TECHNISCHE UNIVERSITÄT CHEMNITZ

On the Dimensionality of the Stochastic Space in the Stochastic Finite Element Method

H.-J. Starkloff

Preprint 2007-29

Fakultät für Mathematik

Preprintreihe der Fakultät für Mathematik
ISSN 1614-8835

On the Dimensionality of the Stochastic Space in the Stochastic Finite Element Method

Hans-Jörg Starkloff

University of Applied Sciences
Zwickau (Germany)

18th December 2007

Abstract

In recent works concerning the solution of various kinds of random equations or the stochastic simulation of random functions often so called (generalized) polynomial chaos expansions are used. Hereby one step is the representation of random variables through independent random variables with specific distributions, e.g., Gaussian variables. The present work addresses the questions how many such variables are needed and what kind of distributions can be generated in such a way. It is shown, that allowing arbitrary measurable transformations, usually one can generate the needed random variables with the help of only one random variable with continuous distribution function, e.g., one standard Gaussian random variable.

Key words polynomial chaos; Gaussian Hilbert space; measurable transformation; random dimension; Monte Carlo methods

MSC (2000) 60E05 ; 60H35 ; 60G12

1 Introduction

In recent time various kinds of polynomial chaos expansions are used to represent random variables or random functions, e.g. in the solution of random equations or in the stochastic simulation of random functions. These polynomial chaos expansion are based on a sequence of orthogonal polynomials with respect to a probability distribution. Although in principle for arbitrary probability distributions with finite moments such sequences of orthogonal polynomials exist and can be taken as a tool for constructing expansions mostly some specific probability distributions are chosen for this task. Among the absolutely continuous distributions there are the uniform, Beta and Gamma distributions. Historically the first and perhaps the most important is the Gaussian distribution. This is partly due to the properties of Gaussian distributions which allow a relatively easy modelling of random functions. All the mentioned probability distributions lead to a family of classical orthogonal polynomials (in a wider sense).

So in many situations one has to consider transformations of random variables or vectors and work with or investigate the corresponding polynomial chaos expansions. In [6] for example the accuracy of such expansions is investigated. Hereby the starting point is the class of random vectors in $\mathbb{R}^d$, which can be written as a deterministic measurable function of a standard Gaussian random vector in $\mathbb{R}^k$. In this situation there are at least two aspects of interest.

1. **Measurability.** A standard Gaussian vector $\boldsymbol{\eta}$ in $\mathbb{R}^k$ is assumed to be given, hence also an underlying probability space $(\Omega, \mathcal{A}, \mathbf{P})$ is implicitly or explicitly given on which this random vector is defined. Assume that we consider a random vector $\boldsymbol{\xi}$ in $\mathbb{R}^d$ which can be written as $\boldsymbol{\xi} = g(\boldsymbol{\eta})$ with a deterministic measurable function $g : \mathbb{R}^k \to \mathbb{R}^d$. Then it follows from elementary measurability properties that $\boldsymbol{\xi}$ is defined on the same basic probability space $(\Omega, \mathcal{A}, \mathbf{P})$, and moreover it is measurable with respect to the σ-algebra $\sigma(\boldsymbol{\eta})$, generated by the random vector $\boldsymbol{\eta}$ on Ω (it is a sub-σ-algebra of $\mathcal{A}$ on Ω). The so called Doob-Dynkin theorem (see e.g. [10], Lemma 1.13) then states that also the converse is true: If $\widetilde{\boldsymbol{\xi}}$ is an $\mathbb{R}^d$-valued random vector, defined on the same basic probability space $(\Omega, \mathcal{A}, \mathbf{P})$ and $\widetilde{\boldsymbol{\xi}}$ is measurable with respect to the σ-algebra $\sigma(\boldsymbol{\eta})$, then there exists a measurable function $\widetilde{g} : \mathbb{R}^k \to \mathbb{R}^d$ with $\widetilde{\boldsymbol{\xi}} = \widetilde{g}(\boldsymbol{\eta})$. Such measurability questions can play a role in the theory of random equations, so e.g. one distinguishes in the investigation of stochastic Itô differential equations between strong and weak solutions (in the stochastic sense), for the strong solutions the probability space is given, for weak solution it can be chosen in some way.

2. **Distributions.** For many problems only the distributions of the two random vectors are given or searched, not the underlying probability space and the specific mappings. Thereby different questions can be posed.

   - The probability distribution of the random vector $\boldsymbol{\eta}$ is known, also the measurable function $g : \mathbb{R}^k \to \mathbb{R}^d$ is given. Then the distribution of the random vector $\boldsymbol{\xi}$ is uniquely determined and one can ask for this probability distribution or some statistical characteristics such as moments or tail probabilities.
   - Generalizing the previous situation we assume that we know the probability distribution of the random vector $\boldsymbol{\eta}$. Now we want to describe all possible probability distributions in $\mathbb{R}^d$ of random vectors $g(\boldsymbol{\eta})$ where the deterministic functions

$g : \mathbb{R}^k \longrightarrow \mathbb{R}^d$ belong to a class of measurable functions with some properties, e.g. the most general class of all measurable functions from $\mathbb{R}^k$ to $\mathbb{R}^d$.

