



Universiteit  
Leiden  
The Netherlands

## Bayesian inference for Gaussian models: Inverse problems and evolution equations

Yan, D.

### Citation

Yan, D. (2020, March 3). *Bayesian inference for Gaussian models: Inverse problems and evolution equations*. Retrieved from <https://hdl.handle.net/1887/86070>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/86070>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/86070> holds various files of this Leiden University dissertation.

**Author:** Yan, D.

**Title:** Bayesian inference for Gaussian models: Inverse problems and evolution equations

**Issue Date:** 2020-03-03

## Chapter 3

# Gaussian Analysis

The study of Gaussian distributions has been an active field since the first invention of the normal distribution by Gauss in the nineteenth century. By generalizing the elementary concepts of mean and covariance of a Gaussian distribution to vector spaces, it leads to the study of Gaussian measure on vector spaces. The importance of Gaussian measure cannot be overstated. One classical example is Brownian motion, investigated by Wiener in the first half of the twentieth century, which is the central achievement in the development of stochastic processes. In 1940s and 1950s, several attempts to develop stochastic integration theory with Brownian motion as the integrator resulted in the well known Ito and Stratonovich calculus, which were later generalized to more general integrators such as Lévy process. In seventies, Malliavin in his celebrated seminal work established the stochastic version of the calculus of variation, now known as Malliavin calculus.

Gaussian analysis is also of great importance in applications. In 1973, Black and Scholes published the famous Black-Scholes formula, which revived the mathematical study of financial market using Gaussian models, which can be traced back to Bachelier's work in 1900. Since then, the stochastic differential equation driven by Wiener process has been a fundamental tool to construct many market models, see e.g. [86].

First, in Section 3.1 we give an overview on measures on Banach spaces. Gaussian measures are introduced in Section 3.2. We also briefly discuss how to radonify cylindrical Gaussian measures in Section 3.3.

In this chapter, we always consider the following situation, unless explicitly stating otherwise. Let  $(E, \|\cdot\|)$  be a *separable* Banach space. Denote its (topological) dual space by  $E^*$  with the (canonical) duality pair by  $\langle \cdot, \cdot \rangle : E^* \times E \rightarrow \mathbb{R}$ , i.e.  $\langle x^*, x \rangle := x^*(x)$ , and its Borel  $\sigma$ -algebra by  $\mathcal{B}(E)$  (or simply  $\mathcal{B}$ ).

### 3.1 Probability Measures on Banach Spaces

A Gaussian measure can be defined at different levels of generality. While it is standard to consider locally convex topological space as the working space when studying Gaussian measures, it is sufficient for us to only consider Gaussian measure on separable Banach spaces. Actually, we even only use the results of Gaussian

measures on Hilbert spaces, but it does not require many extra efforts to state the theory in Banach spaces. Since completeness is standard, we explain in short the reason for imposing separability. First, since the parameter space  $\Theta$  is separable and the transform operator  $\mathcal{A}$  is bounded,  $\mathcal{A}(\Theta)$  is necessarily separable, because separability is topologically invariant. Therefore, it is reasonable to only consider the case that  $\mathcal{A}(\Theta)$  embeds in another separable space. Consequently, it is sufficient to define the Gaussian noise in a separable space. Second, when introducing Gaussian priors, the parameter space  $\Theta$  is also assumed to be separable. Hence, separable Banach spaces are sufficient to serve our purposes.

We first consider some properties of general probability measures on vector spaces. The following type of probability measures is of special interest to us, as many theorems below use it as a premise.

**Definition 3.1** (Radon measure). A finite Borel measure  $\mu$  on a Hausdorff topological space  $E$  is a *Radon measure* if

$$\mu(B) = \sup\{\mu(K) : K \subset B, K \text{ compact}\},$$

for each  $B \in \mathcal{B}(E)$ .

Due to the following result, we will always deal with Radon measures.

**Theorem 3.2** (Theorem 3.1, Chapter II, [96]). *Every Borel probability measure on a Polish space, i.e. a complete separable metric space, is a Radon measure.*

When considering a probability measure, the  $\sigma$ -algebra has great influence on the properties of the probability. The following remark explains why it is always sufficient to consider Borel  $\sigma$ -algebra.

**Remark 3.3** (Choice of  $\sigma$ -algebra). Strictly speaking, probability measures on a vector space  $E$  are defined with the *cylindrical*  $\sigma$ -algebra  $\mathcal{C}(E)$  generated by the cylindrical sets of  $E$  (see Section 3.3), the smallest  $\sigma$ -algebra such that all continuous functionals in  $E^*$  are measurable. However, for a F chet space  $F$ ,  $\mathcal{C}(F) = \mathcal{B}(F)$ . Since all separable Banach spaces are F chet, there is no need to distinguish  $\mathcal{C}$  and  $\mathcal{B}$  in our setting. See Appendix A.3 in [8].

Fourier transform is a useful tool for showing the uniqueness of measures.

**Lemma 3.4** (Uniqueness of measures by Fourier transform (Lemma 7.13.5, [9])). *For a measure  $\mu$  on  $(E, \mathcal{C})$ , its Fourier transform, also known as characteristic functional,  $\hat{\mu}$  is defined by*

$$\hat{\mu} : E^* \rightarrow \mathbb{C}, \quad \hat{\mu}(f) = \int_E e^{if(x)} \mu(dx).$$

*Two measures on  $\mathcal{C}$  are identical, if they have same Fourier transform. In particular, it is true for Radon measures.*

For a random element in an infinite-dimensional Banach space  $E$ , the most fundamental concepts, mean and covariance, are defined as linear actions on the dual space  $E^*$ . While, topological support is also important for measures on vector spaces, on the contrary to measures on Euclidean spaces.

**Definition 3.5.** Let  $\mu$  be a probability measure on  $(E, \mathcal{B})$  such that  $E^* \subset L^2(\mu)$ . The *mean* is an element  $a_\mu$  in the algebraic dual  $(E^*)'$  of  $E^*$  satisfying

$$a_\mu f = \int_E f(x) \mu(dx).$$

The *covariance* operator  $\mathcal{C}_\mu : E^* \rightarrow (E^*)'$  is given by

$$\mathcal{C}_\mu(f)(g) := \int_E [f(x) - a_\mu(f)][g(x) - a_\mu(g)] \mu(dx). \quad (3.1)$$

The *topological support* of  $\mu$  is the smallest closed set  $S_\mu \in E$  such that

$$\mu(E \setminus S_\mu) = 0.$$

**Remark 3.6.** The general definition of mean and covariance is given by replacing  $E$  by a locally convex topological space and  $\mathcal{B}$  by the  $\sigma$ -algebra generated by cylindrical sets, see Section II.3 and III.2 in [96].

