alpha| = k\}}$ of $L^2(X, \gamma)$. See Section 8.1.2.

Proposition 14.1.2. *The spectrum of L_2 is equal to $-\mathbb{N} \cup \{0\}$. For every $k \in \mathbb{N} \cup \{0\}$, $\mathcal{X}_k$ is the eigenspace of L_2 with eigenvalue $-k$. Therefore, $I_k(L_2 f) = L_2(I_k f) = -kI_k(f)$, for every $f \in D(L_2)$.*

Proof. Let us consider the point spectrum. First of all, we prove that $\mathcal{X}_k$ is contained in the eigenspace of L_2 with eigenvalue $-k$.

$\mathcal{X}_0$ consists of the constant functions, which belong to the kernel of L_2. For $k \in \mathbb{N}$, every element $f \in \mathcal{X}_k$ is equal to $\lim_{n\to\infty} f_n$, where each f_n is a linear combination of Hermite polynomials H_α with $|\alpha| = k$. By Proposition 14.1.1, $f_n \in D(L_2)$ and $L_2 f_n = -kf_n$. Since L_2 is a closed operator, $f \in D(L_2)$ and $L_2 f = -kf$.

Let now $f \in D(L_2)$ be such that $L_2 f = \lambda f$ for some $\lambda \in \mathbb{R}$. For every $\alpha \in \Lambda$ we have

$$\lambda \langle f, H_\alpha \rangle_{L^2(X,\gamma)} = \langle L_2 f, H_\alpha \rangle_{L^2(X,\gamma)} = \langle f, L_2 H_\alpha \rangle_{L^2(X,\gamma)} = -|\alpha| \langle f, H_\alpha \rangle_{L^2(X,\gamma)}.$$

Therefore, either $\lambda = -|\alpha|$ or $\langle f, H_\alpha\rangle_{L^2(X,\gamma)} = 0$. If $\lambda = -k$ with $k \in \mathbb{N}\cup\{0\}$, then f is orthogonal to all Hermite polynomials H_β with $|\beta| \neq k$, hence $f \in \mathfrak{X}_k$ is an eigenfunction of L_2 with eigenvalue $-k$. If $\lambda \neq -k$ for every $k \in \mathbb{N}\cup\{0\}$, then f is orthogonal to all Hermite polynomials so that it vanishes. This proves that $\mathfrak{X}_k$ is equal to the eigenspace of L_2 with eigenvalue $-k$.

Since L_2 is self-adjoint, for $f \in D(L_2)$ and $|\alpha| = k$ we have

$$\langle L_2 f, H_\alpha\rangle_{L^2(X,\gamma)} = \langle f, L_2 H_\alpha\rangle_{L^2(X,\gamma)} = -k\langle f, H_\alpha\rangle_{L^2(X,\gamma)}. \tag{14.1.1}$$

Let f_j, $j \in \mathbb{N}$, be any enumeration of the Hermite polynomials H_α with $|\alpha| = k$. The sequence $s_n := \sum_{j=0}^{n}\langle f, f_j\rangle_{L^2(X,\gamma)} f_j$ converges in $D(L_2)$, since $L^2 - \lim_{n\to\infty} s_n = I_k(f)$ and

$$L_2 s_n = \sum_{j=0}^{n}\langle f, f_j\rangle_{L^2(X,\gamma)} L_2 f_j = -k\sum_{j=0}^{n}\langle f, f_j\rangle_{L^2(X,\gamma)} f_j = \sum_{j=0}^{n}\langle L_2 f, f_j\rangle_{L^2(X,\gamma)} f_j,$$

where the last equality follows from (14.1.1). The series in the right hand side converges to $-kI_k(f) = I_k(L_2 f)$, as $n \to \infty$. Then, $L_2 I_k(f) = -kI_k(f) = I_k(L_2 f)$, for every $k \in \mathbb{N}\cup\{0\}$.

It remains to show that the spectrum of L_2 is just $-\mathbb{N}\cup\{0\}$. We notice that $D(L_2)$ is not compactly embedded in $L^2(X,\gamma)$ if X is infinite dimensional, because it has infinite dimensional eigenspaces. So, the spectrum does not necessarily consist of eigenvalues.

If $\lambda \neq -h$ for every $h \in \mathbb{N}\cup\{0\}$, and $f \in L^2(X,\gamma)$, the resolvent equation $\lambda u - L_2 u = f$ is equivalent to $\lambda I_k(u) - I_k(L_2 u) = I_k(f)$ for every $k \in \mathbb{N}\cup\{0\}$, and therefore to $\lambda I_k(u) + kI_k(u) = I_k(f)$, for every $k \in \mathbb{N}\cup\{0\}$. So, we define

$$u = \sum_{k=0}^{\infty}\frac{1}{\lambda + k}\, I_k(f). \tag{14.1.2}$$

The sequence $u_n := \sum_{k=0}^{n} I_k(f)/(\lambda+k)$ converges in $D(L_2)$, since both sequences $1/(\lambda+k)$ and $k/(\lambda+k)$ are bounded. Therefore, $u \in D(L_2)$, and $\lambda u - L_2 u = f$. □

Another consequence is a characterisation of L_2 in terms of Hermite polynomials.

Proposition 14.1.3.

$$\begin{cases} (a) \quad D(L_2) = \Big\{f \in L^2(X,\gamma) : \displaystyle\sum_{k=1}^{\infty} k^2\|I_k(f)\|^2_{L^2(X,\gamma)} < \infty\Big\}, \\[2ex] (b) \quad L_2 f = -\displaystyle\sum_{k=1}^{\infty} kI_k(f), \quad f \in D(L_2). \end{cases} \tag{14.1.3}$$

Proof. Let $f \in D(L_2)$. Then $I_k(L_2 f) = -kI_k(f) = L_2(I_k(f))$ for every $k \in \mathbb{N}\cup\{0\}$, by Proposition 14.1.2. Applying (8.1.9) to $L_2 f$ we obtain

$$L_2 f = \sum_{k=0}^{\infty} I_k(L_2 f) = \sum_{k=1}^{\infty} -kI_k(f)$$

which proves (14.1.3)(b). Moreover,

$$\|L_2 f\|^2_{L^2(X,\gamma)} = \sum_{k=1}^{\infty} k^2 \|I_k(f)\|^2_{L^2(X,\gamma)} < \infty.$$

Conversely, let $f \in L^2(X,\gamma)$ be such that $\sum_{k=1}^{\infty} k^2 \|I_k(f)\|^2_{L^2(X,\gamma)} < \infty$. Then the sequence

$$f_n := \sum_{k=0}^{n} I_k(f)$$

converges to f in $L^2(X,\gamma)$, and it converges in $D(L_2)$ too, since for $n > m$

$$\|L_2(f_n - f_m)\|^2_{L^2(X,\gamma)} = \Big\| \sum_{k=m+1}^{n} -kI_k(f) \Big\|^2_{L^2(X,\gamma)} = \sum_{k=m+1}^{n} k^2 \|I_k(f)\|^2_{L^2(X,\gamma)} \to 0 \text{ as } m \to \infty.$$

Since L_2 is closed, $f \in D(L_2)$. □

As every H_α is an eigenfunction of L_2 with eigenvalue $-|\alpha|$, we ask to verify that

$$T_2(t)H_\alpha = e^{-|\alpha|t} H_\alpha, \quad t \geq 0,\ \alpha \in \Lambda, \tag{14.1.4}$$

see Exercise 14.1. As a consequence, we obtain a very handy expression of $T_2(t)$ in terms of Hermite polynomials.

