Lecture Notes in Mathematics

905

Claudia Prévôt
Michael Röckner

A Concise Course on Stochastic Partial Differential Equations

$$\mathrm{d}X_t = [\Delta\Psi(X_t) + \Phi(X_t)]\,\mathrm{d}t + B(X_t)\,\mathrm{d}W_t$$

Springer

Claudia Prévôt · Michael Röckner

A Concise Course on Stochastic Partial Differential Equations

Springer

Authors

Claudia Prévôt
Fakultät für Mathematik
Universität Bielefeld
Universitätsstr. 25
33615 Bielefeld
Germany
e-mail: cprevot@web.de

Michael Röckner
Fakultät für Mathematik
Universität Bielefeld
Universitätsstr. 25
33615 Bielefeld
Germany
e-mail: roeckner@math.uni-bielefeld.de

Departments of Mathematics
and Statistics
Purdue University
150 N. University St.
West Lafayette, IN 47907-2067
USA
e-mail: roeckner@math.purdue.edu

Library of Congress Control Number: 2007925694

Mathematics Subject Classification (2000): 35-XX, 60-XX

ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692
ISBN-10 3-540-70780-8 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-70780-6 Springer Berlin Heidelberg New York

DOI 10.1007/978-3-540-70781-3

Springer is a part of Springer Science+Business Media
springer.com

Typesetting by the authors and SPi using a Springer LaTeX macro package
Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper SPIN: 11982159 VA41/3100/SPi 5 4 3 2 1 0

Lecture Notes in Mathematics

Edited by J.-M. Morel, F. Takens and B. Teissier

Editorial Policy
for the publication of monographs

1. Lecture Notes aim to report new developments in all areas of mathematics and their applications – quickly, informally and at a high level. Mathematical texts analysing new developments in modelling and numerical simulation are welcome.

   Monograph manuscripts should be reasonably self-contained and rounded off. Thus they may, and often will, present not only results of the author but also related work by other people. They may be based on specialised lecture courses. Furthermore, the manuscripts should provide sufficient motivation, examples and applications. This clearly distinguishes Lecture Notes from journal articles or technical reports which normally are very concise. Articles intended for a journal but too long to be accepted by most journals, usually do not have this „lecture notes" character. For similar reasons it is unusual for doctoral theses to be accepted for the Lecture Notes series, though habilitation theses may be appropriate.

2. Manuscripts should be submitted (preferably in duplicate) either to Springer's mathematics editorial in Heidelberg, or to one of the series editors (with a copy to Springer). In general, manuscripts will be sent out to 2 external referees for evaluation. If a decision cannot yet be reached on the basis of the first 2 reports, further referees may be contacted: The author will be informed of this. A final decision to publish can be made only on the basis of the complete manuscript, however a refereeing process leading to a preliminary decision can be based on a pre-final or incomplete manuscript. The strict minimum amount of material that will be considered should include a detailed outline describing the planned contents of each chapter, a bibliography and several sample chapters.

   Authors should be aware that incomplete or insufficiently close to final manuscripts almost always result in longer refereeing times and nevertheless unclear referees' recommendations, making further refereeing of a final draft necessary.

   Authors should also be aware that parallel submission of their manuscript to another publisher while under consideration for LNM will in general lead to immediate rejection.

3. Manuscripts should in general be submitted in English. Final manuscripts should contain at least 100 pages of mathematical text and should always include

   - a table of contents;
   - an informative introduction, with adequate motivation and perhaps some historical remarks: it should be accessible to a reader not intimately familiar with the topic treated;
   - a subject index: as a rule this is genuinely helpful for the reader.

   For evaluation purposes, manuscripts may be submitted in print or electronic form (print form is still preferred by most referees), in the latter case preferably as pdf- or zipped ps-files. Lecture Notes volumes are, as a rule, printed digitally from the authors' files. To ensure best results, authors are asked to use the LaTeX2e style files available from Springer's web-server at:

   ftp://ftp.springer.de/pub/tex/latex/mathegl/mono/ (for monographs) and

   ftp://ftp.springer.de/pub/tex/latex/mathegl/mult/ (for summer schools/tutorials).

   Additional technical instructions, if necessary, are available on request from lnm@springer-sbm.com.