Similar to the finite-dimensional case, the existence of mean and covariance relies on certain moment conditions. Let the law of  $X$  be  $\mu$  on  $E$ . The  *$p$ th-weak moment* of  $X$  is given by

$$m_p(X) := \sup_{f \in U(E^*)} \int_E |\langle f, x \rangle|^p \mu(dx), \quad (3.2)$$

where  $U(E^*)$  is the unit ball of  $E^*$ , and  $\mathbb{E}\|X\|^p$  is the  *$p$ th-strong moment*. The criteria for the existence of mean and covariance are given as follows.

**Lemma 3.7.** *Let  $X$  be a random element on  $E$  with distribution  $\mu$ .*

- (i) *If  $\mu$  is Borel measurable and  $\mathbb{E}\|X\| < \infty$ , then  $a_\mu$  is an element of  $X$ .*
- (ii) *If  $\mu$  is Radon and  $m_2(X) < \infty$ , then  $\mathcal{C}_\mu(E^*) \subset E$ .*
- (iii) *If  $\mu$  is Radon and  $\mathbb{E}\|X\|^2 < \infty$ , then in addition to (ii),  $\mathcal{C}_\mu : E^* \rightarrow E$  is nuclear, i.e. it admits the representation,*

$$\mathcal{C}_\mu f = \sum_{j \in \mathbb{N}} \kappa_j \langle f, a_j \rangle b_j,$$

where  $\{a_j\}, \{b_j\}$  are sequences in  $E$  such that  $\|a_j\| = \|b_j\| = 1$  for  $j \in \mathbb{N}$ , and  $\{\kappa_j\}$  is a real sequence such that  $\sum_{j \in \mathbb{N}} |\kappa_j| < \infty$ .

*Proof.* The statement (i) is Lemma 2.1 in [100], and the rest are Theorem 2.1 and 2.3 in chapter III, [96].  $\square$

Some properties of the covariance operator  $\mathcal{C}_\mu$  are summarised in the following lemma.

**Lemma 3.8** (Factorization of covariance operators.). *Let  $\mu$  be a Radon probability measure with finite second order strong moment. The covariance operator  $\mathcal{C}_\mu : E^* \rightarrow E$ , given by*

$$\text{Cov}(f(X), g(X)) = \langle f, \mathcal{C}_\mu g \rangle, \quad f, g \in E^*,$$

is symmetric, i.e.

$$\langle f, \mathcal{C}_\mu g \rangle = \langle g, \mathcal{C}_\mu f \rangle, \quad f, g \in E^*,$$

positive<sup>1</sup>, i.e.  $\text{Var } f = \langle f, \mathcal{C}_\mu f \rangle \geq 0$  for all  $f \in E^*$ .

In addition, there exists a Hilbert space  $H$  and a bounded linear operator  $\mathcal{S} : E^* \rightarrow H$  with its dual  $\mathcal{S}^* : H \rightarrow (E^*)^*$ , whose range is  $\mathcal{S}^*(H) \subset E$ , such that the following factorization holds,

$$\mathcal{C}_\mu = \mathcal{S}^* \mathcal{S}.$$

In particular,  $\mathcal{C}_\mu(E^*) = \mathcal{S}^*(H)$ .

*Proof.* The factorization is Lemma 1.1 and 1.2 in chapter III, [96]. The rest are immediate consequences from the definition.  $\square$

In Lemma 3.8, the factorization is not unique in the following sense. If there exists another Hilbert space  $\tilde{H}$  and an isometry  $\mathcal{U} : H \rightarrow \tilde{H}$ , then  $\mathcal{C}_\mu$  can as well be factorized as  $\mathcal{C}_\mu = \tilde{\mathcal{S}}^* \tilde{\mathcal{S}}$  with

$$\tilde{\mathcal{S}} = \mathcal{U} \mathcal{S} : E^* \rightarrow \tilde{H} \quad \text{and} \quad \tilde{\mathcal{S}}^* = \mathcal{S}^* \mathcal{U}^{-1} : \tilde{H} \rightarrow E^*.$$

We will see in the next Section 3.2, that for Gaussian measures there is a natural choice of  $H$ . However, the factorization offers additional flexibility in terms of calculation, see also Section 3.3.

## 3.2 Gaussian Measures

The Gaussianity of a random element on Banach space  $E$  is characterised by the continuous functionals in  $E^*$ . Due to Remark 3.6, it is sufficient to consider measures on  $\mathcal{B}$ .

**Definition 3.9** (Gaussian measure). A probability measure  $\gamma$  on  $(E, \mathcal{B})$  is called *Gaussian* if, for any  $f \in E^*$ , the induced measure  $\gamma \circ f^{-1}$  is a Gaussian distribution on  $\mathbb{R}$ . A random element  $X$  in  $E$  is *Gaussian* if its law is Gaussian.

Directly from the definition, the quantity  $\mathbb{E} \|X\|^p$  is finite for  $p = 1, 2$ . Consequently  $m_2(X)$  also exists. Because of Lemma 3.7, for  $X$ , its mean  $a_\gamma$  of  $X$  is an element of  $E$  and the mapping  $a_\gamma : E^* \rightarrow E$  is bounded. If  $a_\gamma = 0$ , we say it is a *centred* (i.e. *zero mean*) Gaussian measure. In addition, the covariance operator is given by a symmetric positive nuclear operator  $\mathcal{C}_\gamma : E^* \rightarrow E$ .

As an analogy to finite-dimensional Gaussian distributions, one may conjecture that the Gaussian measures on vector spaces are as well uniquely determined by its mean and covariance. It turns out to be true due to the lemma below, taken from Proposition 2.8, Chapter IV, [96].

---

<sup>1</sup>To be precise, it is positive semi-definite or nonnegative definite. However, we use the term ‘positive’ for brevity, and strictly positive if  $\langle f, \mathcal{C}_\mu f \rangle > 0$  for any  $f \neq 0$ .