Corollary 14.1.4. *For every $t > 0$ we have*

$$T_2(t)f = \sum_{k=0}^{\infty} e^{-kt} I_k(f), \quad f \in L^2(X,\gamma), \tag{14.1.5}$$

where the series converges in $L^2(X,\gamma)$. Moreover, $T_2(t)f \in D(L_2)$ and

$$\|L_2 T_2(t) f\|_{L^2(X,\gamma)} \leq \frac{1}{te} \|f\|_{L^2(X,\gamma)}. \tag{14.1.6}$$

The function $t \mapsto T(t)f$ belongs to $C^1((0,\infty); L^2(X,\gamma))$, and

$$\frac{d}{dt} T_2(t)f = L_2 T_2(t) f, \quad t > 0. \tag{14.1.7}$$

Proof. Fix $f \in L^2(X,\gamma)$. By Lemma 14.1.1, for every $k \in \mathbb{N}$, $I_k(f) \in D(L_2)$ and $L_2 I_k(f) = -kI_k(f)$, so that by the above considerations, $T_2(t)I_k(f) = e^{-kt} I_k(f)$. Since $f = \lim_{n\to\infty} \sum_{k=0}^{n} I_k(f)$ in $L^2(X,\gamma)$ and $T_2(t)$ is a bounded operator in $L^2(X,\gamma)$, (14.1.5) follows.

The other statements and estimate $\|L_2 T_2(t) f\|_{L^2(X,\gamma)} \leq c\|f\|_{L^2(X,\gamma)}/t$ follow from the fact that $T_2(t)$ is an analytic semigroup in $L^2(X,\gamma)$, see Theorem 11.4.2. However, we give here a simple independent proof, specifying the constant $c = 1/e$ in (14.1.6).

Since $\sup_{\xi>0} \xi^2 e^{-2\xi} = e^{-2}$, using (14.1.5) we obtain

$$\begin{aligned}\sum_{k=1}^{\infty} k^2 \|I_k(T_2(t)f)\|^2_{L^2(X,\gamma)} &= \sum_{k=1}^{\infty} k^2 e^{-2kt} \|I_k(f)\|^2_{L^2(X,\gamma)} \\ &\leq \frac{1}{e^2 t^2} \sum_{k=1}^{\infty} \|I_k(f)\|^2_{L^2(X,\gamma)} \leq \frac{1}{e^2 t^2} \|f\|^2_{L^2(X,\gamma)}\end{aligned}$$

so that $T_2(t)f \in D(L_2)$ by (14.1.3)(a), and estimate (14.1.6) follows from (14.1.3)(b). Moreover, for every $t > 0$ and $0 < |h| \leq t/2$ we have

$$\left\| \frac{1}{h}(T_2(t+h)f - T_2(t)f) - L_2 T_2(t)f \right\|^2_{L^2(X,\gamma)} = \sum_{k=0}^{\infty} \left(\frac{e^{-k(t+h)} - e^{-kt}}{h} + k \right)^2 \|I_k(f)\|^2_{L^2(X,\gamma)}.$$

Each addend in the right hand side sum converges to 0, and using the Taylor formula for the exponential function we easily obtain

$$\left(\frac{e^{-k(t+h)} - e^{-kt}}{h} + k \right)^2 \leq \frac{t^2}{4} k^4 e^{-2kt} \leq \frac{c}{t^2},$$

with c independent of t, h, k. By the Dominated Convergence Theorem for series, we obtain

$$\lim_{h \to 0^+} \left\| \frac{1}{h}(T_2(t+h)f - T_2(t)f) - L_2 T_2(t)f \right\|_{L^2(X,\gamma)} = 0,$$

namely, the function $T_2(\cdot)f$ is differentiable at t, with derivative $L_2 T_2(t)f$. For $t > t_0 > 0$ we have $L_2 T_2(t)f = L_2 T_2(t - t_0) T(t_0) f = T_2(t - t_0) L_2 T(t_0) f$. Then, $t \mapsto L_2 T_2(t)f$ is continuous in $[t_0, \infty)$. Since t_0 is arbitrary, $T(\cdot)f$ belongs to $C^1((0,\infty); L^2(X,\gamma))$. $\square$

We already know that $D(L_2) = W^{2,2}(X,\gamma)$. So, Proposition 14.1.3 gives a characterisation of $W^{2,2}(X,\gamma)$ in terms of Hermite polynomials. A similar characterisation is available for the space $W^{1,2}(X,\gamma)$.

Proposition 14.1.5.

$$W^{1,2}(X,\gamma) = \Big\{ f \in L^2(X,\gamma) : \sum_{k=1}^{\infty} k \|I_k(f)\|^2_{L^2(X,\gamma)} < \infty \Big\}.$$

Moreover, for every $f \in W^{1,2}(X,\gamma)$,

$$\int_X |\nabla_H f|_H^2 \, d\gamma = \sum_{k=1}^{\infty} k \int_X (I_k(f))^2 d\gamma,$$

and

(i) *for every* $f \in W^{1,2}(X,\gamma)$ *the sequence* $\sum_{k=0}^{n} I_k(f)$ *converges to* f *in* $W^{1,2}(X,\gamma)$, *(ii) the sequence* $\displaystyle\sum_{k=1}^{n}\sqrt{k}I_k$ *converges in* $\mathcal{L}(W^{1,2}(X,\gamma),L^2(X,\gamma))$.