Continued on inside back-cover

Lecture Notes in Mathematics 1905

Contents

1. Motivation, Aims and Examples 1

2. Stochastic Integral in Hilbert Spaces 5
- 2.1. Infinite-dimensional Wiener processes . . . 5
- 2.2. Martingales in general Banach spaces . . . 17
- 2.3. The definition of the stochastic integral . . . 21
  - 2.3.1. Scheme of the construction of the stochastic integral . . 22
  - 2.3.2. The construction of the stochastic integral in detail . . . 22
- 2.4. Properties of the stochastic integral . . . 35
- 2.5. The stochastic integral for cylindrical Wiener processes . . . 39
  - 2.5.1. Cylindrical Wiener processes . . . 39
  - 2.5.2. The definition of the stochastic integral . . . 41

3. Stochastic Differential Equations in Finite Dimensions 43
- 3.1. Main result and a localization lemma . . . 43
- 3.2. Proof of existence and uniqueness . . . 49

4. A Class of Stochastic Differential Equations 55
- 4.1. Gelfand triples, conditions on the coefficients and examples . . 55
- 4.2. The main result and an Itô formula . . . 73
- 4.3. Markov property and invariant measures . . . 91

A. The Bochner Integral 105
- A.1. Definition of the Bochner integral . . . 105
- A.2. Properties of the Bochner integral . . . 107

B. Nuclear and Hilbert–Schmidt Operators 109

C. Pseudo Inverse of Linear Operators 115

D. Some Tools from Real Martingale Theory 119

E. Weak and Strong Solutions: Yamada–Watanabe Theorem 121
- E.1. The main result . . . 121

F. Strong, Mild and Weak Solutions 133

Bibliography **137**

Index **140**

Symbols **143**

1. Motivation, Aims and Examples

These lectures will concentrate on (nonlinear) stochastic partial differential equations (SPDEs) of evolutionary type. All kinds of dynamics with stochastic influence in nature or man-made complex systems can be modelled by such equations. As we shall see from the examples, at the end of this section the state spaces of their solutions are necessarily infinite dimensional such as spaces of (generalized) functions. In these notes the state spaces, denoted by E, will be mostly separable Hilbert spaces, sometimes separable Banach spaces.

There is also enormous research activity on SPDEs, where the state spaces are not linear, but rather spaces of measures (particle systems, dynamics in population genetics) or infinite-dimensional manifolds (path or loop spaces over Riemannian manifolds).

There are basically three approaches to analysing SPDEs: the "martingale (or martingale measure) approach" (cf. [Wal86]), the "semigroup (or mild solution) approach" (cf. [DPZ92], [DPZ96]) and the "variational approach" (cf. [Roz90]). There is an enormously rich literature on all three approaches which cannot be listed here. We refer instead to the above monographs.

The purpose of these notes is to give a concise introduction to the "variational approach", as self-contained as possible. This approach was initiated in pioneering work by Pardoux ([Par72],[Par75]) and further developed by N. Krylov and B. Rozowskii in [KR79] (see also [Roz90]) for continuous martingales as integrators in the noise term and later by I. Gyongy and N. Krylov in [GK81],[GK82],[Gyö82] for not necessarily continuous martingales.

These notes grew out of a two-semester graduate course given by the second-named author at Purdue University in 2005/2006. The material has been streamlined and could be covered in just one semester depending on the pre-knowledge of the attending students. Prerequisites would be an advanced course in probability theory, covering standard martingale theory, stochastic processes in $\mathbb{R}^d$ and maybe basic stochastic integration, though the latter is not formally required. Since graduate students in probability theory are usually not familiar with the theory of Hilbert spaces or basic linear operator theory, all required material from these areas is included in the notes, most of it in the appendices. For the same reason we minimize the general theory of martingales on Hilbert spaces, paying, however, the price that some proofs about stochastic integration on Hilbert space are a bit lengthy, since they have to be done "by bare hands".

In comparison with [Roz90] for simplicity we specialize to the case where the integrator in the noise term is just a cylindrical Wiener process. But everything is spelt out in a way so that it generalizes directly to continuous local martingales. In particular, integrands are always assumed to be predictable rather than just adapted and product measurable. The existence and uniqueness proof (cf. Subsection 4.2) is our personal version of the one in [KR79], [Roz90] and largely taken from [RRW06] presented there in a more general framework. The results on invariant measures (cf. Subsection 4.3) we could not find in the literature for the "variational approach". They are, however, quite straightforward modifications of those in the "semigroup approach" in [DPZ96]. The examples and applications in Subsection 4.1 in connection with the stochastic porous media equation are fairly recent and are modifications from results in [DPRLRW06] and [RRW06].

