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 on 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 $\{\hat{h}_j :\ j \in \mathbb{N}\}$ of X^*_γ. By Lemma 3.1.8, we may assume that each $\hat{h}_j$ belongs to $j(X^*)$. We recall that the set $\{h_j :\ j \in \mathbb{N}\}$, with $h_j = R_\gamma \hat{h}_j$, is an orthonormal basis of H. Notice also that, arguing as in Theorem 2.1.1 we obtain $\mathscr{E}(X) = \mathscr{E}\big(X, \{\hat{h}_j\}_{j\in\mathbb{N}}\big)$. We define

$$P_n x = \sum_{j=1}^{n} \hat{h}_j(x) h_j, \quad n \in \mathbb{N},\ x \in X. \tag{7.1.1}$$

Note that every P_n is a projection, since by (2.3.6) $\hat{h}_j(h_i) = \delta_{ij}$. Moreover, if $x \in H$, then $\hat{h}_j(x) = [x, h_j]_H$ so that $P_n x$ is just a natural extension to X of the orthogonal projection of H on span $\{h_1, \dots, h_n\}$.

We state (without proof) a deep result on finite dimensional approximations.

Theorem 7.1.2. *For γ-a.e. $x \in X$, $\lim_{n\to\infty} P_n x = x$.*

The proof of theorem 7.1.2 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 $\{e_j : j \in \mathbb{N}\}$ of X consisting of eigenvectors of Q, $Qe_k = \lambda_k e_k$. We may assume, without loss of generality, that Q is nondegenerate, i.e. $\lambda_k > 0$ for any $k \in \mathbb{N}$. Then $\{h_j = \lambda_j^{1/2} e_j,\ j \in \mathbb{N}\}$ is an orthonormal basis of H and we have $\hat{h}_j(x) = \langle x, e_j\rangle / \lambda_j^{1/2}$ for every $x \in X$. Indeed, for every $x \in H$,

$$\hat{h}_j(x) h_j = [x, h_j]_H h_j = \langle Q^{-1/2} x, Q^{-1/2} Q^{1/2} e_j\rangle Q^{1/2} e_j = \langle x, e_j\rangle e_j. \tag{7.1.2}$$

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 $P_n x$ is the orthogonal (in X) projection of x on $\operatorname{span}\{e_1, \dots, e_n\} = \operatorname{span}\{h_1, \dots, h_n\}$, which goes to x as $n \to \infty$ for every $x \in X$. The case when Q is degenerate follows similarly using the fact that H is dense in the span of the eigenvectors e_k associated with nonzero eigenvalues $\lambda_k > 0$.

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 1.1.9. □

Remark 7.2.2. Using approximations by simple functions, we have that (7.2.1) implies

$$\int_\Omega gX\, d\mathbb{P} = \int_\Omega g\, \mathbb{E}(X|\mathscr{G}) d\mathbb{P}$$

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. *If* $X \leq Y$*, then* $\mathbb{E}(X|\mathscr{G}) \leq \mathbb{E}(Y|\mathscr{G})$*; in particular, if* $X \geq 0$*, then* $\mathbb{E}(X|\mathscr{G}) \geq 0$*.*
4. *For any* X, Y *and* $\alpha, \beta \in \mathbb{R}$*,* $\mathbb{E}(\alpha X + \beta Y|\mathscr{G}) = \alpha\mathbb{E}(X|\mathscr{G}) + \beta\mathbb{E}(Y|\mathscr{G})$*.*
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, $p > 1$.

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,*

$$\mathbb{E}(\varphi \circ X|\mathscr{G}) \geq \varphi \circ \mathbb{E}(X|\mathscr{G}). \tag{7.2.2}$$

Proof. As φ is convex, we have that for any $x, y \in \mathbb{R}$

$$\varphi(x) \geq \varphi(y) + \varphi'(y)(x - y).$$

We use this inequality with $x = X$ and $y = \mathbb{E}(X|\mathscr{G})$ and we obtain

$$\varphi(X) \geq \varphi(\mathbb{E}(X|\mathscr{G})) + \varphi'(\mathbb{E}(X|\mathscr{G}))(X - \mathbb{E}(X|\mathscr{G})). \tag{7.2.3}$$

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 that

$$\mathbb{E}\Big(\varphi'(\mathbb{E}(X|\mathscr{G}))\big(X - \mathbb{E}(X|\mathscr{G})\big)\Big|\mathscr{G}\Big) = \varphi'(\mathbb{E}(X|\mathscr{G})) \cdot \mathbb{E}\big(X - \mathbb{E}(X|\mathscr{G})\big|\mathscr{G}\big) = 0.$$