Proof. Let $f\in L^2(X,\gamma)$. By Proposition 14.1.2, for every $k\in\mathbb{N}$, $I_k(f)\in D(L_2)$ and $L_2I_k(f)=-kI_k(f)$. Therefore,

$$\int_X|\nabla_H I_k(f)|_H^2d\gamma=-\int_X I_k(f)L_2I_k(f)\,d\gamma=k\int_X(I_k(f))^2d\gamma,\quad k\in\mathbb{N}. \tag{14.1.8}$$

Assume that $\sum_{k=1}^\infty k\|I_k(f)\|^2_{L^2(X,\gamma)}<\infty$. The sequence $s_n:=\sum_{k=0}^n I_k(f)$ converges to f in $L^2(X,\gamma)$. Moreover, $(\nabla_H s_n)$ is a Cauchy sequence in $L^2(X,\gamma;H)$. Indeed, $\nabla_H I_k(f)$ and $\nabla_H I_l(f)$ are orthogonal in $L^2(X,\gamma;H)$ for $l\neq k$, because

$$\int_X[\nabla_H I_k(f),\nabla_H I_l(f)]_Hd\gamma=-\int_X I_k(f)L_2I_l(f)d\gamma=l\int_X I_k(f)I_l(f)d\gamma=0.$$

Therefore, for n, $p\in\mathbb{N}$,

$$\Big\|\sum_{k=n}^{n+p}\nabla_H I_k(f)\Big\|^2_{L^2(X,\gamma;H)}=\sum_{k=n}^{n+p}\int_X|\nabla_H I_k(f)|_H^2d\gamma=\sum_{k=n}^{n+p}k\|I_k(f)\|^2_{L^2(X,\gamma)}.$$

So, $f\in W^{1,2}(X,\gamma)$, $s_n\to f$ in $W^{1,2}(X,\gamma)$, and

$$\int_X|\nabla_H f|_H^2d\gamma=\sum_{k=1}^\infty\int_X|\nabla_H I_k(f)|_H^2d\gamma=\sum_{k=1}^\infty k\|I_k(f)\|^2_{L^2(X,\gamma)}.$$

To prove the converse, first we take $f\in D(L_2)$. Then, by (13.2.5),

$$\begin{aligned}\int_X|\nabla_H f|_H^2d\gamma&=-\int_X fL_2f\,d\gamma=-\int_X\sum_{l=0}^\infty I_l(f)\sum_{k=0}^\infty I_k(L_2f)\,d\gamma\\&=-\int_X\sum_{k=0}^\infty I_k(f)I_k(L_2f)\,d\gamma,\end{aligned}$$

since $I_l(f)\in\mathcal{X}_l$, $I_k(L_2f)\in\mathcal{X}_k$. By Proposition 14.1.2,

$$\int_X|\nabla_H f|_H^2d\gamma=\sum_{k=1}^\infty k\int_X(I_k(f))^2d\gamma.$$

Comparing with (14.1.8), we obtain

$$\int_X|\nabla_H f|_H^2d\gamma=\sum_{k=1}^\infty\int_X|\nabla_H I_k(f)|_H^2d\gamma.$$

So, the mappings $T_n:D(L_2)\to L^2(X,\gamma)$, $T_nf=\sum_{k=1}^n\sqrt{k}I_k(f)$ satisfy

$$\exists L^2(X,\gamma)-\lim_{n\to\infty}T_nf,\quad\|T_nf\|_{L^2(X,\gamma)}\leq\|f\|_{W^{1,2}(X,\gamma)}.$$

Since $D(L_2)$ is dense in $W^{1,2}(X,\gamma)$, the sequence (T_nf) converges in $L^2(X,\gamma)$ for every $f\in W^{1,2}(X,\gamma)$. Since $\|T_nf\|^2_{L^2(X,\gamma)}=\sum_{k=1}^n k\|I_k(f)\|^2_{L^2(X,\gamma)}$, letting $n\to\infty$ we get $\sum_{k=1}^\infty k\|I_k(f)\|^2_{L^2(X,\gamma)}<\infty$. □

Proposition 14.1.3 may be recognized as the spectral decomposition of L. See e.g. [26, §VIII.3], in particular Theorem VIII.6. Accordingly, a functional calculus for L may be defined , namely for every $g : -\mathbb{N} \cup \{0\} \to \mathbb{R}$ we set

$$D(g(L)) := \Big\{ f \in L^2(X,\gamma) : \sum_{k=0}^{\infty} |g(-k)|^2 \|I_k(f)\|^2_{L^2(X,\gamma)} < \infty \Big\}$$

and

$$g(L)(f) = \sum_{k=0}^{\infty} g(-k) I_k(f), \quad f \in D(g(L)).$$

In particular, for $g(\xi) = (-\xi)^{1/2}$, Proposition 14.1.5 says that $D((-L)^{1/2}) = W^{1,2}(X,\gamma)$, and $\|(-L)^{1/2} f\|_{L^2(X,\gamma)} = \| \, |\nabla_H f|_H \|_{L^2(X,\gamma)}$ for every $f \in W^{1,2}(X,\gamma)$.

Corollary 14.1.6. *For every $d \in \mathbb{N}$, the embedding $W^{1,2}(\mathbb{R}^d, \gamma_d) \subset L^2(\mathbb{R}^d, \gamma_d)$ is compact.*

Proof. Let (f_n) be a bounded sequence in $W^{1,2}(\mathbb{R}^d, \gamma_d)$, say $\|f_n\|_{W^{1,2}(\mathbb{R}^d,\gamma_d)} \leq C$. Then there exists a subsequence (f_{n_j}) that converges weakly in $W^{1,2}(\mathbb{R}^d, \gamma_d)$ to an element $f \in W^{1,2}(\mathbb{R}^d, \gamma_d)$, that still satisfies $\|f\|_{W^{1,2}(\mathbb{R}^d,\gamma_d)} \leq C$. We claim that $f_{n_j} \to f$ in $L^2(\mathbb{R}^d, \gamma_d)$.