To keep these notes reasonably self-contained we also include a complete proof of the finite-dimensional case in Chapter 3, which is based on the very focussed and beautiful exposition in [Kry99], which uses the Euler approximation. Among other complementing topics the appendices contain a detailed account of the Yamada–Watanabe theorem on the relation between weak and strong solutions (cf. Appendix E).

The structure of these notes is, as we hope, obvious from the list of contents. We only would like to mention here, that a substantial part consists of a very detailed introduction to stochastic integration on Hilbert spaces (see Chapter 2), major parts of which (as well as Appendices A–C) are taken from the Diploma thesis of Claudia Prévôt and Katja Frieler. We would like to thank Katja Frieler at this point for her permission to do this. We also like to thank all coauthors of those joint papers which form another component for the basis of these notes. It was really a pleasure working with them in this exciting area of probability. We would also like to thank Matthias Stephan and Sven Wiesinger for the excellent typing job, as well as the participants of the graduate course at Purdue University for spotting many misprints and small mistakes.

Before starting with the main body of these notes we would like to give a few examples of SPDE that appear in fundamental applications. We do this in a very brief way, in particular, pointing out which of them can be analysed by the tools developed in this course. We refer to the above-mentioned literature for a more elaborate discussion of these and many more examples and their role in the applied sciences.

Example 1.0.1 (Stochastic quantization of the free Euclidean quantum field).

$$\mathrm{d}X_t = (\Delta - m^2)X_t \,\mathrm{d}t + \mathrm{d}W_t$$

on $E \subset \mathcal{S}'(\mathbb{R}^d)$.

- $m \in [0, \infty)$ denotes "mass",
- $(W_t)_{t \geqslant 0}$ is a cylindrical Brownian motion on $L^2(\mathbb{R}^d) \subset E$ (the inclusion is a Hilbert–Schmidt embedding).

Example 1.0.2 (Stochastic reaction diffusion equations).

$$\mathrm{d}X_t = [\Delta X_t - X_t^3]\,\mathrm{d}t + \sqrt{Q}\,\mathrm{d}W_t$$

on $E := L^p(\mathbb{R}^d)$.

- Q is a trace class operator on $L^2(\mathbb{R}^d)$, can also depend on X_t (then Q becomes $Q(X_t)$),
- $(W_t)_{t \geqslant 0}$ is a cylindrical Brownian motion on $L^2(\mathbb{R}^d)$.

Example 1.0.3 (Stochastic Burgers equation).

$$\mathrm{d}X_t = \Delta X_t - X_t \frac{\mathrm{d}}{\mathrm{d}\xi} X_t + \sqrt{Q}\,\mathrm{d}W_t$$

on $E := L^2\big([0,1]\big)$.

- $\xi \in [0, 1]$,
- Q as above,
- $(W_t)_{t \geqslant 0}$ is a cylindrical Brownian motion on $L^2\big([0,1]\big)$.

Example 1.0.4 (Stochastic Navier–Stokes equation).

$$\mathrm{d}X_t = \big[\nu \Delta_s X_t - \langle X_t, \nabla \rangle X_t\big]\,\mathrm{d}t + \sqrt{Q}\,\mathrm{d}W_t$$

on $E := \big\{x \in L^2(\Lambda \to \mathbb{R}^2,\ \mathrm{d}x) \;\big|\; \operatorname{div} x = 0\big\}$, $\Lambda \subset \mathbb{R}^d$, $d = 2, 3$, $\partial\Lambda$ smooth.

- ν denotes viscosity,
- Δ_s denotes the Stokes Laplacian,
- Q as above,
- $(W_t)_{t \geqslant 0}$ is a cylindrical Brownian motion on $L^2(\Lambda \to \mathbb{R}^d)$,
- div is taken in the sense of distributions.

Example 1.0.5 (Stochastic porous media equation).

$$\mathrm{d}X_t = \big[\Delta\Psi(X_t) + \Phi(X_t)\big]\,\mathrm{d}t + B(X_t)\,\mathrm{d}W_t$$

on $H :=$ dual of $H_0^1(\Lambda)$ (:= Sobolev space of order 1 in $L^2(\Lambda)$ with Dirichlet boundary conditions).