**Lemma 3.10** (Fourier transform and uniqueness of Gaussian measures). *For a Gaussian measure  $\gamma$  on  $E$  with mean  $a_\gamma$  and covariance operator  $\mathcal{C}_\gamma$ , its Fourier transform is given by*

$$\widehat{\gamma}(f) = \exp(i f(a_\gamma) - \frac{1}{2} \langle f, \mathcal{C}_\gamma f \rangle).$$

As shown in Lemma 3.10, the Fourier transform of Gaussian measures only involves the pair  $(a_\gamma, \mathcal{C}_\gamma)$ . Then by Lemma 3.4, Gaussian measures are completely determined by its mean and covariance. Henceforth, we will use the notation  $\mathcal{N}_E(a_\gamma, \mathcal{C}_\gamma)$  to denote a Gaussian measure on  $E$  with mean  $a_\gamma$  and covariance operator  $\mathcal{C}_\gamma$ . In addition, the following corollary is an immediate consequence of Lemma 3.10.

Before starting investigating fine properties of the covariance structure, we collect two celebrated inequalities demonstrating the fundamental properties of Gaussian measures, i.e. isoperimetric inequality and exponential tail property. For the detailed discussion, we refer to Section 3.1 in [67].

**Lemma 3.11.** *Let the law of  $X$  be a Gaussian measure  $\gamma$  on a separable Banach space  $E$ .*

(i) *Borell's inequality. With  $\sigma = m_2(X)$  given in (3.2), we have*

$$\mathbb{P}\left(\left|\|X\| - \mathbb{E}\|X\|\right| > t\right) \leq 2 \exp\left\{-\frac{t^2}{2\sigma^2}\right\}.$$

(ii) *Exponential tail (Fernique). There exists a positive constant  $\alpha$  such that*

$$\int_E e^{\alpha\|x\|^2} \gamma(dx) < \infty.$$

### 3.2.1 Covariance Structure

We are going to look closely on the covariance structure of a Gaussian measure  $\mathcal{N}_E(a_\gamma, \mathcal{C}_\gamma)$ . The following two subspaces of  $E^*$  and  $E$  respectively play an important role in shaping the Gaussian covariance structure.

From the definition,  $\|f\|_{L^2(\gamma)}$  is finite for all  $f \in E^*$ . Hence we can introduce the following definitions.

**Definition 3.12** (Reproducing kernel Hilbert space & Cameron-Martin space). For a Gaussian measure  $\gamma$  on a separable Banach space  $E$ , define an embedding

$$\mathcal{J} : E^* \rightarrow L^2(\gamma), \quad f \mapsto f - a_\gamma f, \tag{3.3}$$

which is continuous because of the finite moments.

(i) The *reproducing kernel Hilbert space* (RKHS) of the measure  $\gamma$  is defined as the closure of  $\mathcal{J}(E^*)$  in  $L^2(\gamma)$ , denoted by  $E_\gamma^*$ , which self is a Hilbert space equipped with  $L^2(\gamma)$  inner product.

(ii) The *Cameron-Martin space* is a subspace in  $E$  defined as

$$\mathbb{H}_\gamma := \left\{ h \in E : |h|_{\mathbb{H}_\gamma} < \infty \right\},$$

where

$$|h|_{\mathbb{H}_\gamma} := \sup \left\{ f(h) : f \in E^*, \|\mathcal{J}(f)\|_{L^2(\gamma)} \leq 1 \right\}. \quad (3.4)$$

**Remark 3.13** (Gaussian covariance operator). Consider the following extension of the covariance operator  $\mathcal{C}_\gamma$  to  $E_\gamma^*$ ,

$$\begin{aligned} \mathcal{K}_\gamma : E_\gamma^* &\rightarrow E \subset (E^*)', \\ (\mathcal{K}_\gamma f)(g) &= \int_E f(x)[g(x) - a_\gamma g] \gamma(dx), \quad f \in E_\gamma^*, g \in E^*. \end{aligned} \quad (3.5)$$

Notice that, if  $\gamma$  is centred,  $\mathcal{K}_\gamma$  is just the continuous extension to  $E_\gamma^*$  by bounded linear transform theorem (Theorem I.7, [81]). In the case of general Gaussian measures, for any  $f \in E^*$ ,  $\mathcal{K}_\gamma f$  coincides with  $\mathcal{C}_\gamma(f - a_\gamma f)$  (the original one defined in (3.1)), where  $f - a_\gamma f \in E_\gamma^*$ , as elements in  $(E^*)'$ , although  $f$  is not necessarily in  $E_\gamma^*$  if  $a_\gamma f \neq 0$ . From now on, for a Gaussian measure,  $\gamma$  the operator  $\mathcal{K}_\gamma$  is understood as in the extension of  $\mathcal{C}_\gamma$  as in (3.5).

Notice that the elements in  $E_\gamma^*$  are centred by the mapping  $\mathcal{J}$ . The following lemma shows that in fact they are also Gaussian.

**Lemma 3.14.** *Let the law of  $X$  be a Gaussian measure  $\gamma$ . For arbitrary  $f \in E_\gamma^*$ ,  $f(X)$  is a centred Gaussian random variable with variance*

$$\mathbb{E}(f(X))^2 = \int_E f^2 d\gamma = \|f\|_{L^2(\gamma)}^2.$$

*Proof.* It follows from the fact that the limit of a Gaussian sequence converging in probability is Gaussian. For the detail, see Lemma 2.2.8, [8].  $\square$

The usage of the term RKHS is not consistent in literature: it may refer to both  $E_\gamma^*$  and  $\mathbb{H}_\gamma$ . But this does not really cause any troubles in practice, as in fact  $E_\gamma^*$  and  $\mathbb{H}_\gamma$  are almost mutually replaceable, as shown in the Lemma 3.15 below.

**Lemma 3.15.** *Let  $\gamma$  be a Gaussian measure on a separable Banach space  $E$ . An element  $h \in E$  belongs to  $\mathbb{H}_\gamma$  if and only if  $h = \mathcal{K}_\gamma \widehat{h}$  for some functional  $\widehat{h} \in E_\gamma^*$ . In addition,*

$$|h|_{\mathbb{H}_\gamma} = \|\widehat{h}\|_{L^2(\gamma)}. \quad (3.6)$$

Consequently,  $\mathcal{K}_\gamma : E_\gamma^* \rightarrow \mathbb{H}_\gamma$  is an isometric isomorphism and  $\mathbb{H}_\gamma = \mathcal{K}_\gamma(E_\gamma^*)$  is a Hilbert space equipped with the inner product

$$\langle h_1, h_2 \rangle_{\mathbb{H}_\gamma} := \langle \widehat{h}_1, \widehat{h}_2 \rangle_{L^2(\gamma)}, \quad (3.7)$$

where  $\mathcal{K}_\gamma \widehat{h}_i = h_i$ , for  $i = 1, 2$ .