For every $N \in \mathbb{N}$ we have (norms and inner products are in $L^2(\mathbb{R}^d, \gamma_d)$)

$$\begin{aligned}
\|f_{n_j} - f\|^2 &= \sum_{k=0}^{\infty} \|I_k(f_{n_j} - f)\|^2 = \sum_{k=0}^{N-1} \|I_k(f_{n_j} - f)\|^2 + \sum_{k=N}^{\infty} \|I_k(f_{n_j} - f)\|^2 \\
&\leq \sum_{k=0}^{N-1} \sum_{\alpha \in (\mathbb{N}\cup\{0\})^d, \, |\alpha| = k} \langle f_{n_j} - f, H_\alpha \rangle^2 + \frac{1}{N} \sum_{k=N}^{\infty} k \|I_k(f_{n_j} - f)\|^2 \\
&\leq \sum_{k=0}^{N-1} \sum_{\alpha \in (\mathbb{N}\cup\{0\})^d, \, |\alpha| = k} \langle f_{n_j} - f, H_\alpha \rangle^2 + \frac{(2C)^2}{N}.
\end{aligned}$$

Fixed any $\varepsilon > 0$, let N be such that $4C^2/N \leq \varepsilon$. The sum in the right hand side consists of a finite number of summands, each of them goes to 0 as $n_j \to \infty$, therefore it does not exceed ε provided n_j is large enough. □

The argument in the proof of Corollary 14.1.6 does not work in infinite dimension, because in this case for every $k \in \mathbb{N}$ the Hermite polynomials H_α with $|\alpha| = k$ are infinitely many. In fact, $W^{1,2}(X,\gamma)$ is not compactly embedded in $L^2(X,\gamma)$ if H is infinite dimensional. It is sufficient to consider the Hermite polynomials H_α with $|\alpha| = 1$, namely the sequence of functions $(\hat{h}_j)$. Their $W^{1,2}(X,\gamma)$ norm is 2 but no subsequence converges in $L^2(X,\gamma)$ since $\|\hat{h}_i - \hat{h}_j\|^2_{L^2(X,\gamma)} = 2$ for $i \neq j$. The same argument shows that $D(L_2)$ is not compactly embedded in $L^2(X,\gamma)$.

14.2 Functional inequalities and asymptotic behaviour

In this section we present two important inequalities, the Logarithmic Sobolev and Poincaré inequality, that hold for functions in Sobolev spaces. The Ornstein-Uhlenbeck semigroup can be used as a tool in their proofs, and, in their turn, they are used to prove summability improving and asymptotic behaviour results for $T_p(t)f$, as $t \to \infty$.

We introduce the mean value $\overline{f}$ of any $f \in L^1(X,\gamma)$,

$$\overline{f} := \int_X f\, d\gamma.$$

If $f \in L^2(X,\gamma)$, $\overline{f} = I_0(f)$ is just the orthogonal projection of f on the kernel $\mathcal{X}_0$ of L_2, that consists of constant functions by Proposition 14.1.2 (see also Exercise 13.4). In any case, we have the following asymptotic behavior result.

Lemma 14.2.1. *For every $f \in C_b(X)$,*

$$\lim_{t\to\infty} T(t)f(x) = \overline{f}, \quad x \in X. \tag{14.2.1}$$

For every $f \in L^p(X,\gamma)$, $1 \le p < \infty$,

$$\lim_{t\to\infty} \|T_p(t)f - \overline{f}\|_{L^p(X,\gamma)} = 0. \tag{14.2.2}$$

Proof. The first assertion is an easy consequence of the definition (12.1.1) of $T(t)f$, through the Dominated Convergence Theorem. Still for $f \in C_b(X)$, we have that (14.2.2) holds again by the Dominated Convergence Theorem. Since $C_b(X)$ is dense in $L^p(X,\gamma)$ and the linear operators $f \mapsto T_p(t)f - \overline{f}$ belong to $\mathcal{L}(L^p(X,\gamma))$ and have norm not exeeding 2, the second assertion follows as well. $\square$

We shall see that the rate of convergence in (14.2.2) is exponential. This fact could be seen as a consequence of general results on analytic semigroups, but here we shall give a simpler and direct independent proof.

14.2.1 The Logarithmic Sobolev inequality

To begin with, we remark that no Sobolev embedding holds for nondegenerate Gaussian measures. Even in dimension 1, the function

$$f(\xi) = \frac{e^{\xi^2/4}}{1+\xi^2}, \quad \xi \in \mathbb{R},$$

belongs to $W^{1,2}(\mathbb{R},\gamma_1)$ but it does not belong to $L^{2+\varepsilon}(\mathbb{R},\gamma_1)$ for any $\varepsilon > 0$. This example may be adapted to show that for every $p \ge 1$, $W^{1,p}(\mathbb{R},\gamma_1)$ is not contained in $L^{p+\varepsilon}(\mathbb{R},\gamma_1)$ for any $\varepsilon > 0$, see Exercise 14.2.

The best result about summability properties in this context is the next Logarithmic Sobolev (Log-Sobolev) inequality. In the following we set $0 \log 0 = 0$.

Theorem 14.2.2. *Let $p>1$. For every $f\in C_b^1(X)$ we have*

$$\int_X |f|^p \log|f|\,d\gamma \le \|f\|^p_{L^p(X,\gamma)}\log\|f\|_{L^p(X,\gamma)} + \frac{p}{2}\int_X |f|^{p-2}|\nabla_H f|_H^2 1\!\!1_{\{f\neq 0\}}d\gamma. \tag{14.2.3}$$

Proof. As a first step, we consider a function f with positive infimum, say $f(x)\ge c>0$ for every x. In this case, also f^p belongs to $C_b^1(X)$, and $(T(t)f^p)(x)\ge c^p$ for every x, by (12.1.1). We define the function

$$F(t)=\int_X (T(t)f^p)\log(T(t)f^p)d\gamma, \quad t\ge 0.$$

Since L_2 is a sectorial operator (or, by Corollary 14.1.4), the function $t\mapsto T(t)f^p$ and $t\mapsto \log(T(t)f^p)$ belong to $C^1((0,\infty);L^2(X,\gamma))$. Consequently, their product is in $C^1((0,\infty);L^1(X,\gamma))$, $F\in C^1(0,\infty)$, and for every $t>0$ we have

$$\begin{aligned} F'(t) &= \int_X [L_2(T(t)f^p)\cdot\log(T(t)f^p) + L_2T(t)f^p]\,d\gamma \\ &= \int_X L_2(T(t)f^p)\cdot\log(T(t)f^p)\,d\gamma. \end{aligned} \tag{14.2.4}$$

The second equality is a consequence of the invariance of γ (Propositions 12.1.5(iii) and 11.3.1). Moreover, $t\mapsto T(t)f^p(x)$ and $t\mapsto \log(T(t)f^p)(x)$ are continuous for every x and bounded by constants independent of x. It follows that F is continuous up to $t=0$, and $F(t)-F(0)=\int_0^t F'(s)ds$. Integrating in the right hand side of (14.2.4) and using (13.2.5) with f replaced by $T(t)f^p$, g replaced by $\log(T(t)f^p)$, we obtain

$$\begin{aligned} F'(t) &= -\int_X [\nabla_H T(s)f^p, \nabla_H \log(T(s)f^p))]_H d\gamma \\ &= -\int_X \frac{1}{T(t)f^p}|\nabla_H(T(t)f^p)|_H^2\,d\gamma. \end{aligned}$$

