PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 39, Fasc. 2 (2019), pp. 279–297
doi:10.19195/0208-4147.39.2.3

KARHUNEN–LOÈVE DECOMPOSITION OF GAUSSIAN MEASURES ON BANACH SPACES*

BY

XAVIER BAY (SAINT-ÉTIENNE) AND JEAN-CHARLES CROIX (SAINT-ÉTIENNE)

Abstract. The study of Gaussian measures on Banach spaces is of active interest both in pure and applied mathematics. In particular, the spectral theorem for self-adjoint compact operators on Hilbert spaces provides a canonical decomposition of Gaussian measures on Hilbert spaces, the so-called Karhunen–Loève expansion. In this paper, we extend this result to Gaussian measures on Banach spaces in a very similar and constructive manner. In some sense, this can also be seen as a generalization of the spectral theorem for covariance operators associated with Gaussian measures on Banach spaces. In the special case of the standard Wiener measure, this decomposition matches with Lévy–Ciesielski construction of Brownian motion.

2010 AMS Mathematics Subject Classification: Primary: 60B11, 60B12; Secondary: 28C20.

Key words and phrases: Gaussian measure, covariance operator, orthogonal decomposition.

1. PRELIMINARIES ON GAUSSIAN MEASURES

Let us first remind a few properties of Gaussian measures on Banach spaces. Our terminology and notation are essentially taken from [2] (alternative presentations can be found in [9], [18] or [5]). In this paper, we consider a separable Banach space X equipped with its Borel σ-algebra $\mathcal{B}(X)$. Note that every probability measure on $\big(X, \mathcal{B}(X)\big)$ is Radon and that Borel and cylindrical σ-algebras are equal in this setting.

A probability measure γ on $\big(X, \mathcal{B}(X)\big)$ is Gaussian if and only if for all $f \in X^*$ (the topological dual space of X), the pushforward measure $\gamma \circ f^{-1}$ (of γ

* Part of this research was conducted within the frame of the Chair in Applied Mathematics OQUAIDO, gathering partners in technological research (BRGM, CEA, IFPEN, IRSN, Safran, Storengy) and academia (CNRS, École Centrale de Lyon, Mines Saint-Étienne, University of Grenoble, University of Nice, University of Toulouse) around advanced methods for Computer Experiments.

through f) is a Gaussian measure on $\big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$. Here, we only consider the case γ centered for simplicity (the general case being obtained through a translation). An important tool in the study of a (Gaussian) measure is its characteristic functional $\hat{\gamma}$ (or Fourier transform)

$$\hat{\gamma} : f \in X^* \to \hat{\gamma}(f) = \int_X e^{i\langle x, f\rangle_{X,X^*}} \gamma(dx) \in \mathbb{C},$$

where $\langle \cdot, \cdot \rangle_{X,X^*}$ is the duality pairing. Since γ is a centered Gaussian measure, we have

$$\forall f \in X^*,\ \hat{\gamma}(f) = \exp\left(-\frac{C_\gamma(f,f)}{2} \right), \tag{1.1}$$

where C_γ is the covariance function

$$C_\gamma : (f,g) \in X^* \times X^* \to \int_X \langle x, f\rangle_{X,X^*} \langle x, g\rangle_{X,X^*} \gamma(dx) \in \mathbb{R}.$$

One of the most striking results concerns integrability. Indeed, using a rotation invariance principle, it has been shown that a Gaussian measure γ admits moments (in a Bochner sense) of all orders (as a simple corollary of Fernique's theorem, see [2]). Consequently, its covariance operator may be defined as

$$R_\gamma : f \in X^* \to \int_X \langle x, f\rangle_{X,X^*} x \gamma(dx) \in X,$$

using Bochner's integral, and is characterized by the following relation:

$$\forall (f,g) \in X^* \times X^*,\ \langle R_\gamma f, g\rangle_{X,X^*} = C_\gamma(f,g). \tag{1.2}$$

Most noticeably, R_γ is a symmetric positive kernel (Hilbertian or Schwartz kernel) in the following sense:

$$\begin{gathered} \forall (f,g) \in X^* \times X^*,\ \langle R_\gamma f, g\rangle_{X,X^*} = \langle R_\gamma g, f\rangle_{X,X^*}, \\ \forall f \in X^*,\ \langle R_\gamma f, f\rangle_{X,X^*} \geqslant 0. \end{gathered}$$

Furthermore, the Cameron–Martin space $H(\gamma)$ associated with γ is the Hilbertian subspace of X with Hilbertian kernel R_γ (see [15] and [1] for the usual case of reproducing kernel Hilbert spaces). In particular, we will extensively use the so-called *reproducing property*

$$\forall h \in H(\gamma),\ \forall f \in X^*,\ \langle h, f\rangle_{X,X^*} = \langle h, R_\gamma f\rangle_\gamma,$$

where $\langle \cdot, \cdot \rangle_\gamma$ denotes the inner product of $H(\gamma)$. Note that $H(\gamma)$ is continuously embedded in X and admits $R_\gamma(X^*)$ as a dense subset. Additionally, the covariance operator has been shown to be nuclear and, in particular, compact (see [18], Chapter 3, for a detailed presentation and proofs).

Our objective is to decompose any Gaussian measure γ on a (separable) Banach space X which, in fact, can be done by considering any Hilbert basis of the Cameron–Martin space $H(\gamma)$. Indeed, let $(h_n)_n$ be any arbitrary orthonormal basis of $H(\gamma)$, and $(\xi_n)_n$ a sequence of independent standard Gaussian random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the series

$$\sum_n \xi_n(\omega) h_n \tag{1.3}$$

converges almost surely in X and the distribution of its sum is the Gaussian measure γ (cf. Theorem 3.5.1, p. 112, in [2]). This idea of using Hilbert bases in the Cameron–Martin space goes back to the early seventieth (see [6] and [8]). In addition, the strongest result in this direction is the existence of such a basis $(h_n)_n$ which achieves

$$\sum_n \|h_n\|_X^2 < +\infty.$$

When X is a Hilbert space, a canonical Hilbert basis of the Cameron–Martin space $H(\gamma)$ is given by the spectral decomposition of the covariance operator R_γ as a self-adjoint compact operator on X (see Mercer's theorem in the special case $X = L^2[a, b]$ with $[a, b]$ any compact interval of $\mathbb{R}$). In this paper, we will show how to define and construct such a basis in the case X being Banach by a direct generalization of the Hilbert case. In particular, this "diagonalizing" basis will be of the form $h_n = R_\gamma h_n^*$, where h_n^* is in the dual space X^* for all n. As a special case of the representation (1.3), the corresponding decomposition in X (for the strong topology) will be

$$x = \sum_n \langle x, h_n^* \rangle_{X,X^*} h_n$$

γ almost everywhere (since $(h_n^*)_n$ is a sequence of independent standard normal random variables by the reproducing property). It means that γ can be seen as the countable product of the standard normal distribution $\mathcal{N}(0, 1)$ on the real line:

$$\gamma = \bigotimes_n \mathcal{N}(0, 1).$$

For a recent review of the interplay between covariance operators and Gaussian measures decomposition, consult [10]. To see how to construct such a basis, we start with the Hilbert case.