*Proof.* suppose  $|h|_{\mathbb{H}_\gamma} < \infty$ . Define a linear mapping

$$\Phi_h : E^* \rightarrow \mathbb{R}, \quad f \mapsto [\mathcal{J}f](h).$$

It is well-defined, since  $|\Phi_h f| \leq \|\mathcal{J}f\|_{L^2(\gamma)} |h|_{\mathbb{H}_\gamma}$  from the definition (3.4), which also implies that  $\Phi_h$  is continuous with respect to  $L^2(\gamma)$  norm. Then by BLT theorem (Lemma A.4), there is a continuous extension of  $\Phi_h$  to  $E_\gamma^*$ , denoted with the same symbol. Consequently, by Riesz representation theorem, there exists an element  $\hat{h}$  in  $E_\gamma^*$  such that  $\Phi_h f = \langle f, \hat{h} \rangle_{L^2(\gamma)}$ , for all  $f \in E_\gamma^*$ . In particular, for  $f \in E^*$ ,

$$f(h) = \Phi_h f = \int_E (\mathcal{J}f) \hat{h} d\gamma = f(\mathcal{K}_\gamma \hat{h}).$$

Using the double dual norm (Lemma A.5), we conclude  $h = \mathcal{K}_\gamma \hat{h}$ , and

$$|h|_{\mathbb{H}_\gamma} = \sup\{f(h) : f \in E^*, \|\mathcal{J}f\|_{L^2(\gamma)} \leq 1\} = \|\hat{h}\|_{L^2(\gamma)}. \quad (3.8)$$

Conversely, assume  $h = \mathcal{K}_\gamma \hat{h}$  for some  $\hat{h} \in E_\gamma^*$ . Then, for all  $f \in E^*$ ,

$$|f(h)| = |\langle \mathcal{J}f, \hat{h} \rangle_{L^2(\gamma)}| \leq \|\mathcal{J}(f)\|_{L^2(\gamma)} \|\hat{h}\|_{L^2(\gamma)},$$

which implies  $|h|_{\mathbb{H}_\gamma} < \infty$ .

The isometry (3.6) follows from (3.8). Since isometry implies injectivity, the claim that  $\mathbb{H}_\gamma$  equipped with (3.7) is a Hilbert space follows from Lemma A.11.  $\square$

As its name implies, the Cameron-Martin space is indeed a Hilbert space. Due to Lemma 3.15, we use the term RKHS for both  $\mathbb{H}_\gamma$  and  $E_\gamma^*$ . In most situations, it does not lead to confusions. Otherwise, we will explicitly specify whether RKHS is a subspace of  $E$  or  $E_\gamma^*$ .

Known from Lemma 3.8, the covariance operator can be factorized with an operator mapping  $E^*$  to a Hilbert space. Again due to Lemma 3.15, a natural choice  $E_\gamma^*$  surfaces for the Gaussian measure  $\gamma$ .

**Corollary 3.16** (Canonical factorization of Gaussian covariance operators). *For a Gaussian measure  $\gamma$  on  $E$ , its covariance operator  $\mathcal{C}_\gamma$  admits the following factorization  $\mathcal{C}_\gamma = \mathcal{J}^* \mathcal{J}$ , where  $\mathcal{J}$  is the embedding of  $E^*$  to  $E_\gamma^*$  defined in (3.3), and  $\mathcal{J}^* = \mathcal{K}_\gamma$ , i.e. the extension of  $\mathcal{C}_\gamma$  to  $E_\gamma^*$  given in (3.5).*

The RKHS contains large information on the Gaussian measure. We present several results with different flavours below. The proofs can be found in Section 2.4 and 3.2, [8].

Seemingly contradictory, the size of RKHS  $\mathbb{H}_\gamma$  can be both small and big at the same time.

**Proposition 3.17.** *Let  $\gamma$  be a Gaussian measure on  $E$ . If  $\dim E = \infty$ , then  $\mathbb{H}_\gamma$  only has zero measure, i.e.  $\gamma(\mathbb{H}_\gamma) = 0$ . On the other hand, the topological support of  $\gamma$  is the closure  $\overline{\mathbb{H}_\gamma}$  of  $\mathbb{H}_\gamma$  in  $E$ . In particular, if  $\gamma$  is a nondegenerate Gaussian measure, i.e. the topological support of  $\gamma$  is the whole space  $E$ ,  $\mathbb{H}_\gamma$  is dense in  $E$ . Furthermore, the unit ball  $U(\mathbb{H}_\gamma) = \{h \in \mathbb{H}_\gamma : |h|_{\mathbb{H}_\gamma} \leq 1\}$  is compact in  $E$ , which implies the embedding  $\mathbb{H}_\gamma$*

Gaussian measures are not so relevant to the underlying space  $E$ , but rather determined by its RKHS, shown as follows.

**Proposition 3.18.** *Let  $\gamma$  be a Gaussian measure on  $E$ . If  $E$  is continuously embedded another Banach space  $\tilde{E}$ , If  $\gamma$  is considered to be a Gaussian measure on  $\tilde{E}$ , the Cameron-Martin space is still  $\mathbb{H}_\gamma$ . Let  $\mu$  be another Gaussian measure on  $E$ . If  $\mathbb{H}_\gamma = \mathbb{H}_\mu$  and  $|h|_{\mathbb{H}_\gamma} = |h|_{\mathbb{H}_\mu}$  for all  $h \in \mathbb{H}_\gamma$ , then  $\gamma = \mu$ .*

The absolute continuity of measures is especially important for statistical study, since it implies a density, which is often desirable for the construction of statistics such as likelihood ratios, etc. For Gaussian measures, this is closely related to RKHS. Consider a Gaussian random element  $X$  with distribution  $\gamma$ . The distribution of the vector  $X + h$  with  $h \in E$  is defined by

$$\gamma_h(B) = \gamma(B - h), \quad B \in \mathcal{B}(E).$$

We call  $\gamma_h$  the *shift* of  $\gamma$  in the direction  $h$ . A shift is admissible if  $\gamma_h \ll \gamma$ . The proposition below shows that  $\mathbb{H}_\gamma$  is exactly the set of admissible shifts.