- Λ as above,
- $\Psi, \Phi : \mathbb{R} \to \mathbb{R}$ "monotone",
- $B(x) : H \to H$ Hilbert–Schmidt operator, $\forall\, x \in H$.

The general form of these equations with state spaces consisting of functions $\xi \mapsto x(\xi)$, where ξ is a spatial variable, e.g. from a subset of $\mathbb{R}^d$, looks as follows:

$$\begin{aligned} \mathrm{d}X_t(\xi) = A\Big(t, X_t(\xi), D_\xi X_t(\xi), D_\xi^2\big(X_t(\xi)\big)\Big)\,\mathrm{d}t \\ + B\Big(t, X_t(\xi), D_\xi X_t(\xi), D_\xi^2\big(X_t(\xi)\big)\Big)\,\mathrm{d}W_t\,. \end{aligned}$$

Here D_ξ and D_ξ^2 mean first and second total derivatives, respectively. The stochastic term can be considered as a "perturbation by noise". So, clearly one motivation for studying SPDEs is to get information about the corresponding (unperturbed) deterministic PDE by letting the noise go to zero (e.g. replace B by $\varepsilon \cdot B$ and let $\varepsilon \to 0$) or to understand the different features occurring if one adds the noise term.

If we drop the stochastic term in these equations we get a deterministic PDE of "evolutionary type". Roughly speaking this means we have that the time derivative of the desired solution (on the left) is equal to a non–linear functional of its spatial derivatives (on the right).

Among others (see Subsection 4.1, in particular the cases, where Δ is replaced by the p-Laplacian) the approach presented in these notes will cover Examples 1.0.2 in case $d = 3$ or 4. (cf. Remark 4.1.10,2. and also [RRW06] without restrictions on the dimension) and 1.0.5 (cf. Example 4.1.11). For Example 1.0.1 we refer to [AR91] and for Examples 1.0.3 and 1.0.4 e.g. to [DPZ92], [DPZ96].

2. The Stochastic Integral in General Hilbert Spaces (w.r.t. Brownian Motion)

This chapter is a slight modification of Chap. 1 in [FK01].

We fix two separable Hilbert spaces $\big(U, \langle\ ,\ \rangle_U\big)$ and $\big(H, \langle\ ,\ \rangle\big)$. The first part of this chapter is devoted to the construction of the stochastic Itô integral

$$\int_0^t \Phi(s)\, \mathrm{d}W(s), \quad t \in [0,T],$$

where $W(t)$, $t \in [0,T]$, is a Wiener process on U and Φ is a process with values that are linear but not necessarily bounded operators from U to H.

For that we first will have to introduce the notion of the standard Wiener process in infinite dimensions. Then there will be a short section about martingales in general Hilbert spaces. These two concepts are important for the construction of the stochastic integral which will be explained in the following section.

In the second part of this chapter we will present the Itô formula and the stochastic Fubini theorem and establish basic properties of the stochastic integral, including the Burkholder–Davis–Gundy inequality.

Finally, we will describe how to transmit the definition of the stochastic integral to the case that $W(t)$, $t \in [0,T]$, is a cylindrical Wiener process. For simplicity we assume that U and H are real Hilbert spaces.

2.1. Infinite-dimensional Wiener processes

For a topological space X we denote its Borel σ-algebra by $\mathcal{B}(X)$.

Definition 2.1.1. A probability measure μ on $\big(U, \mathcal{B}(U)\big)$ is called *Gaussian* if for all $v \in U$ the bounded linear mapping

$$v' : U \to \mathbb{R}$$

defined by

$$u \mapsto \langle u, v\rangle_U, \quad u \in U,$$

has a Gaussian law, i.e. for all $v \in U$ there exist $m := m(v) \in \mathbb{R}$ and $\sigma := \sigma(v) \in [0, \infty[$ such that, if $\sigma(v) > 0$,

$$\left(\mu \circ (v')^{-1}\right)(A) = \mu(v' \in A) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_A e^{-\frac{(x-m)^2}{2\sigma^2}} \,\mathrm{d}x \quad \text{for all } A \in \mathcal{B}(\mathbb{R}),$$