- If the probability distributions $P_{\boldsymbol{\xi}}$ and $P_{\boldsymbol{\eta}}$ of random vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, respectively, are given (the underlying probability spaces do not play here any role, especially they may be distinct), then an important question is to find, if it exists, a suitable measurable function $g : \mathbb{R}^k \longrightarrow \mathbb{R}^d$, such that for random vectors $\widetilde{\boldsymbol{\xi}}$ and $\widetilde{\boldsymbol{\eta}}$ on one common probability space and satisfying $\widetilde{\boldsymbol{\xi}} = g(\widetilde{\boldsymbol{\eta}})$ it holds for the corresponding distributions $P_{\boldsymbol{\xi}} = P_{\widetilde{\boldsymbol{\xi}}}$ and $P_{\boldsymbol{\eta}} = P_{\widetilde{\boldsymbol{\eta}}}$. Usually such a transformation g is not unique so that also the question of finding such a transformation with special properties can be of great interest.

In this article we will not deal with this last question. We will concentrate on the question of finding the class of all probability distributions of random vectors, which are measurable transformations of a random vector with given distribution. Especially we will ask how the dimensionality of the random vectors (expressed in the possibly different dimensionality of the spaces $\mathbb{R}^k, \mathbb{R}^d$) will influence the answer.

The article is structured as follows. In the next section the case of one-dimensional distributions is considered. Here the basic results are well-known. They are reviewed briefly together with interesting conclusions for our problem. After that the case of random vectors with possibly different dimensions k and d is considered followed by a conclusion.

In both cases we will give some formulations, which we find interesting in connection with the investigation of the stochastic finite element method (SFEM), or more generally with the use of polynomial chaos expansions in the solution of random equations. In order to give a correct basis for these statements we will formulate an abstract version of the Cameron-Martin theorem together with used concepts here. This material is taken from Chapter 2 in [8].

Considering the case of representation of random vectors we will use results about probability in metric spaces. Therefore related basic concepts and results are also reviewed briefly afterwards.

1.1 The Cameron-Martin Theorem

Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space which is sufficiently rich such that it is possible to define on it nontrivial normally distributed random variables $\gamma \sim \mathcal{N}(0, \sigma^2)$ with mean value 0 and variance $\sigma^2 > 0$ (otherwise one speaks only about the degenerate random variable which takes on the value 0 with probability 1). A *Gaussian linear space* is a linear subspace of the space $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbf{P})$, consisting of centred (i.e., with mean value 0) Gaussian random variables. A *Gaussian Hilbert space* is a closed linear subspace of the space $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbf{P})$ consisting of centred Gaussian random variables.

For a given Gaussian linear or Hilbert space $\mathcal{H}$ and a number $n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ one considers the following linear subspaces of $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbf{P})$:

$$\mathcal{P}_n(\mathcal{H}) := \{\ p(\xi_1, \dots, \xi_m); p(\cdot) \text{ is a polynomial of degree } \leq n,\ \xi_i \in \mathcal{H},\ i = 1, \dots, m,\ m \in \mathbb{N}_0\}$$

and the corresponding closure (in the space $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbf{P})$)

$$\overline{\mathcal{P}}_n(\mathcal{H}) := \mathbf{clos}\, \mathcal{P}_n(\mathcal{H}).$$

Note that the number of arguments of the polynomials $p(\cdot)$ is arbitrary, also the random variables can be chosen arbitrarily. One can show that these spaces are different for different values of n, so they form a strongly increasing sequence of subspaces $(\overline{\mathcal{P}}_n(\mathcal{H}), n \in \mathbb{N}_0)$ in $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbf{P})$. Taking orthogonal complements one defines for $n \in \mathbb{N}$

$$\mathcal{H}^{:n:} := \overline{\mathcal{P}}_n(\mathcal{H}) \ominus \overline{\mathcal{P}}_{n-1}(\mathcal{H}) = \overline{\mathcal{P}}_n(\mathcal{H}) \cap \overline{\mathcal{P}}_{n-1}(\mathcal{H})^{\perp}$$

so that it holds

$$\overline{\mathcal{P}}_n(\mathcal{H}) = \bigoplus_{k=0}^{n} \mathcal{H}^{:k:}$$

with $\mathcal{H}^{:0:} = \mathcal{P}_0(\mathcal{H}) = \overline{\mathcal{P}}_0(\mathcal{H})$, which is isomorphic to $\mathbb{R}$, and the definition

$$\bigoplus_{n=0}^{\infty} \mathcal{H}^{:n:} = \mathbf{clos} \bigcup_{n=0}^{\infty} \overline{\mathcal{P}}_n(\mathcal{H}).$$

Now we can state the Cameron-Martin theorem.

Theorem 1 *With the above definitions the spaces* $\mathcal{H}^{:n:}, n \geq 0$ *are pairwise orthogonal closed linear subspaces of* $\mathrm{L}^2(\Omega, \mathcal{A}, \mathbf{P})$ *and it holds*

$$\bigoplus_{n=0}^{\infty} \mathcal{H}^{:n:} = \mathrm{L}^2(\Omega, \sigma(\mathcal{H}), \mathbf{P}).$$

Hence in the case of $\mathcal{A} = \sigma(\mathcal{H})$ *the space* $\mathrm{L}^2 = \mathrm{L}^2(\Omega, \mathcal{A}, \mathbf{P})$ *admits the orthogonal expansion*

$$\mathrm{L}^2(\Omega, \mathcal{A}, \mathbf{P}) = \bigoplus_{n=0}^{\infty} \mathcal{H}^{:n:}.$$

Remark

Elements of the spaces L^2 and hence also $\mathcal{H}$ are equivalence classes of random variables. Therefore $\sigma(\mathcal{H})$ means that all of the equivalent functions have to be measurable, i.e., this σ-algebra is generated by one representative from each equivalence class and the events with probability 0. This remark applies also to similar notations below.