2. GAUSSIAN MEASURES ON HILBERT SPACES

Hilbert geometry has nice features that are well understood, including Gaussian measures structure (see [9] and [3] for a recent treatment). First of all, the Riesz representation theorem allows us to identify X^* with X. As a linear operator on a Hilbert space, the covariance operator R_γ of a Gaussian measure γ is self-adjoint and compact. Spectral theory exhibits a particular Hilbert basis of X

given by the set $(x_n)_n$ of eigenvectors of R_γ. Using this specific basis, we obtain the covariance operator in the form

$$R_\gamma : x \in X \to \sum_n \lambda_n \langle x, x_n \rangle_X x_n \in X,$$

where $\langle \cdot, \cdot \rangle_X$ is the inner product of X. A simple normalization, namely $h_n = \sqrt{\lambda_n} x_n$, provides a Hilbert basis of $H(\gamma)$. The nuclear property of R_γ simplifies to the form

$$\sum_n \|h_n\|_X^2 = \sum_n \lambda_n < +\infty.$$

Using the terminology of random elements, let Y be the infinite-dimensional vector defined almost surely by

$$Y(\omega) = \sum_n \xi_n(\omega) h_n = \sum_n \sqrt{\lambda_n} \xi_n(\omega) x_n,$$

where $(\xi_n)_n$ is a sequence of independent standard normal random variables. Then γ is the distribution of the Gaussian vector Y. In the context of stochastic processes, this representation is well known as the Karhunen–Loève expansion ([7], [12]) of the process $Y = (Y_t)_{t \in T}$ (assumed to be mean-square continuous over a compact interval $T = [a, b]$ of $\mathbb{R}$).

To extend this spectral decomposition to the Banach case, let us recall the following simple property (where B_X denotes the unit closed ball of X):

$$\lambda_0 = \sup_{x \in X \setminus \{0\}} \frac{\langle R_\gamma x, x \rangle_X}{\|x\|_X^2} = \max_{x \in B_X} \langle R_\gamma x, x \rangle_X \tag{2.1}$$

is the largest eigenvalue of the covariance operator R_γ and is equal to the Rayleigh quotient $\langle R_\gamma x_0, x_0 \rangle_X / \|x_0\|_X^2$, where x_0 is any corresponding eigenvector. A similar interpretation is valid for every $n \in \mathbb{N}$:

$$\lambda_n = \max_{x \in B_X \cap \operatorname{span}(x_0, \ldots, x_{n-1})^\perp} \langle R_\gamma x, x \rangle_X.$$

Keeping this interpretation in mind, we can now consider the Banach case.

3. GAUSSIAN MEASURES IN BANACH SPACES

In the context of Banach spaces, the previous spectral decomposition of the covariance operator does not make sense anymore. Nevertheless, we will show in Section 3.1 that the Rayleigh quotient is well defined in this context (Lemma 3.1). Combining this and a simple decomposition method (Lemma 3.2), we give in Section 3.2 an iterative decomposition scheme of a Gaussian measure. Main analysis and results are given in the last Section 3.3.

3.1. Rayleigh quotient and split decomposition. The first lemma in this section is an existence result of particular linear functionals based on a compactness property. The second one provides a method to separate a Banach space into two components with respect to a linear functional and a Gaussian measure. These results are given independently to emphasize that Lemma 3.2 could be combined with different linear functionals to define other iterative decomposition schemes (see Section 3.2).

LEMMA 3.1. *Let γ be a Gaussian measure on $\big(X, \mathcal{B}(X)\big)$ a separable Banach space and set $\lambda_0 = \sup_{f \in B_{X^*}} \langle R_\gamma f, f\rangle_{X,X^*} \in [0, +\infty]$. Then*

$$\exists f_0 \in B_{X^*},\ \lambda_0 = \langle R_\gamma f_0, f_0\rangle_{X,X^*}.$$

Moreover, we may assume $\|f_0\|_{X^*} = 1$.

P r o o f. Let $(f_n)_n \in B_{X^*}$ be a maximizing sequence:

$$\langle R_\gamma f_n, f_n\rangle_{X,X^*} \to \lambda_0 \in [0, +\infty].$$

From the weak-star compactness of B_{X^*} (see the Banach–Alaoglu theorem), we can suppose that $f_n \rightharpoonup f_\infty$ for the $\sigma(X^*, X)$-topology, where $f_\infty \in B_{X^*}$. This implies that

$$\hat\gamma(f_n) = \int_X e^{i\langle x, f_n\rangle_{X,X^*}} \gamma(dx) \to \int_X e^{i\langle x, f_\infty\rangle_{X,X^*}} \gamma(dx) = \hat\gamma(f_\infty),$$

by Lebesgue's convergence theorem. From equations (1.1) and (1.2) we conclude that $\langle R_\gamma f_n, f_n\rangle_{X,X^*} \to \langle R_\gamma f_\infty, f_\infty\rangle_{X,X^*}$. Hence $\lambda_0 = \langle R_\gamma f_\infty, f_\infty\rangle_{X,X^*} \in \mathbb{R}_+$. If $\lambda_0 > 0$, then $\|f_\infty\|_{X^*} = 1$ and we can take $f_0 = f_\infty$. In the degenerate case $\lambda_0 = 0$, we have $R_\gamma = 0$ and any f_0 of unit norm is appropriate. ∎

We will now show how to split both X and γ, given any $f \in X^*$ of non-trivial Rayleigh quotient (in the previous sense).

LEMMA 3.2. *Let $\gamma \neq \delta_0$ be a non-trivial Gaussian measure on a separable Banach space $\big(X, \mathcal{B}(X)\big)$. Pick $f_0 \in X^*$ such that $\|f_0\|_{X^*} = 1$ and $\lambda_0 = \langle R_\gamma f_0, f_0\rangle_{X,X^*} > 0$. Set $R_\gamma f_0 = \lambda_0 x_0$, $P_0 : x \in X \to \langle x, f_0\rangle_{X,X^*} x_0$ and $h_0 = \sqrt{\lambda_0} x_0$. Then we have the following properties*:

(1) $\langle x_0, f_0\rangle_{X,X^*} = 1$ *and* $\|h_0\|_\gamma = 1$.

(2) *P_0 is the projection on X with range $\mathbb{R}x_0$ and with the null space* $\ker(f_0) = \{x \in X,\ \langle x, f_0\rangle_{X,X^*} = 0\}$. *Furthermore, the restriction Q_0 of P_0 on $H(\gamma)$ is the orthogonal projection onto $\mathbb{R}h_0$*:

$$h \in H(\gamma),\ \langle h, f_0\rangle_{X,X^*} x_0 = \langle h, h_0\rangle_\gamma h_0.$$

(3) *According to the decomposition* $x = P_0x + (I - P_0)x$ *in* X, *the Gaussian measure* γ *can be decomposed as*

$$\gamma = \gamma_{\lambda_0} * \gamma_1,$$

where $\gamma_{\lambda_0} = \gamma \circ P_0^{-1}$ *and* $\gamma_1 = \gamma_0 \circ (I - P_0)^{-1}$ *are Gaussian measures with respective covariance operators*:

$$R_{\lambda_0} : f \in X^* \to \lambda_0 \langle x_0, f\rangle_{X,X^*} x_0,$$
$$R_{\gamma_1} : f \in X^* \to R_\gamma f - R_{\lambda_0} f.$$

In particular,

$$R_\gamma f = \lambda_0 \langle x_0, f\rangle_{X,X^*} x_0 + R_{\gamma_1} f.$$

(4) *The Cameron–Martin space* $H(\gamma)$ *is decomposed as*

$$H(\gamma) = \mathbb{R}h_0 \oplus H(\gamma_1),$$

where $H(\gamma_1) = (I - Q_0)\big(H(\gamma)\big) = (\mathbb{R}h_0)^\perp$ *equipped with the inner product of* $H(\gamma)$ *is the Cameron–Martin space of* γ_1.