**Proposition 3.19** (Cameron-Martin). *Let  $\gamma$  be a Gaussian measure on  $E$  and  $\mathbb{H}_\gamma$  be its RKHS. Consider the shifted measure  $\gamma_h = \gamma(\cdot - h)$ .*

*If  $h \in \mathbb{H}_\gamma$ , then  $\gamma \sim \gamma_h$ , and the Radon-Nikodym density is given by*

$$\frac{d\gamma_h}{d\gamma}(x) = \exp\left(\widehat{h}(x) - \frac{1}{2}|h|_{\mathbb{H}_\gamma}^2\right), \quad (3.9)$$

where  $\mathcal{K}_\gamma \widehat{h} = h$ ; otherwise  $\gamma \perp \gamma_h$ .

The density (3.9) is also known as *Cameron-Martin formula*. Combining the above Proposition 3.19 with Lemma 3.10, we conclude that if  $h \in \mathbb{H}_\gamma$ , then

$$a_{\gamma_h} = a_\gamma + h.$$

*Kullback-Leibler(KL) divergence* is an important quantity used to quantify the discrepancy between probability measures, which is defined as

$$D_{KL}(\mathbb{P}; \mathbb{Q}) := \int_E \log \frac{d\mathbb{P}}{d\mathbb{Q}} d\mathbb{P}, \quad (3.10)$$

where  $\mathbb{P}$  and  $\mathbb{Q}$  are two probability measures on  $E$ . KL divergence is of great importance to Bayesian nonparametric inference, as it is used to characterise a set around the truth, whose probability mass has a direct implication on the contraction rate of the posterior distribution (see Section 4.3). For Gaussian measures, it can be calculated using RKHS norm, as shown in the lemma below.

**Lemma 3.20** (Kullback-Leibler divergence of Gaussian measures). *Let  $\gamma$  be a centred Gaussian measure on  $E$ , and  $h, g$  be two elements from  $\mathbb{H}_\gamma$ . Consider the shifted measures  $\gamma_h = \gamma(\cdot - h)$  and  $\gamma_g = \gamma(\cdot - g)$ . The KL divergence between  $\gamma_h$  and  $\gamma_g$  is given by*

$$D_{KL}(\gamma_h; \gamma_g) = \frac{1}{2}|h - g|_{\mathbb{H}_\gamma}^2.$$

*Proof.* From Lemma 3.15,  $\mathcal{K}_\gamma : E_\gamma^* \rightarrow \mathbb{H}_\gamma$  is an isometric isomorphism. Because  $\gamma$  is centred, i.e.  $a_\gamma = 0$ , in particular we have

$$|f|_{\mathbb{H}_\gamma}^2 = \|\widehat{f}\|_{L^2(\gamma)}^2 = \langle \widehat{f}, \mathcal{K}_\gamma \widehat{f} \rangle_{E_\gamma^* \times \mathbb{H}_\gamma}.$$

Hence we have  $h = \mathcal{K}_\gamma \widehat{h}$  and  $g = \mathcal{K}_\gamma \widehat{g}$ . Then it follows from Proposition 3.19,

$$\frac{d\gamma_h}{d\gamma_g}(x) = \frac{d\gamma_h/d\gamma}{d\gamma_g/d\gamma}(x) = \exp\left(\widehat{h}(x) - \frac{1}{2}|h|_{\mathbb{H}_\gamma}^2 - \widehat{g}(x) + \frac{1}{2}|g|_{\mathbb{H}_\gamma}^2\right).$$

Consequently,

$$\begin{aligned} D_{KL}(\gamma_h; \gamma_g) &= \int_E \log \frac{d\gamma_h}{d\gamma_g} d\gamma_h = \int_E \widehat{h}(x) - \widehat{g}(x) d\gamma_h - \frac{1}{2}|h|_{\mathbb{H}_\gamma}^2 + \frac{1}{2}|g|_{\mathbb{H}_\gamma}^2 \\ &= \mathbb{E} \widehat{h} - \mathbb{E} \widehat{g} - \frac{1}{2} \langle \widehat{h}, \widehat{h} \rangle_{L^2(\gamma)} + \frac{1}{2} \langle \widehat{g}, \widehat{g} \rangle_{L^2(\gamma)} \\ &= \frac{1}{2} \left[ \langle \widehat{h}, \mathcal{K}_\gamma \widehat{h} \rangle_{E_\gamma^* \times \mathbb{H}_\gamma} - \langle \widehat{g}, \mathcal{K}_\gamma \widehat{g} \rangle_{E_\gamma^* \times \mathbb{H}_\gamma} + \langle \widehat{g}, \mathcal{K}_\gamma \widehat{g} \rangle_{L^2(\gamma)} \right] \\ &= \frac{1}{2} \langle \widehat{h} - \widehat{g}, \mathcal{K}_\gamma (\widehat{h} - \widehat{g}) \rangle_{E_\gamma^* \times \mathbb{H}_\gamma} = \frac{1}{2} |h - g|_{\mathbb{H}_\gamma}^2, \end{aligned}$$

where we use  $\int_E \widehat{h}(x) d\gamma_h = \mathbb{E} \widehat{h} = \widehat{h}(h) = \langle \widehat{h}, \mathcal{K}_\gamma \widehat{h} \rangle_{E_\gamma^* \times \mathbb{H}_\gamma}$  directly from the definition.  $\square$

**Remark 3.21.** All the results stated in Section 3.2 can be generalised to the *Radon* Gaussian measures on a locally convex topological space  $E$ . See Chapter 3 in [8] for the details. In particular, we mention that Proposition 3.19 and Lemma 3.20 remain valid with  $E$  being a locally convex topological space, e.g  $\mathbb{R}^{\mathbb{N}}$ .

### 3.2.2 Examples

We now present several examples of Gaussian elements.

**Example 3.22** (Standard Gaussian measure on  $\mathbb{R}^{\mathbb{N}}$ ). Let  $\gamma_n$  be the standard Gaussian measures on  $\mathbb{R}$ . Then the measure

$$\gamma = \bigotimes_{n \in \mathbb{N}} \gamma_n$$

is centred Gaussian on  $E = \mathbb{R}^{\mathbb{N}}$ . Furthermore,  $E_\gamma^* \simeq \ell^2$  and  $\mathbb{H}_\gamma = \ell^2$ . For the proof, see Example 2.3.5 in [8].