Denoting for a random variable $\xi \in \mathrm{L}^2(\Omega, \sigma(\mathcal{H}), \mathbf{P})$ and $n \in \mathbb{N}_0$ the orthogonal projection on $\mathcal{H}^{:n:}$ with $\xi^{:n:} = \mathbf{proj}_{\mathcal{H}^{:n:}} \xi$ we get the so called Wiener-Hermite polynomial chaos expansion for the random variable

$$\xi = \sum_{k=0}^{\infty} \xi^{:k:}$$

whereby the series converges in the mean square sense. Therefore the random variable ξ can be approximated in mean square sense by partial sums

$$\xi \approx {}^{n}\xi := \sum_{k=0}^{n} \xi^{:k:}.$$

In the simplest case, when the Gaussian Hilbert space is one-dimensional, i.e., $\mathcal{H} = \{c\gamma : c \in \mathbb{R}\}$ with a standard Gaussian random variable γ, the linear spaces $\mathcal{H}^{:n:}$ are also one-dimensional and they are spanned by the Hermite polynomial of degree n with respect to the basis variable, i.e., $\mathcal{H}^{:n:} = \{c\mathrm{He}_n(\gamma) : c \in \mathbb{R}\}$, with $\mathrm{He}_n(x) := (-1)^n \mathrm{e}^{\frac{x^2}{2}} \frac{d^n}{dx^n}\left(\mathrm{e}^{-\frac{x^2}{2}}\right)$ for $n \in \mathbb{N}_0,\ x \in \mathbb{R}$.

In view of the broad use of these facts for SFEM and other numerical methods for random equations some remarks are appropriate.

Remarks

1. The condition $\mathcal{A} = \sigma(\mathcal{H})$ is a necessary one. This follows from measurability properties, see e.g. the Doob-Dynkin theorem. A simple example where this condition is not fulfilled and the conclusion of the theorem is not valid can be given as follows. Take as probability space $\Omega = \mathbb{R}$, $\mathcal{A} = \sigma(\{0\},\{1\})$, $\mathbf{P}(\{1\}) = p$, $\mathbf{P}(\{0\}) = 1 - p$, $0 < p < 1$. Then the only possible nonempty Gaussian Hilbert space for this probability space is trivial, i.e., it consists only of the equivalence class of a.s. constant 0 random variables, $\mathcal{H} = \{\xi_0\}$, with random variable $\xi_0(\cdot)$ with $\xi_0(0) = \xi_0(1) = 0, \xi_0(\omega) = x_0$ for $\omega \notin \{0,1\}$ and with $x_0 \in \mathbb{R}$. The corresponding generated σ-algebra $\sigma(\mathcal{H}) = \sigma(\xi_0)$ consists only of events with probability 0 or 1, it holds $\sigma(\mathcal{H}) = \{\emptyset, \{0,1\}, \mathbb{R}\backslash\{0,1\}, \mathbb{R}\}$. Nevertheless on the given probability space there exist non-degenerate random variables with finite second order moments. For example the random variable ξ with $\xi(0) = 0$, $\xi(1) = 1$ and $\xi(\omega) = 2$ otherwise follows a Bernoulli distribution with parameter p. Completion of this probability space does not change the situation.

2. On the other hand random variables with finite second order moments and arbitrary discrete, singularly continuous or absolutely continuous as well as mixed type of distribution in $\mathrm{L}^2(\Omega, \sigma(\mathcal{H}), \mathbf{P})$ can be approximated in the mean square sense by polynomials in Gaussian random variables, the partial sums of the Wiener-Hermite polynomial chaos expansion.

3. This approximation assures the approximation of corresponding first and second order moments. Because from the mean square convergence it follows the convergence in probability and also the convergence in distribution one can also conclude that in this way an approximation of distribution functions and e.g. quantiles can be achieved. Of course there are also other characteristics which have to be approximated in applications, so e.g. probability densities or higher order moments (see e.g. [5, 6]), also other types of convergence concepts can be of interest.

Similar expansion results are valid for basic random variables with other distributions than a Gaussian one. The corresponding expansion are then called generalized polynomial chaos expansions (see e.g. [14]). Further properties of them will be investigated in a forthcoming paper.

1.2 Random elements in separable metric spaces

Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space. If $(\mathcal{X}, \mathcal{S}_\mathcal{X})$ is another measurable space, then a random element X in $\mathcal{X}$ is a measurable mapping from $(\Omega, \mathcal{A}, \mathbf{P})$ into $(\mathcal{X}, \mathcal{S}_\mathcal{X})$, i.e., $X : \Omega \to \mathcal{X}$ with

$$X^{-1}(B) := \{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A} \quad \forall B \in \mathcal{S}_\mathcal{X}.$$

Other names are "$\mathcal{X}$-valued random element", "generalized random variable in $\mathcal{X}$" or simply "random variable in $\mathcal{X}$".

In the case of $\mathcal{X} = \mathbb{R}$, this gives the definition of a real valued random variable and in case of $\mathcal{X} = \mathbb{R}^d$ $(d \in \mathbb{N})$ the definition of a finite dimensional random vector.