(5) *For each* $t \in \mathbb{R}$, *denote by* $tx_0 + \gamma_1$ *the Gaussian measure on* X *centered at* tx_0 *with covariance operator* R_{γ_1}. *Then*, γ^t *is the conditional probability distribution of* $x \in X$ *given* $f_0(x) = t$:

$$\forall B \in \mathcal{B}(X),\ \gamma^t(B) = \gamma_1(B - tx_0) = \gamma(B|f_0 = t).$$

Moreover, f_0 *is* $\mathcal{N}(0, \lambda_0)$ *and the deconditioning formula is as follows*:

$$\gamma(B) = \int_{\mathbb{R}} \gamma^t(B) \frac{e^{-t^2/(2\lambda_0)}}{\sqrt{2\pi\lambda_0}} dt.$$

P r o o f. (1) Since $R_\gamma f_0 = \lambda_0 x_0$, we have

$$\lambda_0 \langle x_0, f_0\rangle_{X,X^*} = \langle R_\gamma f_0, f_0\rangle_{X,X^*} = \lambda_0$$

and $\lambda_0 > 0$ implies $\langle x_0, f_0\rangle_{X,X^*} = 1$. The second equality is obtained from the definition of h_0 and the reproducing property:

$$\|h_0\|_\gamma^2 = \langle h_0, h_0\rangle_\gamma = \langle x_0, \lambda_0 x_0\rangle_\gamma = \langle x_0, R_\gamma f_0\rangle_\gamma = \langle x_0, f_0\rangle_{X,X^*} = 1.$$

(2) Since $P_0x_0 = \langle x_0, f_0\rangle_{X,X^*} x_0 = x_0$, we have $P_0^2 = P_0$ and P_0 is clearly the projection onto $\mathbb{R}x_0$ along the null space of $f_0 \in X^*$. Now, if $h \in H(\gamma)$, we get by the reproducing property:

$$P_0 h = \langle h, R_\gamma f_0\rangle_\gamma x_0 = \langle h, \lambda_0 x_0\rangle_\gamma x_0 = \langle h, h_0\rangle_\gamma h_0 = Q_0 h.$$

(3) As bounded linear transformations of a (centered) Gaussian measure, both γ_{λ_0} and γ_1 are (centered) Gaussian measures. Consider the decomposition in X^*:

$$f = P_0^* f + (I - P_0^*)f.$$

Now, we see that $P_0^* f = \langle x_0, f\rangle_{X,X^*} f_0$ is Gaussian with variance $\lambda_0 \langle x_0, f\rangle_{X,X^*}^2$ and $(I - P_0^*)f = f - \langle x_0, f\rangle_{X,X^*} f_0$ is Gaussian with variance $\langle R_{\gamma_1} f, f\rangle_{X,X^*}$. To show that $P_0^* f$ and $(I - P_0^*)f$ are independent, we compute their covariance:

$$\begin{aligned}
&\int_X \langle x, P_0^* f\rangle_{X,X^*} \langle x, (I - P_0^*)f\rangle_{X,X^*} \gamma(dx) \\
&= \int_X \langle x_0, f\rangle_{X,X^*} \langle x, f_0\rangle_{X,X^*} \left(\langle x, f\rangle_{X,X^*} - \langle x_0, f\rangle_{X,X^*} \langle x, f\rangle_{X,X^*}\right) \gamma(dx) \\
&= \langle x_0, f\rangle_{X,X^*} \langle R_\gamma f_0, f\rangle_{X,X^*} - \lambda_0 \langle x_0, f\rangle_{X,X^*}^2 \\
&= 0.
\end{aligned}$$

Using the characteristic function of γ, by independence we get

$$\hat{\gamma}(f) = \int_X e^{i\langle x, (P_0^* f + (I - P_0^*)f)\rangle_{X,X^*}} \gamma(dx) = \widehat{\gamma_{\lambda_0}}(f) \hat{\gamma}_1(f).$$

This proves $\gamma = \gamma_{\lambda_0} * \gamma_1$ and also $R_\gamma = R_{\lambda_0} + R_{\gamma_1}$.

(4) Consider the orthogonal decomposition $H(\gamma) = \mathbb{R}h_0 \oplus H_1$, where $H_1 = (\mathbb{R}h_0)^\perp$. Since $R_{\lambda_0} f = \lambda_0 \langle x_0, f\rangle_{X,X^*} x_0 = \langle R_\gamma f, h_0\rangle_\gamma h_0$ is the orthogonal projection of $R_\gamma f$ onto $\mathbb{R}h_0$, we see that $R_\gamma f = R_{\lambda_0} f + R_{\gamma_1} f$ is the corresponding orthogonal decomposition of $R_\gamma f$. Therefore, by the Pythagorean theorem,

$$\|R_\gamma f\|_\gamma^2 = \|R_{\gamma_0} f\|_\gamma^2 + \|R_{\gamma_1} f\|_\gamma^2.$$

Now, using the relation

$$R_{\lambda_0} f = \lambda_0 \langle x_0, f\rangle_{X,X^*} x_0,$$

we get $\|R_{\lambda_0} f\|_\gamma^2 = \lambda_0 \langle x_0, f\rangle_{X,X^*}^2$ (= $\|R_{\lambda_0} f\|_{\gamma_{\lambda_0}}^2$), thus

$$\|R_{\gamma_1} f\|_\gamma^2 = \langle R_\gamma f, f\rangle_{X,X^*} - \langle R_{\lambda_0} f, f\rangle_{X,X^*} = \langle R_{\gamma_1} f, f\rangle_{X,X^*}.$$

Using the reproducing property in the Cameron–Martin space $H(\gamma_1)$, we get

$$\|R_{\gamma_1} f\|_\gamma^2 = \|R_{\gamma_1} f\|_{\gamma_1}^2.$$

Since $R_{\gamma_1}(X^*)$ is dense in $H(\gamma_1)$, we conclude that $H(\gamma_1)$ is a subspace of H_1 and, in particular, $\langle \cdot, \cdot\rangle_{\gamma_1} = \langle \cdot, \cdot\rangle_\gamma$. Finally, $H(\gamma_1) = H_1$ by density of $R_\gamma(X^*)$ in $H(\gamma)$.

(5) Using $\gamma = \gamma_{\lambda_0} * \gamma_1$, we can write for all $B \in \mathcal{B}(X)$:

$$\gamma(B) = \int_X \gamma_1(B - tx_0) \frac{e^{-t^2/\lambda_0}}{\sqrt{2\pi\lambda_0}} dt.$$

Since $f_0 \sim N(0, \lambda_0)$, we deduce that $\gamma(B|f_0 = t) = \gamma_1(B - tx_0)$ (as a regular conditional probability). ∎

Concerning the last property on conditioning, it is worth noting that the conditional covariance operator R_{γ_1} does not depend on the particular value t of the random variable $f_0 \in X^*$. We will now use both of these lemmas to build a complete decomposition of any Gaussian measure γ.

3.2. Iterative decomposition of a Gaussian measure. Consider a (centered) Gaussian measure γ on a separable Banach space X. The initial step of the decomposition is to split X and γ according to Lemma 3.2 using $f_0 \in X^*$ given by Lemma 3.1. The same process is applied to the *residual* Gaussian measure γ_1 defined in Lemma 3.2, and so on and so forth. Now, we formalize the resulting iterative decomposition scheme.

Define $\gamma_0 = \gamma$ (initialization). By induction on $n \in \mathbb{N}$ (iteration), we define the Gaussian measure γ_{n+1} of the covariance operator $R_{\gamma_{n+1}}$ such that

$$\forall f \in X^*,\ R_\gamma f = \sum_{k=0}^{n} \lambda_k \langle x_k, f\rangle_{X,X^*} x_k + R_{\gamma_{n+1}} f,$$

where $\lambda_n = \max_{f \in B_{X^*}} \langle R_{\gamma_n} f, f\rangle_{X,X^*}$ and where x_n is defined by the relation $R_{\gamma_n} f_n = \lambda_n x_n$ with f_n chosen such that $\lambda_n = \langle R_{\gamma_n} f_n, f_n\rangle_{X,X^*}$. By Lemma 3.2, we have the orthogonal decomposition for all n,

$$H(\gamma) = \operatorname{span}(h_0, \ldots, h_n) \oplus H(\gamma_{n+1}),$$

where $h_n = \sqrt{\lambda_n} x_n$. If for some n, $\lambda_{n+1} = 0$, then $R_{\gamma_{n+1}} = 0$ and $H(\gamma_{n+1}) = \{0\}$, which means that R_γ is a finite-rank operator and $H(\gamma) = \operatorname{span}(h_0, \ldots, h_n) = \operatorname{span}(x_0, \ldots, x_n)$ a finite-dimensional linear space. This means that γ is a finite-dimensional Gaussian measure with support equal to its Cameron–Martin space. Lemma 3.3 gives the properties of this decomposition in the general case where $H(\gamma)$ is infinite dimensional.