We recall that for every $x\in X$, $|\nabla_H(T(t)f^p)(x)|_H \le e^{-t}T(t)(|\nabla_H f^p|_H)(x)$ (see Proposition 12.1.6). So,

$$F'(t)\ge -e^{-2t}\int_X \frac{1}{T(t)f^p}(T(t)(|\nabla_H f^p|_H))^2\,d\gamma. \tag{14.2.5}$$

Moreover, the Hölder inequality in (12.1.1) yields

$$|T(t)(\varphi_1\varphi_2)(x)| \le [(T(t)\varphi_1^2)(x)]^{1/2}[(T(t)\varphi_2^2)(x)]^{1/2}, \quad \varphi_i\in C_b(X),\ x\in X.$$

We use this estimate with $\varphi_1 = |\nabla_H f^p|_H/f^{p/2}$, $\varphi_2=f^{p/2}$ and we obtain

$$T(t)(|\nabla_H f^p|_H) = T(t)\bigg(\frac{|\nabla_H f^p|_H}{f^{p/2}}\,f^{p/2}\bigg) \le \bigg(T(t)\bigg(\frac{|\nabla_H f^p|_H^2}{f^p}\bigg)\bigg)^{1/2}(T(t)f^p)^{1/2}.$$

Replacing in (14.2.5) and using (12.1.2), we get

$$\begin{aligned} F'(t) &\ge -e^{-2t}\int_X T(t)\bigg(\frac{|\nabla_H f^p|_H^2}{f^p}\bigg)d\gamma = -e^{-2t}\int_X \frac{|\nabla_H f^p|_H^2}{f^p}\,d\gamma \\ &= -p^2e^{-2t}\int_X f^{p-2}|\nabla_H f|_H^2 d\gamma. \end{aligned}$$

Integrating with respect to time in $(0,t)$ yields

$$\begin{aligned}\int_X (T(t)f^p)\log(T(t)f^p)\,d\gamma - \int_X f^p\log(f^p)d\gamma &= F(t)-F(0)\\ &\geq \frac{p^2}{2}(e^{-2t}-1)\int_X f^{p-2}|\nabla_H f|_H^2 d\gamma.\end{aligned} \tag{14.2.6}$$

Now we let $t\to\infty$. By Lemma 14.2.1, $\lim_{t\to\infty}(T(t)f^p)(x) = \overline{f^p} = \|f\|_{L^p}^p$, and consequently $\lim_{t\to\infty}\log((T(t)f^p)(x)) = p\log(\|f\|_{L^p})$, for every $x\in X$. Moreover, $c^p \leq |(T(t)f^p)(x)| \leq \|f\|_\infty^p$, for every x. By the Dominated Convergence Theorem, the left hand side of (14.2.6) converges to $p\|f\|_{L^p}^p\log(\|f\|_{L^p}) - p\int_X f^p\log f\,d\gamma$ as $t\to\infty$, and (14.2.3) follows.

For $f\in C_b^1(X)$ we approximate $|f|$ in $W^{1,p}(X,\gamma)$ and pointwise by the sequence $f_n=\sqrt{f^2+1/n}$, see Exercise 14.3. Applying (14.2.3) to each f_n we get

$$\begin{aligned}\int_X f_n^p\log f_n\,d\gamma - \|f_n\|_{L^p(X,\gamma)}^p\log\|f_n\|_{L^p(X,\gamma)} &\leq \frac{p}{2}\int_X f^2(f^2+1/n)^{p/2-2}|\nabla_H f|_H^2 d\gamma\\ &\leq \frac{p}{2}\int_X \mathbb{1}_{\{f\neq 0\}}(f^2+1/n)^{p/2-1}|\nabla_H f|_H^2 d\gamma,\end{aligned}$$

and letting $n\to\infty$ yields that f satisfies (14.2.3). Notice that the last integral goes to $\int_X \mathbb{1}_{\{f\neq 0\}}|f|^{p-2}|\nabla_H f|_H^2 d\gamma$ by the Monotone Convergence Theorem, even if $p<2$. □

Corollary 14.2.3. *Let $p\geq 2$. For every $f\in W^{1,p}(X,\gamma)$ we have*

$$\int_X |f|^p\log|f|\,d\gamma \leq \|f\|_{L^p(X,\gamma)}^p\log\|f\|_{L^p(X,\gamma)} + \frac{p}{2}\int_X |f|^{p-2}|\nabla_H f|_H^2 d\gamma. \tag{14.2.7}$$

Proof. We approximate f by a sequence of $\mathcal{F}C_b^1(X)$ functions (f_n) that converges in $W^{1,p}(X,\gamma)$ and pointwise a.e. to f. We apply (14.2.3) to each f_n, and then we let $n\to\infty$. Recalling that $\nabla_H f_n = 0$ a.e. in the set $\{f_n=0\}$ (see Exercise 10.3), we get

$$\int_X |f_n|^{p-2}|\nabla_H f_n|_H^2\mathbb{1}_{\{f_n\neq 0\}}\,d\gamma = \int_X |f_n|^{p-2}|\nabla_H f_n|_H^2\,d\gamma$$

for every n, and

$$\begin{aligned}\int_X |f|^p\log|f|\,d\gamma &\leq \liminf_{n\to\infty}\int_X |f_n|^p\log|f_n|\,d\gamma\\ &\leq \liminf_{n\to\infty}\left(\|f_n\|_{L^p(X,\gamma)}^p\log\|f_n\|_{L^p(X,\gamma)} + \frac{p}{2}\int_X |f_n|^{p-2}|\nabla_H f_n|_H^2 d\gamma\right)\\ &= \|f\|_{L^p(X,\gamma)}^p\log\|f\|_{L^p(X,\gamma)} + \frac{p}{2}\int_X |f|^{p-2}|\nabla_H f|_H^2 d\gamma\end{aligned}$$

□

Note that for $1 < p < 2$ the function $1\!\!1_{\{f\neq 0\}}|f|^{p-2}|\nabla_H f|_H^2$ does not necessarily belong to $L^1(X,\gamma)$ for $f \in W^{1,p}(X,\gamma)$, and in this case (14.2.3) is not meaningful. Take for instance $X = \mathbb{R}$ and $f(x) = x^{1/p}$ for $0 < x < 1$, $f(x) = 0$ for $x \le 0$, $f(x) = 1$ for $x \ge 1$. Then $f \in W^{1,p}(\mathbb{R},\gamma_1)$ but $\int_{\mathbb{R}} |f|^{p-2}|\nabla_H f|_H^2 1\!\!1_{\{f\neq 0\}} d\gamma_1 = \infty$.