With every random element $X : \Omega \to \mathcal{X}$ a probability measure P_X on $(\mathcal{X}, \mathcal{S}_\mathcal{X})$ is related, the distribution of the random element. It is defined by

$$P_X(B) := \mathbf{P}(X \in B) := \mathbf{P}\left(\{\omega \in \Omega : X(\omega) \in B\}\right) \quad (B \in \mathcal{S}_\mathcal{X}).$$

A random element X with values in $\mathcal{X}$ is called a simple random element, if the range is a finite nonempty set in $\mathcal{X}$, i.e., there exists a finite partition of the probability space $\Omega = \bigcup_{k=1}^{N} \Omega_k$ with measurable sets $\Omega_k \in \mathcal{A}, k = 1, \ldots, N$ $(N \in \mathbb{N})$ and elements $x_k, k = 1, \ldots, N$ in $\mathcal{X}$ (they can be assumed to be pairwise distinct), such that $X(\omega) = x_k$ for $\omega \in \Omega_k$. The corresponding probabilities are $\mathbf{P}(\Omega_k) = p_k$, so that it holds $p_k \geq 0, k = 1, \ldots, N$ and $\sum_{k=1}^{N} p_k = 1$. The distribution of a simple random element is a discrete probability measure on $(\mathcal{X}, \mathcal{S}_\mathcal{X})$, it can be written as

$$\mathbf{P}_X = \sum_{k=1}^{N} p_k \delta_{x_k}$$

with the help of Dirac measures δ_{x_k} defined by $\delta_{x_k}(B) = 1$ if $x_k \in B$ and $\delta_{x_k}(B) = 0$ otherwise $(B \in \mathcal{S}_\mathcal{X}, k = 1, \ldots, N)$.

Usually the space $\mathcal{X}$ has a richer structure, often it is a metric space with metric function $\varrho : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ or even a Banach or Hilbert space. Then one can usually work with the measurable space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, where $\mathcal{B}(\mathcal{X})$ is the Borel-σ-algebra, i.e., the σ-algebra generated by the open subsets of $\mathcal{X}$. For such spaces it holds the following basic approximation result.

Theorem 2 *Let $(\mathcal{X}, \varrho)$ be a separable metric space, $\mathcal{B}(\mathcal{X})$ the Borel-σ-algebra on $\mathcal{X}$ and X a $(\mathcal{X}, \mathcal{B}(\mathcal{X}))-$valued random element. Then there exists a sequence $(X_n; n \in \mathbb{N})$ of $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$-valued simple random elements, which converges almost surely to X, i.e.,*

$$\mathbf{P}\left(\omega : \lim_{n\to\infty} \varrho(X_n(\omega), X(\omega)) = 0\right) = 1.$$

A sequence $(X_n; n \in \mathbb{N})$ of $(\mathcal{X}, \varrho)$-valued random elements converges in distribution to the random element X, if for all bounded continuous functions $f : \mathcal{X} \to \mathbb{R}$ it holds

$$\lim_{n\to\infty} \mathbf{E}\{f(X_n)\} = \mathbf{E}\{f(X)\}.$$

If a sequence $(X_n; n \in \mathbb{N})$ of $(\mathcal{X}, \varrho)$-valued random elements converges almost surely to the random element X, then it converges also in distribution. The opposite is not true in general, but the following theorem of Skorokhod is valid. Thereby we denote the usual Lebesgue measure on subsets of $\mathbb{R}$ with **Leb**.

Theorem 3 *Let $(\mathcal{X}, \varrho)$ be a complete separable metric space, and let $(X_n; n \in \mathbb{N})$ be a sequence of $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$-valued random elements, which converges in distribution to the random* *element X. Then there exists a sequence $(\widetilde{X}_n; n \in \mathbb{N})$ of $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$-valued random elements, which is defined on the probability space $((0,1), \mathcal{B}(0,1), \mathbf{Leb})$ and converges almost surely to a random element $\widetilde{X}$, such that it holds for the distributions $P_{X_n} = P_{\widetilde{X}_n}, n \in \mathbb{N}$, and $P_X = P_{\widetilde{X}}$.*

This means, that the convergence in distribution can be realized as almost sure convergence with suitable random elements, such that the distributions of the corresponding members of the sequences coincide. (The distribution of the sequences as a whole do not coincide in general.) Here we use only the fact, that the random elements can be defined on the special probability space $((0,1), \mathcal{B}(0,1), \mathbf{Leb})$. A proof of this result can be found e.g. in [7]. An immediate conclusion is the following:

Corollary 4 *Let $(\mathcal{X}, \varrho)$ be a complete separable metric space, and let X be a $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$-valued random element. Then one can find a random element $\widetilde{X}$ with the same distribution as X, i.e., $P_X = P_{\widetilde{X}}$, which is defined on the probability space $((0,1), \mathcal{B}(0,1), \mathbf{Leb})$.*

These and further results and examples concerning probability or random elements in metric or Banach spaces can be found e.g. in [1, 2, 4, 12].

2 Representation of random variables

Let ξ be a real valued random variable, defined on a probability space. It can be characterized by (and will be given mostly only through) its distribution function

$$F_\xi(x) = \mathbf{P}(\xi \le x), \qquad x \in \mathbb{R}.$$

The basic properties are

1. $\lim_{x \to -\infty} F_\xi(x) = 0, \quad \lim_{x \to \infty} F_\xi(x) = 1;$
2. $F_\xi(x)$ is nondecreasing on $\mathbb{R}$
3. $F_\xi(x)$ is continuous from the right and has limits from the left on $\mathbb{R}$.