LEMMA 3.3. *Suppose $H(\gamma)$ is infinite dimensional and keep previous notation. We have the following properties*:

(1) *$(h_n)_n$ is an orthonormal sequence in $H(\gamma)$.*

(2) *$(x_n)_n$ and $(f_n)_n$ satisfy the following relations*:

$$\forall n \in \mathbb{N},\ \|x_n\|_X = \langle x_n, f_n\rangle_{X,X^*} = 1, \quad \forall (k,l) \in \mathbb{N}^2,\ k > l,\ \langle x_k, f_l\rangle_{X,X^*} = 0.$$

(3) *Let* $Q_n : h \in H(\gamma) \to Q_n h = \sum_{k=0}^{n} \langle h, h_k \rangle_\gamma h_k$ *be the orthogonal projection onto the linear space* $\operatorname{span}(h_0, \ldots, h_n) = \operatorname{span}(x_0, \ldots, x_n)$ *in* $H(\gamma)$. *Then, we have* $Q_n h = \sum_{k=0}^{n} \langle h - Q_{k-1}h, f_k \rangle_{X,X^*} x_k$ *with the convention that* $Q_{-1} = 0$.

(4) *Define* P_n *on* X *by* $P_n x = \sum_{k=0}^{n} \langle x - P_{k-1}x, f_k \rangle_{X,X^*} x_k$ *with the same convention* $P_{-1} = 0$. *Then,* P_n *is the projection onto* $\operatorname{span}(x_0, \ldots, x_n)$ *and null space* $\{x \in X : \langle x, f_k \rangle_{X,X^*} = 0 \text{ for } k = 0, \ldots, n\}$. *Furthermore, the operator* P_n *restricted to* $H(\gamma)$ *is equal to* Q_n.

(5) *According to the decomposition* $x = P_n x + (I - P_n)x$ *in* X, *the Gaussian measure* γ *can be decomposed as* $\gamma = \gamma_{\lambda_0,\ldots,\lambda_n} * \gamma_{n+1}$, *where* $\gamma_{\lambda_0,\ldots,\lambda_n} = \gamma \circ P_n^{-1}$ *is a Gaussian measure with covariance operator*

$$R_{\lambda_0,\ldots,\lambda_n} : f \in X^* \to \sum_{k=0}^{n} \lambda_k \langle x_k, f \rangle_{X,X^*} x_k.$$

Furthermore, we have $\gamma_{n+1} = \gamma \circ (I - P_n)^{-1}$ *and the relation*

$$R_\gamma = R_{\lambda_0,\ldots,\lambda_n} + R_{\gamma_{n+1}}.$$

(6) *The Cameron–Martin space* $H(\gamma)$ *is decomposed as*

$$H(\gamma) = \operatorname{span}(h_0, \ldots, h_n) \oplus H(\gamma_{n+1}),$$

where $H(\gamma_{n+1}) = (I - Q_n)\big(H(\gamma)\big)$ *equipped with the inner product of* $H(\gamma)$ *is the Cameron–Martin space of the Gaussian measure* γ_{n+1}.

(7) *Let* $x_n^* = (I - P_{n-1})^* f_n$ *for* $n \geqslant 0$. *Then, for all* n *we have* $R_\gamma x_n^* = \lambda_n x_n$. *The random variables* x_n^* *are independent* $\mathcal{N}(0, \lambda_n)$, *and*

$$\forall n,\ P_n x = \sum_{k=0}^{n} \langle x, x_k^* \rangle_{X,X^*} x_k.$$

For the computation of the dual basis $(x_n^*)_n$, *we have the recurrence formula*

$$x_n^* = f_n - P_{n-1}^* f_n \quad \text{with } P_{n-1}^* f_n = \sum_{k=0}^{n-1} \langle x_k, f_n \rangle_{X,X^*} x_k^* \text{ and } x_0^* = f_0.$$

Furthermore, $\gamma_{\lambda_0,\ldots,\lambda_n} = \gamma_{\lambda_0} * \ldots * \gamma_{\lambda_n}$, *where* γ_{λ_n} *is the distribution of the random vector* $x \to \langle x, x_n^* \rangle_{X,X^*} x_n$ *for all* n.

(8) *Let* $h_n^* = (\sqrt{\lambda_n})^{-1} x_n^*$ *for* $n \geqslant 0$. *Then, we have* $R_\gamma h_n^* = h_n$, *and the random variables* h_n^* *are independent* $\mathcal{N}(0, 1)$.

(9) *For each* $t = (t_0, \ldots, t_n) \in \mathbb{R}^{n+1}$, *denote by* $\sum_{k=0}^{n} t_k x_k + \gamma_{n+1}$ *the Gaussian measure on* X *centered at* $\sum_{k=0}^{n} t_k x_k$ *with covariance operator* $R_{\gamma_{n+1}}$. *Then,* $\gamma^t = \sum_{k=0}^{n} t_k x_k + \gamma_{n+1}$ *is the conditional probability distribution of* $x \in X$ *given* $x_0^*(x) = t_0, \ldots, x_n^*(x) = t_n$:

$$\forall B \in \mathcal{B}(X),\ \gamma^t(B) = \gamma_{n+1}\big(B - \sum_{k=0}^{n} t_k x_k\big) = \gamma(B | x_0^* = t_0, \ldots, x_n^* = t_n).$$

The deconditioning formula is

$$\gamma(B) = \int_{\mathbb{R}^n} \gamma^t(B) \prod_{k=0}^{n} \frac{e^{-t_k^2/(2\lambda_k)}}{\sqrt{2\pi\lambda_k}} dt_k.$$

P r o o f. (1) For $n \in \mathbb{N}$, $\|h_n\|_\gamma = 1$ by construction. If $n < m$, remark that

$$h_n \in \operatorname{span}(h_0, \ldots, h_{m-1}) = H(\gamma_m)^\perp$$

to get $\langle h_n, h_m \rangle_\gamma = 0$.

(2) By the definition of x_n, we have $\langle x_n, f_n \rangle_{X,X^*} = 1$. Now, the reproducing property gives

$$\forall f \in B_{X^*},\ \langle x_n, f \rangle_{X,X^*} = \langle x_n, R_{\gamma_n} f \rangle_{\gamma_n} \leqslant \|x_n\|_{\gamma_n} \sqrt{\langle R_{\gamma_n} f, f \rangle_{X,X^*}}.$$

Using the relations $\langle R_{\gamma_n} f, f \rangle_{X,X^*} \leqslant \lambda_n$ and $\|\sqrt{\lambda_n} x_n\|_{\gamma_n} = \|h_n\|_\gamma = 1$, we get $\langle x_n, f \rangle_{X,X^*} \leqslant 1$. This proves that $\|x_n\|_X = \langle x_n, f_n \rangle_{X,X^*} = 1$.

For $k > l$, $h_k \in H(\gamma_l)$ and the reproducing property gives

$$\sqrt{\lambda_k} \langle x_k, f_l \rangle_{X,X^*} = \langle h_k, R_{\gamma_l} f_l \rangle_{\gamma_l} = \sqrt{\lambda_l} \langle h_k, h_l \rangle_\gamma = 0.$$

Hence $\langle x_k, f_l \rangle_{X,X^*} = 0$ since $\lambda_k > 0$.