Although at the first glance, Example 3.22 seems to be quite specific, it is the ‘only’ centred Radon Gaussian measure on locally convex spaces (in particular Banach spaces) in the sense that there always exists an isomorphism between it and another centred Radon Gaussian measure on a Banach space (see Theorem 3.4.4, [8]).

**Example 3.23** (Gaussian measures on Hilbert spaces). Let  $(H, \langle \cdot, \cdot \rangle)$  be a separable Hilbert space and we identify  $H^*$  with  $H$ . For simplicity, we consider the zero mean case. Let  $\mathcal{C}$  be a covariance operator. Then by spectral theorem of self-adjoint compact operators, there exists an orthonormal basis  $\varphi_k$  diagonalising  $\mathcal{C}$ , i.e.  $\mathcal{C}\varphi_k = \kappa_k\varphi_k$ . In particular,  $\sum_k \kappa_k < \infty$ , since  $\mathcal{C}$  is nuclear. Due to the diagonalisation, a Gaussian element  $X$  in  $H$  with distribution  $\mathcal{N}_H(0, \mathcal{C})$ , admits the representation

$$X = \sum_k \sqrt{\kappa_k} \varphi_k \xi_k,$$

where  $\{\xi_k\}$  is a sequence of i.i.d.  $\mathcal{N}_{\mathbb{R}}(0, 1)$  random variables. This representation is known as *Karhunen-Loève expansion*. Conversely, the representation also defines a Gaussian measure on  $H$ .

Now we identify the RKHS  $\mathbb{H}$  of  $\gamma$  using the canonical factorisation of  $\mathcal{C}$  (see Corollary 3.16). The closure of  $\mathcal{J}(H^*)$  follows from explicit calculations,

$$\mathcal{J}(H^*) = H_\gamma^* = \left\{ f : f(h) = \langle f, h \rangle_H, \|\mathcal{J}(f)\|_{L^2(\gamma)}^2 := \sum_k (\sqrt{\kappa_k} f_k)^2 < \infty \right\}.$$

Then, using Lemma 3.15, we obtain

$$\mathbb{H}_\gamma = \left\{ h \in H : |h|_{\mathbb{H}}^2 = \sum_k \left( \frac{h_k}{\sqrt{\kappa_k}} \right)^2 < \infty \right\},$$

by applying change of variables.

**Remark 3.24.** A Gaussian measure on Hilbert space can be connected to a Hilbert scale (see Section 2.2). Using the notations from the previous example, define  $\Lambda = \mathcal{C}^{-1/2}$  using spectral theorem on the domain  $\text{Dom } \Lambda = \mathbb{H}_\gamma$ , which is dense in  $X$ . Let  $\{H_s\}_{s \in \mathbb{R}}$  be the Hilbert scale generated by  $\Lambda$  with  $H_0 = H$ . Then, we have  $H_1 = \mathbb{H}_\gamma$  and  $H_{-1} = H_\gamma^*$ . Furthermore,  $\Lambda^{-2} : H_{-1} \rightarrow H_1$  is an isometric isomorphism. In fact,  $\Lambda^{-2}$  is the extension  $\mathcal{K}_\gamma$  of covariance operator  $\mathcal{C}$ , c.f. Lemma 3.15.

Next we consider the Gaussian measure induced by Gaussian processes. Recall that a random process is a family  $\{X_t\}_{t \in \mathfrak{T}}$  of random elements indexed a parameter set  $\mathfrak{T}$ , defined on a common probability space  $(\Omega, \mathbb{P})$ . A process is *Gaussian* if  $(X_{t_1}, \dots, X_{t_n})$  is a  $\mathbb{R}^n$ -valued Gaussian random element, for any  $\{t_1, \dots, t_n\} \subset \mathfrak{T}$ . Similar to Gaussian measures, the property of Gaussian processes are completely determined by the expectation  $\mathbb{E} X_t, t \in \mathfrak{T}$  and covariance  $\text{Cov}(X_s, X_t), s, t \in \mathfrak{T}$ .

**Example 3.25** (Classical Wiener measure). Wiener process  $W_t$  on  $[0, 1]$  is characterised by

$$\mathbb{E} W_t = 0, \quad \text{Cov}(W_s, W_t) = s \wedge t,$$

for  $s, t \in [0, 1]$ . It is well know that  $W_t$  is almost surely in  $E = C[0, 1]$ , the Banach space of continuous functions on  $[0, 1]$ , equipped with supremum norm

$\|f\|_E = \sup_{0 \leq x \leq 1} |f(x)|$ . The dual space  $E^*$  is the space of sign measures of finite variation on  $[0, 1]$ , with duality given by

$$\langle \mu, f \rangle_{E^* \times E} := \int_{[0,1]} f d\mu.$$

Consider  $W = W_{(\cdot)}$  as a random element in  $E$ . It is obvious  $\mathbb{E}W = 0$ . In addition,

$$\begin{aligned} \mathbb{Cov}(\langle \mu, W \rangle, \langle \nu, W \rangle) &= \mathbb{E} \int_{[0,1]^2} W(s)W(t) \mu(ds)\nu(dt) \\ &= \int_{[0,1]^2} \mathbb{E} W(s)W(t) \mu(ds)\nu(dt) = \int_{[0,1]^2} s \wedge t \mu(ds)\nu(dt) = \langle \mu, \mathcal{K}\nu \rangle, \end{aligned}$$

where

$$(\mathcal{K}\nu)(s) = \int_{[0,1]} s \wedge t \nu(dt).$$

It can be shown that the operator  $\mathcal{K}$  is indeed a covariance operator for the centred Gaussian measure on  $E = C[0, 1]$  induced by Wiener process on  $[0, 1]$ .