These properties characterize the class of distribution functions of real random variables, i.e., each real valued function on $\mathbb{R}$ with these properties is a distribution function of a random variable (see e.g. [13], Chapter II, §3).

One can define the generalized inverse function

$$F_\xi^-(u) := \sup\{x \in \mathbb{R} : F_\xi(x) \le u\} \qquad (u \in (0,1)).$$

In many cases or allowing $+\infty$ and $-\infty$ as possible values the domain of definition can be extended (by continuity) to the closed interval $[0,1]$. If the distribution function $F_\xi(\cdot)$ is strongly increasing on $\mathbb{R}$ or on the interval $[a,b]$ with $F_\xi(a) = 0$ and $F_\xi(b) = 1$, then the inverse function $F_\xi^{-1}(\cdot)$ exists in the usual sense and it coincides with the generalized inverse.

Now it holds (see e.g. [11], Sect.5-2 or [3], Sect.3.1 for the following results)

Theorem 5 *Let η be a random variable with uniform distribution on the interval* $(0,1)$ *and* $F(\cdot)$ *an arbitrary distribution function of a real random variable with generalized inverse function* $F^{-}(\cdot)$. *Then the random variable* $\xi := F^{-}(\eta)$ *has the distribution function* $F(\cdot)$, *i.e., it holds*

$$F_{\xi}(x) = \mathbf{P}(\xi \leq x) = \mathbf{P}(F^{-}(\eta) \leq x) = F(x).$$

This basic result is used for example in Monte Carlo simulations in order to generate realizations of random variables with prescribed distribution functions.

A corresponding result in the opposite direction is the following.

Theorem 6 *Let* $F_{\xi}(\cdot)$ *be a continuous distribution function of a real valued random variable* ξ. *Then the random variable* $\eta := F_{\xi}(\xi)$ *is uniformly distributed on the interval* $(0,1)$, *i.e., it holds*

$$F_{\eta}(x) = \mathbf{P}(\eta \leq x) = \mathbf{P}(F_{\xi}(\xi) \leq x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & 1 \leq x. \end{cases}$$

Here the assumption, that $F_{\xi}(\cdot)$ is continuous on $\mathbb{R}$ is necessary, for a discontinuous distribution function $F_{\xi}(\cdot)$ the random variable $\eta = F_{\xi}(\xi)$ is not uniformly distributed on the interval $(0,1)$.

Example
Let ξ be a random variable with exponential distribution with parameter 1, i.e.,

$$F_{\xi}(x) = \mathbf{P}(\xi \leq x) = \begin{cases} 0 & x < 0 \\ 1 - \mathrm{e}^{-x} & x \geq 0. \end{cases}$$

Then the (generalized) inverse function is

$$F_{\xi}^{-1}(u) = -\ln(1-u) \qquad 0 \leq u < 1.$$

Hence the random variable $\eta := F_{\xi}(\xi) = 1 - \mathrm{e}^{-\xi}$ is uniformly distributed on the interval $(0,1)$.
Otherwise, for a uniformly on the interval $(0,1)$ distributed random variable $\widetilde{\eta}$ the random variable $\widetilde{\xi} := F_{\xi}^{-1}(\eta) = -\ln(1-\widetilde{\eta})$ follows an exponential distribution with parameter 1 (here one can also take the variable $\hat{\xi} = -\ln(\widetilde{\eta})$ because also $1-\widetilde{\eta}$ is uniformly distributed on the interval $(0,1)$).

From Theorem 6 we conclude the following.

Theorem 7 *Let* $(\Omega, \mathcal{A}, \mathbf{P})$ *be a probability space, on which a real-valued random variable* ξ *with continuous distribution function* $F_{\xi}(\cdot)$ *may be defined. Furthermore let* $F(\cdot)$ *be an arbitrary distribution function of a real-valued random variable. Then one can define on the probability space* $(\Omega, \mathcal{A}, \mathbf{P})$ *a random variable with this distribution function* $F(\cdot)$.

Proof
The distribution function of the random variable $\zeta_F := F^{-}(F_{\xi}(\xi))$ is the given function $F(\cdot)$ by the preceding theorems. □

Corollary 8 *If the Gaussian Hilbert space $\mathcal{H}$ is nontrivial ($\mathcal{H} \neq \{0\}$), then on the probability space $(\Omega, \sigma(\mathcal{H}), \mathbf{P})$ there exist random variables with arbitrary distribution functions.*

Hence for a nontrivial Gaussian Hilbert space random variables with arbitrary distributions with existing second order moments can be approximated through a Wiener-Hermite polynomial chaos expansion (given the right measurability properties).

In connections with applications of Theorem 7 it can also be remarked, that there is usually not only one random variable on the given probability space $(\Omega, \mathcal{A}, \mathbf{P})$ with the prescribed distribution function $F(\cdot)$. So e.g. for a random variable η with uniform distribution on the interval $(0,1)$ also the random variables $\eta_k = g_k(\eta), k = 1,2,3$ are uniformly distributed on the interval $(0,1)$ and defined on the same underlying probability space, if the deterministic functions are

$$g_1(x) = 1 - x \qquad x \in (0,1)$$
$$g_2(x) = \begin{cases} 2x & 0 < x \leq \frac{1}{2} \\ 2x - 1 & \frac{1}{2} < x < 1 \end{cases}$$
$$g_3(x) = \begin{cases} x & 0 < x \leq \frac{1}{2} \\ \frac{3}{2} - x & \frac{1}{2} < x < 1. \end{cases}$$

Then e.g. all the random variables $\zeta_1 := F^-(g_1(F_\xi(\xi)))$, $\zeta_2 := F^-(g_2(F_\xi(\xi)))$ and $\zeta_3 := F^-(g_3(F_\xi(\xi)))$ possess the same given distribution function $F(\cdot)$.