(3) For $h \in H(\gamma)$, we have

$$Q_n h = \sum_{k=0}^{n} \langle h, \lambda_k x_k \rangle_\gamma x_k = \sum_{k=0}^{n} \langle h, R_{\gamma_k} f_k \rangle_\gamma x_k.$$

According to the orthogonal decomposition

$$H(\gamma) = \operatorname{span}(h_0, \ldots, h_{k-1}) \oplus H(\gamma_k),$$

we get

$$\langle h, R_{\gamma_k} f_k \rangle_\gamma = \langle h - Q_{k-1} h, R_{\gamma_k} f_k \rangle_{\gamma_k} = \langle h - Q_{k-1} h, f_k \rangle_{X,X^*},$$

which proves the result.

(4) Let $x \in X$. Then $P_n x \in \operatorname{span}(x_0, \ldots, x_n) = \operatorname{range}(Q_n)$, thus $P_n x \in H(\gamma)$ and $P_n(P_n x) = Q_n(P_n x) = P_n x$. Clearly, we have

$$\bigcap_{k=0}^{n} \ker(f_k) \subset \ker(P_n).$$

Conversely, if $P_n x = 0$, then $P_k x = 0$ for all $k \in [0, n]$ and $0 = \langle x - P_{k-1} x, f_k \rangle = \langle x, f_k \rangle_{X,X^*}$, hence $\ker(P_n) \subset \bigcap_{k=0}^{n} \ker(f_k)$.

(5) Since $Q_n = P_n$ on $H(\gamma)$, remark first that $R_{\gamma_{\lambda_0,\dots,\lambda_n}} = R_\gamma P_n^* = Q_n R_\gamma$ and also $R_{\gamma_{n+1}} = R_\gamma(I - P_n)^* = (I - Q_n)R_\gamma$. In particular,

$$R_{\gamma_{\lambda_0,\dots,\lambda_n}} f = \sum_{k=0}^{n} \langle R_\gamma f, h_k\rangle_\gamma h_k = \sum_{k=0}^{n} \lambda_k \langle x_k, f\rangle_{X,X^*} x_k.$$

Consider now the decomposition for $f \in X^*$:

$$f = P_n^* f + (I - P_n)^* f.$$

The random variable $P_n^* f$ is Gaussian with variance $\langle R_{\gamma_{\lambda_0,\dots,\lambda_n}} f, f\rangle_{X,X^*}$ and $(I - P_n)^* f$ is Gaussian with variance $\langle R_{\gamma_{n+1}} f, f\rangle_{X,X^*}$. Since

$$\langle R_\gamma P_n^* f, (I - P_n)^* f\rangle_{X,X^*} = \langle (I - Q_n) Q_n R_\gamma f, f\rangle_{X,X^*} = 0,$$

the random variables $P_n^* f$ and $(I - P_n)^* f$ are independent and we conclude as in Lemma 3.2.

(6) The proof is similar to the proof of (4) in Lemma 3.2. Introduce the space $H_{n+1} = \operatorname{span}(h_0, \dots, h_n)^\perp$; then we see that $\big(R_{\gamma_{n+1}}(X^*), \langle\cdot,\cdot\rangle_{\gamma_{n+1}}\big)$ is a subspace of H_{n+1}, which is sufficient to prove $H(\gamma_{n+1}) = H_{n+1}$ as Hilbert spaces.

(7) For $n \geqslant 0$ and $h \in H(\gamma)$, we write $\langle h, R_\gamma x_n^*\rangle_\gamma = \langle h, (I - P_{n-1})^* f_n\rangle_{X,X^*}$, thus $\langle h, R_\gamma x_n^*\rangle_\gamma = \langle (I - Q_{n-1})h, f_n\rangle_{X,X^*} = \langle (I - Q_{n-1})h, R_{\gamma_n} f_n\rangle_\gamma$. Using now the relation $R_{\gamma_n} f_n = \lambda_n x_n$, we finally get $\langle h, R_\gamma x_n^*\rangle_\gamma = \langle h, \lambda_n x_n\rangle_\gamma$, which proves $R_\gamma x_n^* = \lambda_n x_n$. In particular, $\langle R_\gamma x_n^*, x_n^*\rangle_{X,X^*} = \lambda_n$. In the same way, we get $\langle R_\gamma x_m^*, x_n^*\rangle_{X,X^*} = 0$ if $m \neq n$. Hence, the random variables x_n^* are independent with respective variance λ_n. The computation of this sequence comes from the identity $P_n^* f = \sum_{k=0}^n \langle x_k, f\rangle_{X,X^*} x_k^*$.

(8) This is a reformulation of the previous statement about the sequence $(x_n^*)_n$.

(9) This last assertion is a direct consequence of (5) and (7). ∎

Interpretation of the pairs (λ_n, x_n) is the following: for $n = 0$, x_0 is a (unit) direction vector for a line in X that has the largest variance possible ($= \lambda_0$) by a projection of norm one (namely, the projection P_0 in Lemma 3.3). Remark that P_0 of norm one means P_0 orthogonal or self-adjoint in the Hilbert case. By considering the measure $\gamma_1 = \gamma \circ (I - P_0)^{-1}$, the vector x_1 is the direction vector for a line in the subspace $(I - P_0)X$ that has the largest variance possible and so on. In the Hilbert case, this decomposition process is known as (functional) principal component analysis.

In this paper, we have supposed the Radon measure γ to be Gaussian. By a slight modification of the proof of Lemma 3.1, the decomposition is valid if we assume only $\int_X \|x\|^2 \gamma(dx) < +\infty$ and results have to be interpreted in a mean-square sense (in particular, independence becomes non-correlation and the last

parts of Lemmas 3.2 and 3.3 on conditioning are valid only in the Gaussian case). Now, it remains to see that this decomposition is complete, namely that we have

$$\gamma = *_n \gamma_{\lambda_n}$$

according to the decomposition of the covariance operator

$$R_\gamma = \sum_n \lambda_n \langle x_n, \cdot \rangle_{X,X^*} x_n.$$

3.3. Asymptotic analysis and the main result. In this section, we suppose that $H(\gamma)$ is infinite dimensional and we use the notation of the previous section. The following two lemmas will be essential for the main result of this paper (Theorem 3.1).

LEMMA 3.4. *We have* $B_{H(\gamma_n)} = B_{H(\gamma)} \cap \operatorname{span}(h_0, \dots, h_{n-1})^\perp$ *for all* n *and*

$$\sqrt{\lambda_n} = \sup_{f \in B_{X^*}} \sup_{h \in B_{H(\gamma_n)}} \langle h, f \rangle_{X,X^*}. \tag{3.1}$$

P r o o f. Since $H(\gamma) = \operatorname{span}(h_0, \dots, h_{n-1}) \oplus H(\gamma_n)$ and $\|\cdot\|_{\gamma_n} = \|\cdot\|_\gamma$ on $H(\gamma_n)$ (see Lemma 3.3, assertion (6)), we get

$$B_{H(\gamma)} \cap \operatorname{span}(h_0, \dots, h_{n-1})^\perp = B_{H(\gamma_n)}.$$

But, for $h \in H(\gamma_n)$, $\langle h, f \rangle_{X,X^*} = \langle h, R_{\gamma_n} f \rangle_{\gamma_n}$ and $\sup_{h \in B_{H(\gamma_n)}} \langle h, f \rangle_{X,X^*}$ is attained for $h = R_{\gamma_n} f / \|R_{\gamma_n} f\|_{\gamma_n}$ (if $R_{\gamma_n} f \neq 0$). Thus,

$$\sup_{h \in B_{H(\gamma_n)}} \langle h, f \rangle_{X,X^*} = \sqrt{\langle R_{\gamma_n} f, f \rangle_{X,X^*}}. \quad \blacksquare$$

LEMMA 3.5. *The sequence* $(\lambda_n)_n$ *is non-increasing and* $\lambda_n \to 0$.