and, if $\sigma(v) = 0$,

$$\mu \circ (v')^{-1} = \delta_{m(v)}.$$

Theorem 2.1.2. *A measure μ on $(U, \mathcal{B}(U))$ is Gaussian if and only if*

$$\hat{\mu}(u) := \int_U e^{i\langle u,v\rangle_U} \,\mu(\mathrm{d}v) = e^{i\langle m,u\rangle_U - \frac{1}{2}\langle Qu,u\rangle_U}, \quad u \in U,$$

where $m \in U$ and $Q \in L(U)$ is nonnegative, symmetric, with finite trace (see Definition B.0.3; here $L(U)$ denotes the set of all bounded linear operators on U).

In this case μ will be denoted by $N(m, Q)$ where m is called mean and Q is called covariance (operator). The measure μ is uniquely determined by m and Q.

Furthermore, for all $h, g \in U$

$$\int \langle x, h\rangle_U \,\mu(\mathrm{d}x) = \langle m, h\rangle_U,$$

$$\int \big(\langle x, h\rangle_U - \langle m, h\rangle_U\big)\big(\langle x, g\rangle_U - \langle m, g\rangle_U\big) \,\mu(\mathrm{d}x) = \langle Qh, g\rangle_U,$$

$$\int \|x - m\|_U^2 \,\mu(\mathrm{d}x) = \operatorname{tr} Q.$$

Proof. (cf. [DPZ92]) Obviously, a probability measure with this Fourier transform is Gaussian. Now let us conversely assume that μ is Gaussian. We need the following:

Lemma 2.1.3. *Let ν be a probability measure on $(U, \mathcal{B}(U))$. Let $k \in \mathbb{N}$ be such that*

$$\int_U \left|\langle z, x\rangle_U\right|^k \nu(\mathrm{d}x) < \infty \quad \forall\, z \in U.$$

Then there exists a constant $C = C(k, \nu) > 0$ such that for all $h_1, \ldots, h_k \in U$

$$\int_U \left|\langle h_1, x\rangle_U \cdots \langle h_k, x\rangle_U\right| \nu(\mathrm{d}x) \leqslant C\, \|h_1\|_U \cdots \|h_k\|_U.$$

In particular, the symmetric k-linear form

$$U^k \ni (h_1, \ldots, h_k) \mapsto \int \langle h_1, x\rangle_U \cdots \langle h_k, x\rangle_U \,\nu(\mathrm{d}x) \in \mathbb{R}$$

is continuous.

Proof. For $n \in \mathbb{N}$ define

$$U_n := \left\{ z \in U \;\middle|\; \int_U \left|\langle z, x\rangle_U\right|^k \nu(\mathrm{d}x) \leqslant n \right\}.$$

By assumption

$$U = \bigcup_{n=1}^{\infty} U_n.$$

Since U is a complete metric space, by the Baire category theorem, there exists $n_0 \in \mathbb{N}$ such that U_{n_0} has non-empty interior, so there exists a ball (with centre z_0 and radius r_0) $B(z_0, r_0) \subset U_{n_0}$. Hence

$$\int_U \left|\langle z_0 + y, x\rangle_U\right|^k \nu(\mathrm{d}x) \leqslant n_0 \quad \forall\, y \in B(0, r_0),$$

therefore for all $y \in B(0, r_0)$

$$\begin{aligned}
\int_U \left|\langle y, x\rangle_U\right|^k \nu(\mathrm{d}x) &= \int_U \left|\langle z_0 + y, x\rangle_U - \langle z_0, x\rangle_U\right|^k \nu(\mathrm{d}x) \\
&\leqslant 2^{k-1} \int_U \left|\langle z_0 + y, x\rangle_U\right|^k \nu(\mathrm{d}x) + 2^{k-1} \int_U \left|\langle z_0, x\rangle_U\right|^k \nu(\mathrm{d}x) \\
&\leqslant 2^k n_0.
\end{aligned}$$

Applying this for $y := r_0 z$, $z \in U$ with $|z|_U = 1$, we obtain

$$\int_U \left|\langle z, x\rangle_U\right|^k \nu(\mathrm{d}x) \leqslant 2^k n_0 r_0^{-k}.$$