Instead (see [24]), it is possible to show that for any $p \in (1,\infty)$

$$-(p-1)\int_X |f|^{p-2}|\nabla_H f|_H^2 1\!\!1_{\{f\neq 0\}} d\gamma = \int_X f|f|^{p-2} L_p f\, d\gamma \tag{14.2.8}$$

for every $f \in D(L_p)$, so that $\int_X |f|^{p-2}|\nabla_H f|_H^2 1\!\!1_{\{f\neq 0\}} d\gamma \le C_p \|f\|_{D(L_p)}$. See Exercise 14.4. So, if $f \in D(L_p)$ (14.2.3) may be rewritten as

$$\int_X |f|^p \log|f|\, d\gamma \le \|f\|^p_{L^p(X,\gamma)} \log\|f\|_{L^p(X,\gamma)} - \frac{p}{2(p-1)}\int_X f|f|^{p-2} L_p f\, d\gamma. \tag{14.2.9}$$

An important consequence of the Log-Sobolev inequality is the next summability improving property of $T(t)$, called *hypercontractivity.*

Theorem 14.2.4. *Let $p > 1$, and set $p(t) = e^{2t}(p-1)+1$ for $t > 0$. Then $T_p(t)f \in L^{p(t)}(X,\gamma)$ for every $f \in L^p(X,\gamma)$, and*

$$\|T_p(t)f\|_{L^{p(t)}(X,\gamma)} \le \|f\|_{L^p(X,\gamma)}, \quad t > 0. \tag{14.2.10}$$

Proof. Let us prove that (14.2.10) holds for every $f \in \Sigma$ with positive infimum (the set Σ was introduced at the beginning of Section 13.2, and it is dense in $L^p(X,\gamma)$). For such f's, since they belong to $D(L_q)$ for any q, we have that $T_p(f) = T(t)f$ and we can drop the idex p in the semigroup. We shall show that the function

$$\beta(t) := \|T(t)f\|_{L^{p(t)}(X,\gamma)}, \quad t \ge 0$$

decreases in $[0,\infty)$.

It is easily seen that β is continuous in $[0,\infty)$. Our aim is to show that $\beta \in C^1(0,\infty)$, and $\beta'(t) \le 0$ for every $t > 0$. Indeed, by Proposition 13.2.1 we know that for every $x \in X$ the function $t \mapsto T(t)f(x)$ belongs to $C^1(0,\infty)$, as well as $t \mapsto (T(t)f(x))^{p(t)}$, and

$$\begin{aligned}\frac{d}{dt}(T(t)f(x))^{p(t)} &= p'(t)(T(t)f(x))^{p(t)}\log(T(t)f(x)) + p(t)(T(t)f(x))^{p(t)-1}\frac{d}{dt}(T(t)f(x))\\ &= p'(t)(T(t)f(x))^{p(t)}\log(T(t)f(x)) + p(t)(T(t)f(x))^{p(t)-1}(L_2T(t)f(x)).\end{aligned}$$

We have used the operator L_2, but any other L_q can be equivalently used. Moreover, $|d/dt(T(t)f(x))^{p(t)}|$ is bounded by $c(t)(1+\|x\|)$ for some continuous function $c(\cdot)$. So, $t \mapsto \int_X |T(t)f|^{p(t)} d\gamma$ is continuously differentiable, with derivative equal to

$$p'(t)\int_X (T(t)f)^{p(t)}\log(T(t)f)d\gamma - p(t)(p(t)-1)\int_X T(t)f)^{p(t)-2}|\nabla_H T(t)f|_H^2\, d\gamma.$$

It follows that β is differentiable and

$$\beta'(t) = \beta(t)\Big[-\frac{p'(t)}{p(t)^2}\log\int_X (T(t)f)^{p(t)}d\gamma + +\frac{p'(t)}{p(t)}\frac{\int_X (T(t)f)^{p(t)}\log(T(t)f)d\gamma}{\int_X (T(t)f)^{p(t)}d\gamma}$$
$$- (p(t)-1)\frac{\int_X (T(t)f)^{p(t)-2}|\nabla_H T(t)f|_H^2\,d\gamma}{\int_X (T(t)f)^{p(t)}d\gamma}\Big].$$

The Logarithmic Sobolev inequality (14.2.3) yields

$$\int_X (T(t)f)^{p(t)}\log(T(t)f)d\gamma \le$$
$$\le \frac{1}{p(t)}\int_X (T(t)f)^{p(t)}d\gamma \log\int_X (T(t)f)^{p(t)}d\gamma + \frac{p(t)}{2}\int_X (T(t)f)^{p(t)-2}|\nabla_H T(t)f|_H^2\,d\gamma,$$

and replacing we obtain

$$\beta'(t) \le \left(\frac{p'(t)}{2} - (p(t)-1)\right)\frac{\int_X (T(t)f)^{p(t)-2}|\nabla_H T(t)f|_H^2\,d\gamma}{\int_X (T(t)f)^{p(t)}d\gamma}.$$

The function $p(t)$ was chosen in such a way that $p'(t) = 2(p(t)-1)$. Therefore, $\beta'(t) \le 0$, and (14.2.10) follows.

Let now $f \in \Sigma$ and set $f_n = (f^2 + 1/n)^{1/2}$. For every $x \in X$ and $n \in \mathbb{N}$ we have, by (12.1.1), $|(T(t)f)(x)| \le (T(t)|f|)(x) \le (T(t)f_n)(x)$, so that

$$\|T(t)f\|_{L^{p(t)}(X,\gamma)} \le \liminf_{n\to\infty}\|T(t)f_n\|_{L^{p(t)}(X,\gamma)} \le \liminf_{n\to\infty}\|f_n\|_{L^p(X,\gamma)} = \|f\|_{L^p(X,\gamma)},$$

and (14.2.10) holds. Since Σ is dense in $L^p(X,\gamma)$, (14.2.10) holds for every $f \in L^p(X,\gamma)$. □

We notice that in the proof of Theorem 14.2.4 we have not used specific properties of the Ornstein-Uhlenbeck semigroup: the main ingredients were the integration by parts formula, namely the fact that the infinitesimal generator L_2 is the operator associated to the quadratic form (13.2.4), and the Log-Sobolev inequality (14.2.3) for good functions. In fact, the proof may be extended to a large class of semigroups in spaces $L^p(\Omega,\mu)$, (Ω,μ) being a probability space, see [16]. In [16] a sort of converse is proved, namely under suitable assumptions if a semigroup $T(t)$ is a contraction from $L^p(\Omega,\mu)$ to $L^{q(t)}(\Omega,\mu)$, with q differentiable and increasing, then a logarithmic Sobolev inequality of the type (14.2.9) holds in the domain of the infinitesimal generator of $T(t)$ in $L^p(X,\mu)$.