P r o o f. By Lemma 3.4 and (3.1), we see that $\lambda_{n+1} \leqslant \lambda_n$. Moreover, (h_n) is an orthonormal system in $H(\gamma)$, hence

$$\forall f \in X^*, \ \langle h_n, f \rangle_{X,X^*} = \langle h_n, R_\gamma f \rangle_\gamma \to 0,$$

as a consequence of Bessel's inequality. In other words, we have $h_n \rightharpoonup 0$ for the weak topology of X. Since the unit ball of $H(\gamma)$ is precompact in X ([2], Corollary 3.2.4, p. 101), we can extract a subsequence $(h_{n_k})_k$ such that $h_{n_k} \to_k h_\infty$ for the strong topology of X. By unicity of limit in the topological vector space X equipped with the weak topology, we deduce that $h_\infty = 0$ in X. Therefore, $\|h_{n_k}\|_X = \sqrt{\lambda_{n_k}} \to_k 0$, which completes the proof. $\blacksquare$

The above two lemmas are the final ingredients to prove the main result of the paper (Theorem 3.1), namely that the orthonormal family $(h_n)_n$ is a Hilbert basis of $H(\gamma)$ in $R_\gamma(X^*)$ as it is discussed in [17].

THEOREM 3.1. *Let γ be a centered Gaussian measure on a separable Banach space $\big(X, \mathcal{B}(X)\big)$ with covariance operator R_γ and Cameron–Martin space $H(\gamma)$. Consider the sequence $(x_n)_n$ of unit direction vectors associated with "principal" functionals $(x_n^*)_n$ with non-increasing variance $(\lambda_n)_n$. Let $h_n^* = \big(\sqrt{\lambda_n}\big)^{-1} x_n^*$ and $h_n = R_\gamma h_n^*$. Then, $(h_n)_n$ is a Hilbert basis of $H(\gamma)$.*

P r o o f. Let $h \in H(\gamma)$ such that for all $n \in \mathbb{N}$, $\langle h, h_n\rangle_\gamma = 0$. Then, using Lemma 3.4, we have

$$\forall n \in \mathbb{N},\ \forall f \in B_{X^*},\ \langle h, f\rangle_{X,X^*} \leqslant \sqrt{\lambda_n}\|h\|_\gamma,$$

which implies $\langle h, f\rangle_{X,X^*} = 0$ for all $f \in X^*$. Thus, $h = 0$ and $\operatorname{span}(h_n,\ n \geqslant 0)$ is dense in $H(\gamma)$. ∎

We give now the claimed two results of this paper.

COROLLARY 3.1. *The covariance operator R_γ can be decomposed as*

$$R_\gamma = \sum_n \lambda_n \langle x_n, \cdot\rangle_{X,X^*} x_n,$$

where the convergence is in $\mathcal{L}(X^*, X)$. More precisely, the nth step truncation error is

$$\Big\|R_\gamma - \sum_{k=0}^{n} \lambda_k \langle x_k, \cdot\rangle_{X,X^*} x_k\Big\| = \lambda_{n+1},$$

where $\|\cdot\|$ stands for the operator norm in $\mathcal{L}(X^*, X)$.

P r o o f. From Theorem 3.1 we know that $(h_n)_n$ is a Hilbert basis of $H(\gamma)$. It suffices to write

$$\forall f \in X^*,\ R_\gamma f = \sum_n \langle R_\gamma f, h_n\rangle_\gamma h_n,$$

and use the reproducing property. The truncation error norm is

$$\Big\|R_\gamma - \sum_{k=0}^{n} \lambda_k \langle x_k, .\rangle_{X,X^*} x_k\Big\| = \sup_{f \in B_{X^*}} \|R_{\gamma_{n+1}} f\|_X.$$

However,

$$\|R_{\gamma_{n+1}} f\|_X = \sup_{g \in B_{X^*}} \langle R_{\gamma_{n+1}} f, g\rangle_{X,X^*} \leqslant \lambda_{n+1}$$

by the Cauchy–Schwarz inequality. Since $R_{\gamma_{n+1}} f_{n+1} = \lambda_{n+1} x_{n+1}$, $\|x_{n+1}\|_X = 1$, we have $\|R_{\gamma_{n+1}} f_{n+1}\|_X = \lambda_{n+1}$. Hence

$$\Big\|R_\gamma - \sum_{k=0}^{n} \lambda_k \langle x_k, \cdot\rangle_{X,X^*} x_k\Big\| = \lambda_{n+1} \to 0. \quad ∎$$

For γ a Gaussian measure on a separable Hilbert space X, Corollary 3.1 is equivalent to the spectral theorem applied to the self-adjoint compact operator R_γ. In the Banach case, Corollary 3.1 says that

$$R_\gamma = \sum_n \lambda_n \langle x_n, \cdot \rangle_{X,X^*} x_n,$$

where $(\lambda_n)_n$ is a non-increasing sequence that converges to zero and $(x_n)_n$ is a sequence of unit norm vectors in X and orthogonal in $H(\gamma)$. Furthermore, we have the same formula for the error. The Gaussian hypothesis is motivated by applications both in Gaussian process regression (or Kriging, see [14]) and Bayesian inverse problems (see [16]). As Lemma 3.3 indicates, we are interested in an efficient algorithm to construct a design of experiments (see [4]) or a training set (functionals $(f_n)_n$ or, equivalently, $(x_n^*)_n$). The error expression

$$\Big\| R_\gamma - \sum_{k=0}^{n} \lambda_k \langle x_k, \cdot \rangle_{X,X^*} x_k \Big\| = \sup_{f \in B_{X^*}} \| R_{\gamma_{n+1}} f \|_X = \lambda_{n+1}$$

in Corollary 3.1 says that we have a precise quantification of uncertainty in terms of confidence interval in the Gaussian case.

COROLLARY 3.2. *Let us remind the definition* $h_n^* = \left(\sqrt{\lambda_n}\right)^{-1} x_n^*$ *with* $x_n^* = (I - P_{n-1})^* f_n$ *for* $n \geqslant 1$ *and* $x_0^* = f_0$. *Then, we have the decomposition in* X:

$$x = \sum_n \langle x, h_n^* \rangle_{X,X^*} h_n, \ \gamma \ a.e.,$$

where the random variables h_n^* *are independent* $\mathcal{N}(0,1)$. *In the equivalent form, let* $(\xi_n)_n$ *be a sequence of independent standard normal variables on* $(\Omega, \mathcal{F}, \mathbb{P})$. *Then the random series*

$$\sum_n \sqrt{\lambda_n} \xi_n(\omega) x_n$$

defines an X*-valued random Gaussian vector with distribution* γ.

The random series representation $\sum_n \sqrt{\lambda_n} \xi_n(\omega) x_n$ is a generalization of the Karhunen–Loève expansion based on the corresponding decomposition of the covariance operator $R_\gamma = \sum_n \lambda_n \langle x_n, \cdot \rangle_{X,X^*} x_n$.

4. EXAMPLES

4.1. Decomposition of the classical Wiener measure. Let γ be the standard Wiener measure on $X = \mathcal{C}([0,1], \mathbb{R})$, the space of all real continuous functions on the interval $[0,1]$ which is a Banach space if equipped with the supremum norm. The Riesz–Markov representation theorem allows us to identify X^* with the linear space of all bounded signed measures on $[0,1]$ equipped with the norm of total variation. In this context, the dual pairing is

$$\forall x \in X, \ \forall \mu \in X^*, \ \langle x, \mu \rangle_{X,X^*} = \int_0^1 x(t) \mu(dt).$$

The Cameron–Martin space associated with γ is the Sobolev-type space $H_0^1([0,1],\mathbb{R})$ defined by

$$H_0^1([0,1],\mathbb{R}) = \big\{ f \in X,\ \forall t \in [0,1],\ f(t) = \int_0^t f'(s)ds,\ f' \in L^2([0,1],\mathbb{R}) \big\}$$

and the associated inner product $\langle f_1, f_2 \rangle_\gamma = \langle f_1', f_2' \rangle_{L^2}$. The covariance operator R_γ satisfies

$$\langle R_\gamma \mu, \mu \rangle_{X,X^*} = \mathrm{Var}\big(\int_0^1 W_t \mu(dt) \big),$$

where $(W_t)_{t\in[0,1]}$ is the standard Wiener process. Using Fubini's theorem, we easily get

$$\langle R_\gamma \mu, \mu \rangle_{X,X^*} = \iint_{[0,1]^2} t \wedge s \mu(dt)\mu(ds) = \int_0^1 \mu([u,1])^2 du.$$