Hence, if $h_1, \ldots, h_k \in U \setminus \{0\}$, then by the generalized Hölder inequality

$$\begin{aligned}
&\int_U \left| \left\langle \frac{h_1}{|h_1|_U}, x \right\rangle_U \cdots \left\langle \frac{h_k}{|h_k|_U}, x \right\rangle_U \right| \nu(\mathrm{d}x) \\
&\leqslant \left(\int_U \left| \left\langle \frac{h_1}{|h_1|_U}, x \right\rangle_U \right|^k \nu(\mathrm{d}x) \right)^{1/k} \cdots \left(\int_U \left| \left\langle \frac{h_k}{|h_k|_U}, x \right\rangle_U \right|^k \nu(\mathrm{d}x) \right)^{1/k} \\
&\leqslant 2^k n_0 r_0^{-k},
\end{aligned}$$

and the assertion follows. □

Applying Lemma 2.1.3 for $k = 1$ and $\nu := \mu$ we obtain that

$$U \ni h \;\mapsto\; \int \langle h, x\rangle_U \,\mu(\mathrm{d}x) \;\in \mathbb{R}$$

is a continuous linear map, hence there exists $m \in U$ such that

$$\int_U \langle x, h\rangle_U \,\mu(\mathrm{d}x) = \langle m, h\rangle \quad \forall\, h \in H.$$

Applying Lemma 2.1.3 for $k = 2$ and $\nu := \mu$ we obtain that

$$U^2 \ni (h_1, h_2) \mapsto \int \langle x, h_1\rangle_U \langle x, h_2\rangle_U \; \mu(\mathrm{d}x) - \langle m, h_1\rangle_U \langle m, h_2\rangle_U$$

is a continuous symmetric bilinear map, hence there exists a symmetric $Q \in L(U)$ such that this map is equal to

$$U^2 \ni (h_1, h_2) \mapsto \langle Qh_1, h_2\rangle_U.$$

Since for all $h \in U$

$$\langle Qh, h\rangle_U = \int \langle x, h\rangle_U^2 \; \mu(\mathrm{d}x) - \Big(\int \langle x, h\rangle_U \; \mu(\mathrm{d}x)\Big)^2 \geqslant 0,$$

Q is positive definite. It remains to prove that Q is trace class (i.e.

$$\operatorname{tr} Q := \sum_{i=1}^{\infty} \langle Qe_i, e_i\rangle_U < \infty$$

for one (hence every) orthonormal basis $\{e_i \mid i \in \mathbb{N}\}$ of U, cf. Appendix B). We may assume without loss of generality that μ has mean zero, i.e. $m = 0$ $(\in U)$, since the image measure of μ under the translation $U \ni x \mapsto x - m$ is again Gaussian with mean zero and the same covariance Q. Then we have for all $h \in U$ and all $c \in (0, \infty)$

$$\begin{aligned} 1 - e^{-\frac{1}{2}\langle Qh,h\rangle_U} &= \int_U \big(1 - \cos\langle h, x\rangle_U\big) \; \mu(\mathrm{d}x) \\ &\leqslant \int_{\{|x|_U \leqslant c\}} \big(1 - \cos\langle h, x\rangle_U\big) \; \mu(\mathrm{d}x) + 2\mu\big(\{x \in U \mid |x|_U > c\}\big) \\ &\leqslant \frac{1}{2} \int_{\{|x|_U \leqslant c\}} \big|\langle h, x\rangle_U\big|^2 \; \mu(\mathrm{d}x) + 2\mu\big(\{x \in U \mid |x|_U > c\}\big) \end{aligned} \tag{2.1.1}$$

(since $1 - \cos x \leqslant \frac{1}{2}x^2$). Defining the positive definite symmetric linear operator Q_c on U by

$$\langle Q_c h_1, h_2\rangle_U := \int_{\{|x|_U \leqslant c\}} \langle h_1, x\rangle_U \cdot \langle h_2, x\rangle_U \; \mu(\mathrm{d}x), \quad h_1, h_2 \in U,$$

we even have that Q_c is trace class because for every orthonormal basis $\{e_k \mid k \in \mathbb{N}\}$ of U we have (by monotone convergence)

$$\begin{aligned} \sum_{k=1}^{\infty} \langle Q_c e_k, e_k\rangle_U &= \int_{\{|x|_U \leqslant c\}} \sum_{k=1}^{\infty} \langle e_k, x\rangle_U^2 \; \mu(\mathrm{d}x) = \int_{\{|x|_U \leqslant c\}} |x|_U^2 \; \mu(\mathrm{d}x) \\ &\leqslant c^2 < \infty. \end{aligned}$$

Claim: There exists $c_0 \in (0, \infty)$ (large enough) so that $Q \leqslant 2\log 4\, Q_{c_0}$ (meaning that $\langle Qh, h\rangle_U \leqslant 2\log 4 \langle Q_{c_0}h, h\rangle_U$ for all $h \in U$).