Then, taking conditional expectation in (7.2.3), we have

$$\mathbb{E}(\varphi(X)|\mathscr{G}) \geq \mathbb{E}(\varphi(\mathbb{E}(X|\mathscr{G}))|\mathscr{G}) + \mathbb{E}\Big(\varphi'(\mathbb{E}(X|\mathscr{G}))(X - \mathbb{E}(X|\mathscr{G}))\Big|\mathscr{G}\Big) = \varphi(\mathbb{E}(X|\mathscr{G})).$$

□

Corollary 7.2.5. *Let $X \in L^p(\Omega, \mathscr{F}, \mathbb{P})$, $p > 1$, 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.*

Proof. Theorem 7.2.4 with $\varphi(x) = |x|^p$ yields

$$\int_\Omega |\mathbb{E}(X|\mathscr{G})|^p \, d\mathbb{P} \le \int_\Omega \mathbb{E}(|X|^p|\mathscr{G}) \, d\mathbb{P} = \int_\Omega |X|^p \, d\mathbb{P}.$$

□

Notice that the properties of the conditional expectation listed in Proposition 7.2.3 hold also in $L^p(X, \gamma)$, $p \ge 1$.

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 on 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 on X with image in 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 $\zeta \in X^*$

$$\begin{aligned}\widehat{\gamma_F}(\zeta) &= \int_X \exp\{i\zeta(P_F(x))\} \, \gamma(dx) = \int_X \exp\{iP_F^*\zeta(x)\} \, \gamma(dx) \\ &= \exp\left\{-\frac{1}{2}B_\gamma(P_F^*\zeta, P_F^*\zeta)\right\},\end{aligned}$$

then γ_F is a centred Gaussian measure by Corollary 2.2.7(i), with

$$\begin{aligned}B_{\gamma_F}(\zeta_1, \zeta_2) =& B_\gamma(P_F^*\zeta_1, P_F^*\zeta_2) = \int_X P_F^*\zeta_1(x)P_F^*\zeta_2(x)\gamma(dx) \\ &= \int_X \zeta_1(P_F x)\zeta_2(P_F x) \, \gamma(dx) = \int_X \zeta_1(z)\zeta_2(z) \, \gamma_F(dz) \\ &= \int_F \zeta_1(z)\zeta_2(z) \, \gamma_F(dz) = \langle \zeta_1, \zeta_2 \rangle_{L^2(F, \gamma_F)},\end{aligned} \tag{7.3.1}$$

$\gamma_F(F) = 1$, 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 $F = \text{span}\{h_1, \dots, h_n\}$ with $h_1, \dots, h_n$ 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 $\xi\in X^*$

$$\begin{aligned}\widehat{\tilde\gamma_F\otimes\tilde\gamma_F^\perp}(\xi)=&\int_{F\times X_F}\exp\{i\xi(z+y)\tilde\gamma_F\otimes\tilde\gamma_F^\perp(d(z,y))\\ =&\int_F\exp\{i\xi(z)\}\gamma_F(dz)\cdot\int_{X_F}\exp\{i\xi(y)\}\gamma_F^\perp(dy)\\ =&\exp\left\{-\frac12\Big(B_{\gamma_F}(\xi,\xi)+B_{\gamma_F^\perp}(\xi,\xi)\Big)\right\}.\end{aligned}$$

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 1.1.14) $A_z=\{y\in X_F:\ (z,y)\in A\}$, 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*

$$(\mathbb{E}_nf)(x)=\int_X f(P_nx+(I-P_n)y)\gamma(dy),\quad x\in X. \tag{7.4.1}$$

Moreover, the conditional expectation is a contraction, i.e.

$$\|\mathbb{E}_nf\|_{L^p(X,\gamma)}\le\|f\|_{L^p(X,\gamma)}. \tag{7.4.2}$$

Proof. Let us define

$$f_n(x) = \int_X f(P_n x + (I - P_n)y)\gamma(dy), \quad n \in \mathbb{N},\ x \in X.$$

Using the factorisation $\gamma = \tilde{\gamma}_F \otimes \tilde{\gamma}_F^{\perp}$, we may also write

$$f_n(x) = \int_{X_F} f(P_n x + y)\tilde{\gamma}_F^{\perp}(dy).$$

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 $f_n = \mathbb{E}(f|\Sigma_n)$.