Hence, $(R_\gamma \mu)'(t) = \mu([t,1])$ almost everywhere in $[0,1]$, and $R_\gamma \mu : t \in [0,1] \to \int_0^t \mu([u,1])du$. Consider now the initial step of the decomposition, that is, find $f_0 = \mu_0 \in B_{X^*}$ such that

$$\langle R_\gamma \mu_0, \mu_0 \rangle_{X,X^*} = \sup_{\mu \in B_{X^*}} \langle R_\gamma \mu, \mu \rangle_{X,X^*}.$$

Since $\forall \mu \in B_{X^*}, \forall u \in [0,1],\ |\mu([u,1])| \leqslant 1$, the unique measure (up to sign) into B_{X^*} maximizing $\langle R_\gamma \mu, \mu \rangle_{X,X^*} = \int_0^1 \mu([u,1])^2 du$ is $\mu_0 = \delta_1$. Moreover,

$$\lambda_0 = \langle R_\gamma \mu_0, \mu_0 \rangle_{X,X^*} = \mathrm{Var}(W_1) = 1$$

is the variance of the Wiener process at the point $t = 1$. Since $\mu \to \langle R_\gamma \mu, \mu \rangle_{X,X^*}$ is a non-negative quadratic functional, a usual argument shows directly that μ_0 must be an extremal point of B_{X^*}. Thus $\mu_0 = \delta_{t_0}$ for some point $t_0 \in [0,1]$. And clearly, $t_0 = 1$, corresponding to the maximum of variance of the Wiener process. So, we have $\lambda_0 = 1$, $f_0 = \mu_0 = \delta_1$. Using the fact that

$$R_\gamma \delta_t : s \in [0,1] \to \langle R_\gamma \delta_t, \delta_s \rangle_{X,X^*} = \mathrm{Cov}(W_t, W_s) = t \wedge s,$$

we get $x_0 = (t \in [0,1] \to t)$ and $h_0 = x_0$ (since $\lambda_0 = 1$). Now, we have $P_0 x : t \in [0,1] \to \langle x, f_0 \rangle_{X,X^*} x_0(t) = x(1)t$ and $(I - P_0)x$ is the function $t \in [0,1] \to x(t) - x(1)t$. Consequently, we see that $\gamma_1 = \gamma \circ (I - P_0)^{-1}$ is the Gaussian measure associated with the Brownian bridge $(B_t)_{t\in[0,1]}$ with covariance kernel

$$K_1 : (t,s) \in [0,1]^2 \to \mathrm{Cov}(B_t, B_s) = t \wedge s - ts.$$

Using now the fact that $\mu \to \langle R_{\gamma_1} \mu, \mu \rangle_{X,X^*}$ is a non-negative quadratic functional, we see that $f_1 = \mu_1 = \delta_{t_1}$, where $t_1 = \frac{1}{2}$ is the maximum of variance of

the Brownian bridge B. Hence, we get $\lambda_1 = \frac{1}{4}$, $x_1 = \big(t \to 4\big(t \wedge \frac{1}{2} - \frac{t}{2}\big)\big)$ (by the relation $\lambda_1 x_1 = R_{\gamma_1}\delta_{1/2}$) and $h_1 = \frac{1}{2}x_1$. Furthermore, $x_1^* = \delta_{t_1} - \frac{1}{2}\delta_{t_0}$ and $\gamma_2 = \gamma \circ (I - P_1)^{-1}$ is the Gaussian distribution of the process $(I - P_1)W : t \to W_t - W_1 x_0(t) - \big(W_{1/2} - \frac{1}{2}W_1\big)x_1(t)$. By the assertion (9) of Lemma 3.3, γ_2 is the conditional distribution of W given $W_1 = 0, W_{1/2} = 0$. Using this interpretation, scale-invariance and spatial Markov properties of the Wiener process, we immediately get

$$\lambda_n = \frac{1}{2^{p+2}} \qquad \text{for } n = 2^p + k,\ k = 0, \dots, 2^p - 1 \text{ and } p \geqslant 0.$$

Furthermore, the Hilbert basis $(h_n)_n$ of $H(\gamma)$ is given by $h_0(t) = t$ and $h_n(t) = \int_0^t h_n'(s)ds,\ n \geqslant 1$, where

$$h_n'(s) = \begin{cases} \sqrt{2^p} & \text{for } 2k/2^{p+1} \leqslant s \leqslant (2k+1)/2^{p+1}, \\ -\sqrt{2^p} & \text{for } (2k+1)/2^{p+1} < s \leqslant (2k+2)/2^{p+1}, \\ 0 & \text{otherwise} \end{cases}$$

if $n = 2^p + k,\ k = 0, \dots, 2^p - 1$ and $p \geqslant 0$. The family $(h_n')_{n\in\mathbb{N}}$ is the usual Haar basis of $L^2([0,1], \mathbb{R})$. The functions $(x_n)_n$ are Schauder's functions

$$x_n(t) = \sqrt{2^{p+2}} h_n(t)$$

corresponding to *hat functions* of height one and lying above the intervals $\big[\frac{k}{2^p}, \frac{k+1}{2^p}\big]$

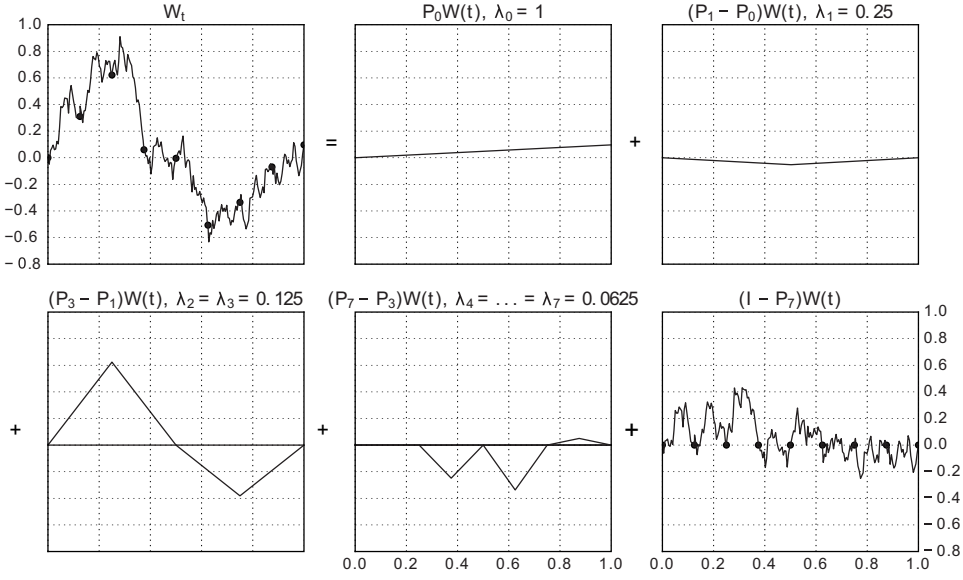

FIGURE 1. Decomposition of the standard Wiener measure on the eight first steps.

($n = 2^p + k$). The resulting decomposition $\sum_n \sqrt{\lambda_n}\xi_n(\omega)x_n$ is the famous Lévy–Ciesielski construction of Brownian motion on the interval $[0, 1]$ (see [13]). The eight first steps (and the associated residual) of this decomposition are illustrated in Figure 1.