The result in Example 3.25 can be extended to general Gaussian processes with continuous paths as follows. Let  $E = C(\mathfrak{X}; \mathbb{R})$ , where  $\mathfrak{X}$  is a compact metric space, equipped with the supremum norm. Then the dual space  $E^*$  is again given by

$$\langle \varphi_\mu, f \rangle_{E^* \times E} := \int_{\mathfrak{X}} f d\mu,$$

where  $\mu$  is a signed measure with finite variation  $|\mu|(E) = \|\varphi_\mu\|$  (see Chapter 6, [83]). For a Gaussian process  $X_t$  with  $\mathbb{E}X_t = a(t)$  and  $\mathbb{Cov}(X_s, X_t) = k(s, t)$ , the distribution of  $X = X_{(\cdot)}$  is a Gaussian measure on  $E$  with expectation  $a$  and covariance operator

$$(\mathcal{K}\nu)(s) = \int_{\mathfrak{X}} k(s, t) \nu(dt).$$

Conversely, every Gaussian measure on  $E$  induces a continuous Gaussian process on its dual space  $E^*$ .

**Example 3.26** (Gaussian process induced by Gaussian measures). A Gaussian measure  $\gamma$  on a separable Banach space  $E$  always induces a continuous Gaussian process on the dual space  $E^*$ . Since  $E$  is separable, the unit ball  $U(E^*)$  of its dual is metrizable for the weak\* topology on  $\sigma(E^*, E)$  (where  $E$  is considered as a subset of the double dual  $E^{**}$  of  $E$ ). Denote by  $(z_k)$  a weak dense sequence in  $U(E^*)$ . For  $x$  in  $E$ , we have  $\|x\| = \sup_k |z_k(x)|$ . Consider each  $f$  in  $U(E^*)$  as a Gaussian random variable on  $(E, \gamma)$ . This defines a weak\* continuous Gaussian process, since for each  $x$  in  $E$ , the map  $f \mapsto f(x)$  is weak\* continuous.

As usual, the situation becomes more transparent, when the underlying space is a Hilbert space.

**Example 3.27** (Gaussian process indexed by Hilbert space). Let  $X$  be a centred Gaussian random element on  $H$  with covariance operator  $\mathcal{C}$ . Denote the functional  $\langle h, \cdot \rangle_H$  by  $\ell_h$ . The covariance

$$\text{Cov } \ell_h(X) \ell_g(X) = \langle h, \mathcal{C}g \rangle_H = K(h, g).$$

Hence,  $\{W(h) := \langle h, X \rangle_H, h \in H\}$  is a Gaussian process indexed by  $H$  with covariance function  $K(h, g) = \langle h, \mathcal{C}g \rangle_H$ . From the previous example, the process  $\{W(h)\}_{h \in H}$  is a (weak-)continuous Gaussian process on  $H$ .

### 3.3 Radonification of Cylindrical Measures

In this section we continue the study of the relations between Gaussian measures and processes. We are interested in the following Gaussian process, which is of great importance in statistics.

**Definition 3.28** (Isonormal process). An *isonormal* process (also known *white noise* process)  $\{W(h) : h \in H\}$  indexed by a Hilbert space  $H$  is a stochastic process such that:

- (i)  $W(h_1), \dots, W(h_n)$  are jointly centred Gaussian for all  $h_1, \dots, h_n \in H$ ;
- (ii)  $\text{Cov}(W(h), W(g)) = \langle h, g \rangle_H$ .

Analogous to Example 3.23, the isonormal process on  $H$  can be constructed using series representation. Let  $\{\varphi_k\}$  be an orthonormal basis on  $H$  and  $\{\xi_k\}$  be a sequence of i.i.d.  $\mathcal{N}_{\mathbb{R}}(0, 1)$  random variables. Then the process

$$W(h) := \sum_k \xi_k \langle \varphi_k, h \rangle \tag{3.11}$$

is isonormal. However, it does not induce a genuine Gaussian measure as the covariance operator  $\mathcal{C} = \text{id}$  is not nuclear if  $\dim H = \infty$ .

We are going to demonstrate that the ‘measure’ induced by an isonormal process can be extended to a larger space such that it becomes an authentic measure, which is known as *Radonification*. To show this, we first introduce the following concepts.

The *cylindrical* sets in a Banach space  $E$  with its dual  $E^*$  has the following form

$$C = \{x \in E : (f_1(x), \dots, f_n(x)) \in C_0\}, \quad f_i \in E^*,$$

where  $C_0 \in \mathcal{B}(\mathbb{R}^n)$  is a base of  $C$ . A nonnegative additive function  $\nu$  is a *cylindrical measure* on  $E$  if for every continuous linear operator  $\mathcal{P} : E \rightarrow \mathbb{R}^n$ , the set function

$$\nu \circ \mathcal{P}^{-1} : B \mapsto \nu(\mathcal{P}^{-1}(B))$$

on the algebra  $\mathcal{A}(E)$  of cylindrical sets is countably additive. Fourier transform of a cylindrical measure  $\nu$  is defined similarly to the one of measures,

$$\widehat{\nu}(f) = \int_{\mathbb{R}} e^{it} \nu \circ f^{-1}(dt), \quad f \in E^*.$$

Now we see that an isonormal process from Definition 3.28 induces a cylindrical Gaussian measure.

**Example 3.29** (Cylindrical Gaussian measure on Hilbert space). Let  $H$  be an infinite dimensional separable Hilbert space. For all cylindrical sets  $C$  in the form  $C = P^{-1}(C_0)$ , where  $P$  is an orthogonal projection onto a  $n$ -dimensional subspace  $E_n$  of  $E$ , define

$$\nu(C) = \gamma_n(C_0),$$

where  $\gamma_n$  is the standard Gaussian measure on  $E_n$  (identified with  $\mathbb{R}^n$ ). The set function  $\nu$  is called the *canonical cylindrical Gaussian measure* on  $E$ , whose Fourier transform is

$$\widehat{\nu}(f) = e^{-\frac{1}{2}\|f\|_H^2}.$$

It is straightforward from the definition that an isonormal process induces  $\nu$  on  $E$ .

We recall a general result on radonification of Gaussian measures, known as Sazonov's theorem (see Theorem 3.4, [13]).