This rises new questions. Consider for example the case of Wiener-Hermite polynomial chaos expansions for a random variable $\xi \in \mathrm{L}^2(\Omega, \sigma(\gamma), \mathbf{P})$ with a standard Gaussian random variable γ. Then there are many other standard Gaussian random variables $\gamma_i = g_i(\gamma), i \in \mathbf{I}$, with deterministic bijective measurable functions $g_i : \mathbb{R} \to \mathbb{R}$, $\mathbf{I}$ is a nonempty index set. In general the corresponding Wiener-Hermite polynomial chaos expansions of ξ with respect to Hermite polynomials in γ_i will be different for different indices $i \in \mathbf{I}$. So the expansion of the random variable γ with respect to Hermite polynomials in γ will be exact for $n = 1$, whereas the expansion with respect to Hermite polynomials in $\gamma_i \neq \gamma$ will not. Thus we can use expansions which converge slower or faster and the question of the best choice can be posed. We will not consider this question here.

3 Representation of random vectors

Now we investigate the case of transformations of random vectors, especially the question of the influence of different dimensions. In the Gaussian case different dimensions of the random vectors are characterized by different numbers of stochastically independent components. Hence there is a relationship to the possible numbers of independent random variables in the result of such a deterministic measurable transformation. So one can pose for example the following question. Is it possible to generate a random vector $\boldsymbol{\xi}$ in $\mathbb{R}^d$ with prescribed distribution $P_{\boldsymbol{\xi}}$ by a deterministic measurable transformation from a standard Gaussian vector $\boldsymbol{\gamma}$ in $\mathbb{R}^k$ but not from a standard Gaussian vector $\widetilde{\boldsymbol{\gamma}}$ in $\mathbb{R}^{\tilde{k}}$ with $\tilde{k} < k$?

A first question in this direction which one can ask is the following. Is it possible to find to a given random variable η a transformed random variable $\xi = g(\eta)$, such that η and ξ

are stochastically independent, i.e., is it possible to transform a random variable η into a random vector $(\eta, g(\eta))$ with independent components? The answer on this question can be given easily.

Theorem 9 *Let η be a random variable on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ which is not almost surely constant. Then any random variable $\xi = g(\eta)$ with a deterministic measurable function $g : \mathbb{R} \to \mathbb{R}$, which is stochastically independent of η, is almost surely constant.*

Proof
For the generated σ-algebras it holds $\sigma(\xi) \subseteq \sigma(\eta)$, furthermore the σ-algebras $\sigma(\xi)$ and $\sigma(\eta)$ are independent if the random variables ξ and η are stochastically independent. Hence for any event $A \in \sigma(\xi)$ we have

$$\mathbf{P}(A) = \mathbf{P}(A \cap A) = \mathbf{P}(A)\mathbf{P}(A).$$

This is possible only if $\mathbf{P}(A) = 0$ or $\mathbf{P}(A) = 1$ which in turn means that the random variable ξ is almost surely constant. □

So in this way it is not possible to generate new nontrivial independent random variables to a given one. But also another situation is of interest. Assume we are given a random variable η which is non-degenerate, i.e., not almost surely constant. Can we find several measurable functions $g_i : \mathbb{R} \to \mathbb{R}$, such that the non-degenerate random variables $\xi_i = g_i(\eta)$ are mutual stochastically independent? We will here restrict ourselves to random variables η with continuous distribution function (this is the case we are mainly interested in) and give at first an abstract result, from which the affirmative answer to the question above follows.

Theorem 10 *Let η be a real random variable with continuous distribution function on an underlying probability space $(\Omega, \mathcal{A}, \mathbf{P})$ and let $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ be a measurable space, consisting of a complete separable metric space $\mathcal{X}$ endowed with its σ-algebra of Borel subsets $\mathcal{B}(\mathcal{X})$. Then for any probability distribution $\widetilde{\mathbf{P}}$ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ there exists a measurable mapping $g : \mathbb{R} \to \mathcal{X}$ such that the random element $\widetilde{X} := g(\eta)$ has the distribution $\widetilde{\mathbf{P}}$.*

Proof
Let $F_\eta(\cdot)$ be the continuous distribution function of the random variable η. By Theorem 6 the random variable $\xi := F_\eta(\eta)$ is uniformly distributed on the interval $(0, 1)$.