Remark that

$$\sum_n \lambda_n = 1 + \frac{1}{4} + 2 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + \ldots = +\infty$$

due to the "multiplicity" of the values λ_n. In the Hilbert case, this sum is always finite and is the trace of the operator R_γ. Furthermore, this finite-trace property is characteristic of Gaussian measures on Hilbert spaces. Such a characterization in the Banach case is still an open problem.

4.2. Expansion of the Ornstein–Uhlenbeck process. Let us consider now the Ornstein–Uhlenbeck process (Z_t) on the time interval $[0, 1]$ defined as the stationary solution of the following stochastic differential equation:

$$dZ_t = -\beta Z_t dt + \sigma dW_t, \quad \beta > 0,\ \sigma > 0. \tag{4.1}$$

One can show that the solution is a centered, continuous Gaussian process with covariance kernel

$$K(t, s) = \mathrm{Cov}(Z_t, Z_s) = \frac{\sigma^2}{2\beta} e^{-\beta|t-s|}.$$

This kernel is also known as the exponential covariance kernel or Matérn covariance kernel of order $\nu = \frac{1}{2}$. To get an expansion of the corresponding process, we consider its distribution as a Gaussian measure γ on the Banach space $X = \mathcal{C}([0, 1], \mathbb{R})$ as we did for the Brownian motion. Taking $\sigma^2 = 2\beta$ (with no loss of generality), we have $\langle R_\gamma \delta_t, \delta_s \rangle_{X,X^*} = e^{-\beta|t-s|}$, where R_γ is the covariance operator of γ. In comparison with the previous example, any extremal point δ_t ($t \in [0, 1]$) maximizes the function $\mu \in \mathcal{B}_{X^*} \to \langle R_\gamma \mu, \mu \rangle_{X,X^*}$ (since the process has constant variance). Here, it is natural to choose $f_0 = \delta_0$ associated with an initial condition of equation (4.1). It immediately follows that $\lambda_0 = 1$ and $x_0 : t \to e^{-\beta t} = K(t, 0)$. The conditional Gaussian measure γ_1 is the distribution of the centered Gaussian process with covariance kernel

$$K_1 : (t, s) \in [0, 1]^2 \to e^{-\beta|t-s|} - e^{-\beta t} e^{-\beta s}.$$

The maximum of $t \to K_1(t, t)$ is obtained with $f_1 = \mu_1 = \delta_1$ and we get

$$\lambda_1 = 1 - e^{-2\beta},$$
$$x_1(t) = \frac{1}{\lambda_1}(e^{-\beta(1-t)} - e^{-\beta} e^{-\beta t}).$$

The next iteration will be the last needed to obtain the full decomposition of the process. The conditional Gaussian measure γ_2 is of covariance kernel

$$K_2(t,s) = K_1(t,s) - \lambda_1^{-1} K_1(t,1) K_1(s,1).$$

A straightforward computation shows that the variance function $t \to K_2(t,t)$ is maximum at $t = \frac{1}{2}$ that leads to $f_2 = \mu_2 = \delta_{1/2}$ and

$$\lambda_2 = 1 - 2\frac{e^{-\beta}}{1+e^{-\beta}},$$

$$x_2(t) = \frac{1}{\lambda_2}\left(e^{-\beta(t-1/2)} - e^{-\beta/2}\frac{e^{-\beta t} + e^{-\beta(1-t)}}{1+e^{-\beta}}\right).$$

From this and since the Ornstein–Uhlenbeck process is Markovian, we deduce the shape of all other functions. Indeed, all further steps will be given on dyadic intervals of the form $\left[\frac{k-1}{2^p}, \frac{k}{2^p}\right]$ with $k \in [1, \ldots, 2^p]$, the process being independent of these intervals. Here is the general form of the basis for $n = 2^p + k$ with $p \geqslant 0$ and $k \in [1, \ldots, 2^p]$:

$$\lambda_n = 1 - 2\frac{e^{-\beta/2^p}}{1+e^{-\beta/2^p}},$$

$$x_n(t) = \frac{1}{\lambda_n}\left[e^{-\beta(t-(k-1/2)/2^p)} - e^{-\beta/2^{p+1}}\frac{e^{-\beta(t-(k-1)/2^p)} + e^{-\beta(k/2^p-t)}}{1+e^{-\beta/2^p}}\right].$$

This gives an analytical (weak) solution $\frac{\sigma}{\sqrt{2\beta}}\sum_n \sqrt{\lambda_n}\xi_n x_n$ of the Langevin stochastic differential equation (equation (4.1)).

5. CONCLUSION

In this paper, we suggest a Karhunen–Loève expansion for a Gaussian measure on a separable Banach space based on a corresponding decomposition of its covariance operator. In some sense, this decomposition generalizes the Hilbert case. Lévy–Ciesielski's construction of Brownian motion appears to be a particular case of such an expansion. Finally, we believe that this result will be useful both in pure and applied mathematics since it provides a canonical representation of Gaussian measures on separable Banach spaces.

Acknowledgments. The authors would like to thank the anonymous referee for his careful reading of the paper, suggestions to improve clarity of the initial text and his recommendation for a second analytical illustration.

REFERENCES

[1] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. 68 (1950), pp. 337–404.
[2] V. I. Bogachev, *Gaussian Measures*, American Mathematical Society, Providence, RI, 1998.
[3] G. Da Prato, *An Introduction to Infinite-Dimensional Analysis*, Springer, Berlin 2006.
[4] V. V. Fedorov, *Theory of Optimal Experiments*, Academic Press, New York–London 1972.
[5] M. Hairer, *An Introduction to Stochastic PDEs*, The University of Warwick / Courant Institute, Lecture notes, 2009.
[6] N. C. Jain and G. Kallianpur, *Norm convergent expansions for Gaussian processes in Banach spaces*, Proc. Amer. Math. Soc. 25 (1970), pp. 890–895.
[7] K. Karhunen, *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, Ann. Acad. Sci. Fenn. Ser. AI Math.-Phys. 37 (1947), pp. 1–79.
[8] J. Kuelbs, *Expansions of vectors in a Banach space related to Gaussian measures*, Proc. Amer. Math. Soc. 27 (1971), pp. 364–370.
[9] H.-H. Kuo, *Gaussian Measures in Banach Spaces*, Springer, Berlin–New York 1975.
[10] V. V. Kvaratskhelia, *Unconditional convergence of functional series in problems of probability theory*, J. Math. Sci. (N.Y.) 200 (2) (2014), pp. 143–294.
[11] S. Kwapień and B. Szymański, *Some remarks on Gaussian measures in Banach spaces*, Probab. Math. Statist. 1 (1) (1980), pp. 59–65.
[12] M. Loève, *Probability Theory*, second edition, D. Van Nostrand Co., Inc., Princeton, NJ, 1960.
[13] H. P. McKean Jr., *Stochastic Integrals*, Academic Press, New York 1969.
[14] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, MA, 2006.
[15] L. Schwartz, *Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants)*, J. Anal. Math. 13 (1964), pp. 115–256.
[16] A. M. Stuart, *Inverse problems: A Bayesian perspective*, Acta Numer. 19 (2010), pp. 451–559.
[17] N. Vakhania, *Canonical factorization of Gaussian covariance operators and some of its applications*, Theory Probab. Appl. 38 (3) (1994), pp. 498–505.
[18] N. N. Vakhania, V. I. Tarieladze, and S. A. Chobanyan, *Probability Distributions on Banach Spaces*, D. Reidel Publishing Co., Dordrecht 1987.

Xavier Bay
Mines Saint-Étienne, Institut Fayol
Génie Mathématique et Industriel
LIMOS UMR 6158
158 cours Fauriel, 42023 Saint-Étienne, France
E-mail: bay@emse.fr

Jean-Charles Croix
Mines Saint-Étienne, Institut Fayol
Génie Mathématique et Industriel
LIMOS UMR 6158
158 cours Fauriel, 42023 Saint-Étienne, France
E-mail: jean-charles.croix@emse.fr

Received on 16.6.2017;
revised version on 3.2.2018