The contractivity estimate $\|\mathbb{E}_n f\|_p \le \|f\|_p$ easily follows from (7.4.1), and the properties of conditional expectation:

$$\begin{aligned}
\|\mathbb{E}_n f\|^p_{L^p(X,\gamma)} &= \int_X |\mathbb{E}_n f(x)|^p\, \gamma(dx) = \int_X \Big| \int_X f(P_n x + (I - P_n)y)\gamma(dy)\Big|^p \gamma(dx) \\
&\le \int_X \mathbb{E}_n(|f|^p)\gamma(dx) = \|f\|^p_{L^p(X,\gamma)}.
\end{aligned}$$

□

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*

$$E = \Big\{x \in X : \ f(x) := (f_n(x)) \in B\Big\}. \tag{7.4.3}$$

Proof. For every fixed sequence $(f_n) \subset X^*$ the sets of the form (7.4.3) are a σ-algebra, see Exercise 7.5. Then, the family of the sets as in (7.4.3) 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 (7.4.3) 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 (7.4.3) with $B \in \mathscr{G}_f$ is contained in $\mathscr{E}(X)$ for every f as above. Then $\mathscr{E}(X) \subset \mathscr{F}$ and the proof is complete. □

Remark 7.4.3. We notice that the σ–algebra $\mathscr{E}(X)$ coincides also with the σ–algebra of cylinders based on elements ox X^*_γ; indeed, for any $F = \{f_1, \dots, f_d\} \subset X^*_\gamma$ the sets of the type

$$E = \{x \in X : (f_1(x), \dots, f_d(x)) \in B\},$$

with $B \in \mathscr{B}(\mathbb{R}^d)$, belong to $\mathscr{B}(X)$ for any choice of Borel representatives $f_1, \dots, f_d$. Hence $\mathscr{E}(X, X^*_\gamma) \subset \mathscr{B}(X)$. The reverse inclusion is trivial, whence the fact that $\mathscr{E}(X, X^*_\gamma) = \mathscr{B}(X)$ by the separability of X, see Theorem 2.1.1.

Lemma 7.4.2 easily implies further useful approximation results.

Lemma 7.4.4. *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* (7.4.3) *verifies $E \subset A$ and $\gamma(A \setminus E) < \varepsilon$.*

Proof. Let A be as in (7.4.3). 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 1.1.5 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 7.4.5. *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. It is sufficient to check L^p-convergence of indicator functions, as the general case follows by the density of simple functions. Taking into account that $\mathscr{E}(X) = \mathscr{E}(X, (\hat{h}_j)_{j\in\mathbb{N}})$, for every measurable set $B \subset X$ and $\varepsilon > 0$ there is $n \in \mathbb{N}$ and $B_\varepsilon \in \Sigma_n$ such that $\gamma(B \triangle B_\varepsilon) < (\varepsilon/2)^p$. Equality $\mathbb{E}_n 1\!\!1_{B_\varepsilon} = 1\!\!1_{B_\varepsilon}$ and estimate (7.4.2) yields

$$\begin{aligned}\|\mathbb{E}_n 1\!\!1_B - 1\!\!1_B\|_{L^p(X,\gamma)} &\le \|\mathbb{E}_n 1\!\!1_B - 1\!\!1_{B_\varepsilon}\|_{L^p(X,\gamma)} + \|1\!\!1_B - 1\!\!1_{B_\varepsilon}\|_{L^p(X,\gamma)} \\ &= \|\mathbb{E}_n (1\!\!1_B - 1\!\!1_{B_\varepsilon})\|_{L^p(X,\gamma)} + \|1\!\!1_B - 1\!\!1_{B_\varepsilon}\|_{L^p(X,\gamma)} \le \varepsilon.\end{aligned}$$

The convergence γ-a.e. of a subsequence follows from the L^p convergence, whereas the convergence of the whole sequence easily follows from Theorem 7.1.2 (which we have stated without proof) via dominated convergence. □

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 kerP_n, hence it can be identified with a function defined on $F = P_n X$ setting $f_n(y) = \mathbb{E}_n f(x)$, $y \in F$, $y = P_n x$.

Theorem 7.4.6. *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 $f_n \in L^p(\mathbb{R}^n, \gamma \circ P_n^{-1})$. Each f_n can be approached 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 $g_{n.j}(x) = \psi_{n,j}(P_n x)$, 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 $h_1, \ldots, h_d \in H$, the map $P : X \to \mathbb{R}^d$, $P(x) = (\hat{h}_1(x), \ldots, \hat{h}_d(x))$ 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 (7.4.3) 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 $\varphi \equiv 0$.

Bibliography

[B] V. I. Bogachev: *Gaussian Measures.* American Mathematical Society, 1998.

[D] R. M. Dudley: *Real Analysis and Probabillity,* Cambridge University Press, 2004.