**Theorem 3.30** (Sazonov). *Let  $H, G$  be two Hilbert space and  $\nu$  be the canonical cylindrical Gaussian measure on  $H$ . A bounded operator  $\mathcal{T} : H \rightarrow G$  is  $\gamma$ -radonifying if the push forward  $\nu \circ \mathcal{T}^{-1}$  on  $G$  is a genuine Gaussian measure. We have,  $\mathcal{T} : H \rightarrow G$  is  $\gamma$ -radonifying if and only if  $\mathcal{T}$  is a Hilbert-Schmidt operator.*

We outline one application of Theorem 3.30 to diagonalisable operators. The main motivation for this is to cover the white noise. A bounded operator  $\mathcal{Q} : H \rightarrow H$  is called *diagonalisable*, if with an orthonormal basis  $\{\varphi_k\}_{k \in \mathbb{N}}$  for  $H$ , for all  $f \in H$ ,

$$\mathcal{Q}f = \sum_{k \in \mathbb{N}} q_k f_k \varphi_k \quad \text{with } f_k = \langle f, \varphi_k \rangle_H.$$

Let  $\mathcal{Q} : H \rightarrow H$  be a self-adjoint, positive definite, and diagonalisable operator on  $H$  with a basis  $\{\varphi_k\}_{k \in \mathbb{N}}$ . If  $\mathcal{Q}$  is of trace class, i.e.

$$\sum_{k \in \mathbb{N}} \langle \mathcal{Q}\varphi_k, \varphi_k \rangle_H = \sum_{k \in \mathbb{N}} \|\mathcal{Q}^{1/2}\varphi_k\|_H^2 < \infty,$$

it is the covariance of a centred Gaussian measure. On the other hand, if  $\mathcal{Q}$  is not of trace class, i.e.  $\sum_{k \in \mathbb{N}} q_k = \infty$ , it cannot be the covariance of a Gaussian distribution on  $H$ . However, as in Example 3.27, it defines a centred Gaussian process  $\{W(h) : h \in H\}$  with covariance function

$$\text{Cov}(W(h), W(g)) = \langle h, \mathcal{Q}g \rangle_H.$$

Denote

$$\mathbb{H} = \mathcal{Q}^{1/2}(H) = \left\{ f = \sum_k f_k \varphi_k : |f|_{\mathbb{H}} := \|\mathcal{Q}^{-1/2}f\|_H < \infty \right\},$$

which is a Hilbert space with the inner product  $\langle \cdot, \cdot \rangle_{\mathbb{H}} = \langle \mathcal{Q}^{-1/2} \cdot, \mathcal{Q}^{-1/2} \cdot \rangle_H$  (by Lemma A.11). By restricting the previous process to  $\mathbb{H}$ , the new process  $\{W(h) :$

$h \in \mathbb{H}$  is isonormal on  $\mathbb{H}$ , whose distribution is equivalent to the cylindrical measure on  $\mathbb{H}$ .

Notice that  $\{e_k = \mathcal{Q}^{1/2}\varphi_k : k \in \mathbb{N}\}$  forms a complete orthonormal basis of  $\mathbb{H}$ . Introduce an inner product  $\langle \cdot, \cdot \rangle_X$  on  $\mathbb{H}$  such that,

$$\langle e_j, e_k \rangle_X := w_j w_k \delta_{jk},$$

where the weights  $\{w_k\}_{k \in \mathbb{N}}$  satisfy  $\sum_{k \in \mathbb{N}} w_k^2 < \infty$  and  $\delta_{jk}$  is the Kronecker delta. For example, any sequence of weights that converges to zero faster than a square root rate, i.e.  $w_k \ll k^{-1/2}$  as  $k$  goes to infinity, satisfies the previous condition of summability. Let  $X$  be the closure of  $\mathbb{H}$  under the norm introduced by the inner product above. Then the embedding  $\iota : \mathbb{H} \hookrightarrow X$  is Hilbert-Schmidt. Furthermore, because  $\mathcal{Q}$  is diagonalised with  $\{\varphi_k\}$ , it is easy to examine that  $H \subset X$ . Consequently, by Theorem 3.30, there is a centred Gaussian measure with the covariance operator  $\iota \mathcal{Q} \iota^*$  on  $X$ , with the Cameron-Martin space  $\mathbb{H}$ . In particular, the technique discussed above is applicable to the white noise. Moreover, since the identity operator is diagonalisable with any orthonormal basis, we have the freedom to choose the most convenient one to work with.

**Example 3.31** (White noise process). Let  $H = L^2(\mathfrak{D})$  with a bounded domain  $\mathfrak{D} \subset \mathbb{R}^d$ . Let  $\xi : \mathfrak{D} \rightarrow \mathbb{R}$  be a Gaussian process such that  $\mathbb{E} \xi(s) \xi(t) = \delta(s - t)$ , where  $\delta$  is the Dirac delta function. Let  $\langle \xi, u \rangle := \int_{\mathfrak{D}} \xi(s) u(s) ds$ , where  $\xi(s) ds$  is understood in the distributional sense. We have

$$\mathbb{E} \langle \xi, u \rangle \langle \xi, v \rangle = \int_{\mathfrak{D}} \int_{\mathfrak{D}} u(s) v(t) \mathbb{E} [\xi(s) \xi(t)] ds dt = \int_{\mathfrak{D}} u(t) v(t) dt = \langle u, v \rangle_{L^2},$$

for  $u, v \in L^2(\mathfrak{D})$ . Therefore, the Gaussian process  $\xi$  is isonormal on  $L^2(\mathfrak{D})$ . By the results in this section, while  $\xi$  is not a  $L^2$ -valued Gaussian process, it can be realised in a space  $G$  such that the embedding of  $L^2(\mathfrak{D})$  to  $G$  is Hilbert-Schmidt, and the stochastic integral  $\int_{\mathfrak{D}} \xi(t) f(t) dt$  is well-defined. In particular,

$$\int_{\mathfrak{D}} \xi(s) \mathbb{1}_{\{s_i \leq t_i : 1 \leq i \leq d\}} ds_1 \cdots ds_d = \prod_{i \leq d} t_i$$

and

$$\mathbb{E} \langle \xi, \mathbb{1}_{\{r \leq s\}} \rangle \langle \xi, \mathbb{1}_{\{r \leq t\}} \rangle = \prod_{i \leq d} s_i \wedge t_i.$$

Consequently,  $\int_0^t \xi(s) ds$  is the Brownian sheet, which is often stated as white noise is the (distributional) derivative of Brownian motion. See also Section 4.1.5 in [23].

### 3.4 Notes

For a systematic (and more general) treatment on Gaussian measures, we refer to [8]. Radonification of a cylindrical Gaussian measure can also be stated in the framework of nuclear spaces, in which Minlos' theorem is used as the major tool, see [47]. The fields of white noise analysis, generalised stochastic processes were originated from this approach. A closely related object is the abstract Wiener space, see [74] for details.