Now we apply Corollary 4 and find on the probability space $((0, 1), \mathcal{B}(0, 1), \mathbf{Leb})$ a random element $\hat{X}$ with the given distribution $P_{\hat{X}} = \widetilde{\mathbf{P}}$. The searched measurable function $g : \mathbb{R} \to \mathbb{R}$ can then be defined as the composite mapping $g(x) := \hat{X}(F_\eta(x))$, the random element is $\widetilde{X} = \hat{X}(F_\eta(\eta))$. □

Corollary 11 *Let η be a real random variable with continuous distribution function on an underlying probability space $(\Omega, \mathcal{A}, \mathbf{P})$. Then for an arbitrary sequence $(F_i; i \in \mathbb{N})$ of distribution functions there exist functions $g_i : \mathbb{R} \to \mathbb{R}, i \in \mathbb{N}$, such that the random variables $\xi_i = g_i(\eta)$ are stochastically independent.*

Proof
Apply the previous theorem to the complete separable metric space $\mathcal{X} = \mathbb{R}^{\mathbb{N}}$ and the product

probability distribution $\widetilde{\mathbf{P}} = \prod_{i=1}^{\infty} \mathbf{P}_{F_i}$ on it. Here $\mathbf{P}_{F_i}$ denotes the probability distribution on $\mathbb{R}$ with distribution function F_i. □

Theorem 12 *Let η be a real random variable with continuous distribution function on an underlying probability space $(\Omega, \mathcal{A}, \mathbf{P})$. Furthermore let $(g_i; i \in \mathbf{I})$ be a family of deterministic measurable functions $g_i : \mathbb{R} \to \mathbb{R}$, such that the non-degenerate random variables $\xi_i = g_i(\eta), i \in \mathbf{I}$, are stochastically independent. Then the index set $\mathbf{I}$ is at most denumerable.*

Proof

If the random variables $\xi_i = g_i(\eta), i \in \mathbf{I}$, are independent, then also the transformed bounded random variables $\widetilde{\xi}_i = \arctan(\xi_i) = \arctan(g_i(\eta)) = \widetilde{g}_i(\eta)$, $i \in \mathbf{I}$. For these bounded random variables all moments exist. We assume without loss of generality that these random variables are centred. The corresponding deterministic functions $\widetilde{g}_i(x) = \arctan\left(g_i(F_\eta^-(x))\right)$ are defined and measurable on $(0,1)$, they belong to the Hilbert space $\mathrm{L}^2((0,1), \mathcal{B}(0,1), \mathbf{Leb})$ and they are mutually orthogonal there. But the space $\mathrm{L}^2((0,1), \mathcal{B}(0,1), \mathbf{Leb})$ is separable, hence the number of such functions is at most countable. □

From Corollary 11 an interesting theoretical result for polynomial chaos expansions can be infered. We begin with an auxiliary result.

Theorem 13 *Let $\mathcal{H}$ be a separable Gaussian Hilbert space with basis $(\gamma_i; i \in \mathbf{I} \subseteq \mathbb{N})$. Then there exists a standard Gaussian random variable $\gamma \in \mathrm{L}^2(\Omega, \sigma(\mathcal{H}), \mathbf{P})$ such that the basis random variables can be generated from γ with deterministic measurable functions $g_i : \mathbb{R} \to \mathbb{R}$, i.e., $\gamma_i = g_i(\gamma), i \in \mathbf{I}$. Particularly it holds $\sigma(\mathcal{H}) = \sigma(\gamma)$.*

Proof

We can transform any standard Gaussian random variable $\widetilde{\gamma}$ into a uniformly on the interval $(0,1)$ distributed random variable via $\widetilde{\xi} = \Phi(\widetilde{\gamma})$ and vice versa. Thereby independent random variables are transformed into independent random variables. So we can deal with random variables $\xi_i = \Phi(\gamma_i), i \in \mathbf{I}$, with uniform distribution on the interval $(0,1)$. Furthermore we can assume, that all random variables are defined on the probability space $((0,1), \mathcal{B}(0,1), \mathbf{Leb})$.

Every number $\omega \in (0,1)$ can be expanded in the dual number system,

$$\omega = \sum_{n=1}^{\infty} \frac{b_n(\omega)}{2^n},$$

thereby the functions $b_n(\cdot)$ take on only the values 0 and 1 and this expansion is almost surely unique. This leads to corresponding expansions

$$\xi_i(\omega) = \sum_{n=1}^{\infty} \frac{b_n(\xi_i(\omega))}{2^n},$$

here the random variables $b_n(\xi_i(\cdot)), n, i \in \mathbb{N}$, are stochastically independent and Bernoulli distributed with parameter $p = 0.5$ because the random variables $\xi_i, i \in \mathbb{N}$, are uniformly distributed on $(0,1)$. Now we consider a bijection $k : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ and define the random variable

$$\xi(\omega) := \sum_{i,j=1}^{\infty} \frac{b_j(\xi_i(\omega))}{2^{k(i,j)}} = \sum_{k=1}^{\infty} \frac{c_k(\omega)}{2^k}$$

with $c_k(\omega) := c_{k(i,j)}(\omega) = b_j(\xi_i(\omega)), i, j \in \mathbb{N}$.

Then the random variables $c_k(\cdot), k \in \mathbb{N}$ are independent and Bernoulli distributed with parameter $p = 0.5$, they coincide almost surely with the expansion term of $\xi(\cdot)$ in the dual system, i.e., $c_k(\cdot) = b_k(\xi(\cdot))$ a.s., furthermore the random variable $\xi(\cdot)$ is uniformly distributed on the interval $(0, 1)$. Then the random variables $\xi_i(\cdot), i \in \mathbb{N}$, can be reconstructed from $\xi(\cdot)$ via the expansions

$$\xi_i(\omega) = \sum_{n=1}^{\infty} \frac{c_{k(i,n)}(\omega)}{2^n} = \sum_{n=1}^{\infty} \frac{b_{k(i,n)}(\xi(\omega))}{2^n}.$$

This leads in connection with the transformation from and into Gaussian random variables to the searched functions.