14.2.2 The Poincaré inequality and the asymptotic behaviour

The Poincaré inequality is the following.

Theorem 14.2.5. *For every $f \in W^{1,2}(X,\gamma)$,*

$$\int_X (f-\overline{f})^2 d\gamma \le \int_X |\nabla_H f|_H^2 d\gamma. \qquad (14.2.11)$$

Proof. There are several proofs of (14.2.11). One of them follows from Theorem 14.2.4, see Exercise 14.5. The simplest proof uses the Wiener Chaos decomposition. By (8.1.9) and (14.1.3), for every $f \in D(L_2)$ we have $f = \sum_{k=0}^{\infty} I_k(f)$ and $L_2 f = \sum_{k=1}^{\infty} -kI_k(f)$, where both series converge in $L^2(X,\gamma)$. Using (13.2.5) and these representation formulas we obtain

$$\begin{aligned}\int_X |\nabla_H f|_H^2\, d\gamma &= -\int_X f\, L_2 f\, d\gamma \\ &= \sum_{k=1}^{\infty} k\|I_k(f)\|^2_{L^2(X,\gamma)} \geq \sum_{k=1}^{\infty} \|I_k(f)\|^2_{L^2(X,\gamma)} \\ &= \|f\|^2_{L^2(X,\gamma)} - \|I_0(f)\|^2_{L^2(X,\gamma)} \\ &= \|f\|^2_{L^2(X,\gamma)} - \overline{f}^2 = \|f-\overline{f}\|^2_{L^2(X,\gamma)}.\end{aligned}$$

Since $D(L_2)$ is dense in $W^{1,2}(X,\gamma)$, (14.2.11) follows. □

An immediate consequence of the Poincaré inequality is the following: if $f \in W^{1,2}(X,\gamma)$ and $\nabla_H f \equiv 0$, then f is constant a.e. (compare with Exercise 13.4).

An L^p version of (14.2.11) is

$$\int_X |f-\overline{f}|^p \gamma \leq c_p \int_X |\nabla_H f|_H^p d\gamma. \tag{14.2.12}$$

that holds for $p > 2$, $f \in W^{1,p}(X,\gamma)$ (Exercise 14.6).

Now we are able to improve Lemma 14.2.1, specifying the decay rate of $T_q(t)f$ to $\overline{f}$.

Proposition 14.2.6. *For every $q > 1$ there exists $c_q > 0$ such that $c_2 = 1$ and for every $f \in L^q(X,\gamma)$,*

$$\|T_q(t)f - \overline{f}\|_{L^q(X,\gamma)} \leq c_q e^{-t}\|f\|_{L^q(X,\gamma)}, \quad t > 0. \tag{14.2.13}$$

Proof. As a first step, we prove that the statement holds for $q = 2$. By (14.1.5), for every $f \in L^2(X,\gamma)$ and $t > 0$ we have $T(t)f = \sum_{k=0}^{\infty} e^{-kt} I_k(f)$. We already know that for $k = 0$, $I_0(f) = \overline{f}$. Therefore,

$$\|T(t)f - \overline{f}\|^2_{L^2(X,\gamma)} = \left\| \sum_{k=1}^{\infty} e^{-kt} I_k(f) \right\|^2_{L^2(X,\gamma)} \leq e^{-2t} \sum_{k=1}^{\infty} \|I_k(f)\|^2_{L^2(X,\gamma)} \leq e^{-2t}\|f\|^2_{L^2(X,\gamma)}.$$

For $q \neq 2$ it is enough to prove that (14.2.13) holds for every $f \in C_b(X)$. For such functions we have $T_q(t)f = T(t)f$ for every $t > 0$.

Let $q > 2$. Set $\tau = \log\sqrt{q-1}$, so that $e^{2\tau} + 1 = q$, and by Theorem 14.2.4 $T_q(\tau)$ is a contraction from $L^2(X,\gamma)$ to $L^q(X,\gamma)$. Then, for every $t \geq \tau$,

$$\begin{aligned}\|T(t)f - \overline{f}\|_{L^q(X,\gamma)} &= \|T(\tau)(T(t-\tau)f - \overline{f})\|_{L^q(X,\gamma)} \\ &\leq \|T(t-\tau)f - \overline{f}\|_{L^2(X,\gamma)} \quad \text{(by (14.2.10))} \\ &\leq e^{-(t-\tau)}\|f\|_{L^2(X,\gamma)} \quad \text{(by (14.2.13) with } q=2) \\ &\leq e^{-(t-\tau)}\|f\|_{L^q(X,\gamma)} \quad \text{(by the Hölder inequality)} \\ &= \sqrt{q-1}\, e^{-t}\|f\|_{L^q(X,\gamma)},\end{aligned}$$

while for $t \in (0,\tau)$ we have

$$\|T(t)f - \overline{f}\|_{L^q(X,\gamma)} \le 2\|f\|_{L^q(X,\gamma)} = 2e^t e^{-t}\|f\|_{L^q(X,\gamma)} \le 2\sqrt{q-1}\, e^{-t}\|f\|_{L^q(X,\gamma)}.$$

So, (14.2.13) holds with $c_q = 2\sqrt{q-1}$.

Let now $q<2$ and set $\tau = -\log\sqrt{q-1}$, so that $e^{2\tau}(q-1)+1=2$, and by Theorem 14.2.4 $T_q(\tau)$ is a contraction from $L^q(X,\gamma)$ to $L^2(X,\gamma)$. For every $t\ge\tau$ we have

$$\begin{aligned}
\|T(t)f - \overline{f}\|_{L^q(X,\gamma)} &\le \|T(t)f - \overline{f}\|_{L^2(X,\gamma)} \quad \text{(by the Hölder inequality)}\\
&= \|T(t-\tau)(T(\tau)f - \overline{T(\tau)f})\|_{L^2(X,\gamma)}\\
&\le e^{-(t-\tau)}\|T(\tau)f\|_{L^2(X,\gamma)} \quad \text{(by (14.2.13) with } q=2)\\
&\le e^{-(t-\tau)}\|f\|_{L^q(X,\gamma)} \quad \text{(by (14.2.10))}\\
&= \frac{1}{\sqrt{q-1}}\, e^{-t}\|f\|_{L^q(X,\gamma)},
\end{aligned}$$

while for $t \in (0,\tau)$ we have, as before,

$$\|T(t)f - \overline{f}\|_{L^q(X,\gamma)} \le 2\|f\|_{L^q(X,\gamma)} = 2e^t e^{-t}\|f\|_{L^q(X,\gamma)} \le \frac{2}{\sqrt{q-1}}\, e^{-t}\|f\|_{L^q(X,\gamma)}.$$

So, (14.2.13) holds with $c_q = 2/\sqrt{q-1}$. □

In fact, estimate (14.2.13) could be deduced also by the general theory of (analytic) semigroups, but we prefer to give a simpler self-contained proof.

14.3 Exercises

Exercise 14.1. Prove the equality (14.1.4).

Exercise 14.2. Show that for every $p\ge 1$, $W^{1,p}(\mathbb{R},\gamma_1)$ is not contained in $L^{p+\varepsilon}(\mathbb{R},\gamma_1)$ for any $\varepsilon>0$.

Exercise 14.3. Prove that for every $f \in W^{1,p}(X,\gamma)$ the sequence $f_n = \sqrt{f^2+1/n}$ converges to $|f|$ in $W^{1,p}(X,\gamma)$.

Exercise 14.4. Prove that for every $p>1$ and $f\in D(L_p)$, (14.2.8) holds.
Hint: for every $f\in\Sigma$ and $\varepsilon>0$, apply formula (13.2.5) with $g = f(f^2+\varepsilon)^{1-p/2}$ and then let $\varepsilon\to 0$.

Exercise 14.5. Prove the Poincaré inequality (14.2.11) for functions $f\in C^1_b(X)$ such that $\overline{f}=0$, in the following alternative way: apply (14.2.7) with $p=2$ to the functions $f_\varepsilon := 1+\varepsilon f$, for $\varepsilon>0$, and then divide by ε^2 and let $\varepsilon\to 0$.

Exercise 14.6. Prove that (14.2.12) holds for every $f\in W^{1,p}(X,\gamma)$ with $p>2$.
Hint: For $p\le 4$, apply (14.2.11) to $|f|^{p/2}$ and estimate $(\int_X |f|^{p/2}d\gamma)^2$ by $\|f\|^p_{L^2(X,\gamma)}$, then

estimate $(\int_X |\nabla_H f|_H^2 |f|^{p/2-1} d\gamma)^2$ by $\varepsilon \int_X |f|^p d\gamma + C(\varepsilon)(\int_X |\nabla_H f|_H^p d\gamma)$. Taking ε small, arrive at

$$\int_X |f|^p d\gamma \le \|f\|_{L^2(X,\gamma)}^p + K \int_X |\nabla_H f|_H^p d\gamma.$$

(14.2.12) follows applying such estimate to $f - \overline{f}$, and using (14.2.11) to estimate $\|f - \overline{f}\|_{L^2(X,\gamma)}$. For $p \ge 4$, use a bootstrap procedure.

Bibliography

[1] S. Banach: *Théorie des operérations linéaires,* Monografje Matematyczne, Warsaw, 1932.

[2] F. Baudoin: *Diffusion Processes and Stochastic Calculus,* E.M.S. Textbooks in Mathematics, 2014.

[3] V. I. Bogachev: *Gaussian Measures.* American Mathematical Society, 1998.

[4] H. Brezis: *Functional Analysis, Sobolev spaces and partial differential equations,* Springer, 2011.

[5] G. Da Prato: *An introduction to infinite-dimensional analysis,* Springer, 2006.

[6] G. Da Prato, J. Zabczyk: *Second order partial differential equations in Hilbert spaces,* London Mathematical Society Lecture Note Series, 293, Cambridge U. P., 2002.

[7] B. Davies: *Heat kernels and spectral theory.* Cambridge University Press, 1989.

[8] J. Diestel, J. J. Uhl: *Vector measures,* Mathematical Surveys and Monographs, 15. American Mathematical Society, 1977.

[9] R. M. Dudley: *Real Analysis and Probability,* Cambridge University Press, 2004.

[10] N. Dunford, J. T. Schwartz: *Linear operators I*, Wiley, 1958.

[11] N. Dunford, J. T. Schwartz: *Linear operators II*, Wiley, 1963.

[12] R.E. Edwards: *Functional Analysis.Theory and applications*, Holt, Rinehard and Winston, 1965 and Dover, 1995.

[13] K. Engel, R. Nagel: *One-parameter semigroups for linear evolution equations,* Graduate Texts in Mathematics, 194, Springer, 1999.

[14] D. Feyel, A. De la Pradelle: *Hausdorff measures on the Wiener space*, Potential Analysis, **1** (1992), 177-189.

[15] M. Fukushima, Y. Oshima, M. Takeda: *Dirichlet forms and symmetric Markov processes,* de Gruyter Studies in Mathematics, 19, de Gruyter, 1994.

[16] L. Gross: *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), 1061-1083.

[17] E. Kreyszig: *Introductory functional analysis with applications,* Wiley, 1978.

[18] S. Lang: Linear Algebra, Springer, 1987.

[19] M. Ledoux: Isoperimetry and Gaussian analysis, in: *Lectures on Probability Theory and Statistics,* Saint Flour, 1994, Lecture Notes in Mathematics, 1648, Springer, 1996, 165-294.

[20] M. Ledoux: *The concentration of measure phenomenon.* Mathematical Surveys and Monographs, 89. American Mathematical Society, 2001.

[21] A. Lunardi: *Analytic semigroups and optimal regularity in parabolic problems,* Birkhäuser, 1995. Second edition, Modern Birkhäuser Classics, 2013.

[22] G. Metafune, D. Pallara, E. Priola: *Spectrum of Ornstein–Uhlenbeck operators in L^p spaces with respect to invariant measures*, J. Funct. Anal. **(196)** (2002) 40–60.

[23] G. Metafune, J. Prüss, A. Rhandi, R. Schnaubelt: *The domain of the Ornstein-Uhlenbeck operator on a L^p space with invariant measure*, Ann. Sc. Norm. Sup. Pisa Cl. Sci. (5) **1** (2002), 471-485.

[24] G. Metafune, C. Spina: *An integration by parts formula in Sobolev spaces*, Mediterr. J. Math. **5** (2008), 357-369.

[25] P.A. Meyer: *Transformations de Riesz pour les lois gaussiennes*, Lect. Notes in Math. **1059** (1984), 179-193.

[26] M. Reed, B. Simon: *Methods of modern mathematical physics. I: Functional analysis*, Academic Press, Inc., 1980.

[27] N. N. Vakhania, V. I. Tarieladze, S. A. Chobanyan: *Probability distribution in Banach spaces,* Kluwer, 1987.

[28] N. Wiener: The homogeneous chaos, *Amer. J. Math.* **60** (1938), 897-936.

[29] K. Yosida: *Functional Analysis*, 6^{th} ed., Springer, 1980.