□

The reasoning in the proof above follows the probabilistic interpretation of the interval $(0, 1)$, as developed e.g. by Steinhaus (see e.g. [9], Chapter I or [13], Chapter II, §11).

From this and the abstract Martin-Cameron Theorem 1 we conclude

Corollary 14 *Let $\mathcal{H}$ be a separable Gaussian Hilbert space. Then there exists a standard Gaussian random variable $\gamma \in \mathrm{L}^2(\Omega, \sigma(\mathcal{H}), \mathbf{P})$ such that it holds $\sigma(\mathcal{H}) = \sigma(\gamma)$ and each random variable $\xi \in \mathrm{L}^2(\Omega, \sigma(\mathcal{H}), \mathbf{P})$ can be expanded in this space in an abstract Fourier series with respect to Hermite polynomials in γ, i.e.*

$$\xi = \sum_{n=0}^{\infty} c_n \frac{\mathrm{He}_n(\gamma)}{\sqrt{n!}} \qquad \text{with} \qquad c_n = \mathbf{E}\left\{\xi \frac{\mathrm{He}_n(\gamma)}{\sqrt{n!}}\right\}.$$

So, theoretically one can restrict oneself to polynomial chaos expansions with respect to *only one* standard Gaussian random variable. But it must be remarked, that the transformation from a vector of independent standard Gaussian random variables to one standard Gaussian random variable, which stands behind this fact, is a general measurable one and in general one cannot expect further smoothness or similar properties. It is expected that these series converge slower than corresponding series for smooth functions of Gaussian random variables. On the other hand the number of needed term in Wiener-Hermite chaos expansions rises very fast for several basic random variables. The possible effect of the reduction to only one random variable has to be investigated separately.

In connection with the results above the following can be stated.

Theorem 15 *Let $\mathcal{H}$ be a nontrivial Gaussian Hilbert space. Then there exists a standard Gaussian random variable $\gamma \in \mathrm{L}^2(\Omega, \sigma(\mathcal{H}), \mathbf{P})$ which do not belong to the Gaussian Hilbert space $\mathcal{H}$.*

Proof

There exist random vectors (ξ_1, ξ_2), for which the one-dimensional marginal distributions of ξ_1 and ξ_2 are Gaussian, but the vector is not a Gaussian random vector, i.e., there exist linear combinations $a_1\xi_1 + a_2\xi_2$ which are not normally distributed. An example is given e.g. in [13], Chapter II, §13. Random variables $\widetilde{\xi}_1$ and $\widetilde{\xi}_2$ such that the distributions of

$(\widetilde{\xi}_1, \widetilde{\xi}_2)$ and (ξ_1, ξ_2) coincide belong to $L^2(\Omega, \sigma(\mathcal{H}), \mathbf{P})$, but they cannot be elements of $\mathcal{H}$ because the linear combination $a_1\widetilde{\xi}_1 + a_2\widetilde{\xi}_2$ is not normally distributed. $\square$

This theorem means with other words, that in general a Gaussian Hilbert space cannot contain all Gaussian random variables on a given probability space.

4 Conclusion

It is shown in this article, that if one works with separable Hilbert spaces of random variables which contain random variables with continuous distribution functions, theoretically one Gaussian random variable can generate all other random variables (with the help of deterministic measurable functions) and polynomial chaos expansions are valid with respect to this single random variable. The possible use of this result for the approximate solution of random equations requires further investigations.

A further result states that allowing arbitrary measurable transformations random variables or vectors with arbitrary distributions can be generated from a random variable with continuous distribution function.

A forthcoming paper is addressed to the question of convergence of generalized polynomial chaos expansions and related problems.

References

[1] A.T. Bharucha-Reid. *Random Integral Equations.* Academic Press, New York, 1972.

[2] P. Billingsley. *Convergence of Probability Measures.* John Wiley & Sons, New York et al., 1968.

[3] K.L. Chung. *A Course in Probability Theory.* Academic Press, New York et al, 2001.

[4] G. da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, Cambridge, 1992.

[5] B.J. Debusschere, H.N. Najm, Ph.P. Pébay, O.M. Knio, R.G. Ghanem, and O.P. le Maître. Numerical challenges in the use of polynomial chaos representations for stochastic processes. *SIAM J. Sci. Comput.*, 26(2):698–719, 2004.

[6] R.V. Field Jr. and M. Grigoriu. On the accuracy of the polynomial chaos expansion. *Probabilistic Engineering Mechanics*, 19:65–80, 2004.

[7] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam, 1989.

[8] S. Janson. *Gaussian Hilbert Spaces.* Cambridge University Press, Cambridge, 1997.

[9] J.-P. Kahane. *Some Random Series of Functions.* D.C. Heath and Company, Lexington, 1968 (Russian Translation: Mir, Moscow 1973).

[10] O. Kallenberg. *Foundations of Modern Probability. Second edition.* Springer, Berlin et al., 2001.

[11] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* Mc Graw-Hill Inc., New York et al., 1991.

[12] K.R. Parthasarathy. *Probability Measures on Metric Spaces.* Academic Press, New York, 1967.

[13] A.N. Shiryaev. *Probability.* Springer, Berlin et al., 1996.

[14] D. Xiu and G.E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.