To prove the claim let c_0 be so big that $\mu\big(\{x \in U \mid |x|_U > c_0\}\big) \leqslant \frac{1}{8}$. Let $h \in U$ such that $\langle Q_{c_0}h, h\rangle_U \leqslant 1$. Then (2.1.1) implies

$$1 - e^{-\frac{1}{2}\langle Qh,h\rangle_U} \leqslant \frac{1}{2} + \frac{1}{4} = \frac{3}{4},$$

hence $4 \geqslant e^{\frac{1}{2}\langle Qh,h\rangle_U}$, so $\langle Qh, h\rangle_U \leqslant 2\log 4$. If $h \in U$ is arbitrary, but $\langle Q_{c_0}h, h\rangle_U \neq 0$, then we apply what we have just proved to $h/\langle Q_{c_0}h, h\rangle_U^{\frac{1}{2}}$ and the claim follows for such h. If, however, $\langle Q_{c_0}h, h\rangle = 0$, then for all $n \in \mathbb{N}$, $\langle Q_{c_0}nh, nh\rangle_U = 0 \leqslant 1$, hence by the above $\langle Qh, h\rangle_U \leqslant n^{-2} 2\log 4$. Therefore, $\langle Q_{c_0}h, h\rangle_U = 0$ and the claim is proved, also for such h.

Since Q_{c_0} has finite trace, so has Q by the claim and the theorem is proved, since the uniqueness part follows from the fact that the Fourier transform is one-to-one. □

The following result is then obvious.

Proposition 2.1.4. *Let X be a U-valued Gaussian random variable on a probability space $(\Omega, \mathcal{F}, P)$, i.e. there exist $m \in U$ and $Q \in L(U)$ nonnegative, symmetric, with finite trace such that $P \circ X^{-1} = N(m, Q)$.*

Then $\langle X, u\rangle_U$ is normally distributed for all $u \in U$ and the following statements hold:

- $E\big(\langle X, u\rangle_U\big) = \langle m, u\rangle_U$ *for all* $u \in U$,
- $E\big(\langle X - m, u\rangle_U \cdot \langle X - m, v\rangle_U\big) = \langle Qu, v\rangle_U$ *for all* $u, v \in U$,
- $E\big(\|X - m\|_U^2\big) = \operatorname{tr} Q$.

The following proposition will lead to a representation of a U-valued Gaussian random variable in terms of real-valued Gaussian random variables.

Proposition 2.1.5. *If $Q \in L(U)$ is nonnegative, symmetric, with finite trace then there exists an orthonormal basis e_k, $k \in \mathbb{N}$, of U such that*

$$Qe_k = \lambda_k e_k, \quad \lambda_k \geqslant 0,\ k \in \mathbb{N},$$

and 0 is the only accumulation point of the sequence $(\lambda_k)_{k\in\mathbb{N}}$.

Proof. See [RS72, Theorem VI.21; Theorem VI.16 (Hilbert–Schmidt theorem)]. □

Proposition 2.1.6 (Representation of a Gaussian random variable). *Let $m \in U$ and $Q \in L(U)$ be nonnegative, symmetric, with $\operatorname{tr} Q < \infty$. In addition, we assume that e_k, $k \in \mathbb{N}$, is an orthonormal basis of U consisting of eigenvectors of Q with corresponding eigenvalues λ_k, $k \in \mathbb{N}$, as in Proposition 2.1.5, numbered in decreasing order.*

Then a U-valued random variable X on a probability space $(\Omega, \mathcal{F}, P)$ is Gaussian with $P \circ X^{-1} = N(m, Q)$ if and only if

$$X = \sum_{k \in \mathbb{N}} \sqrt{\lambda_k} \beta_k e_k + m \quad \textit{(as objects in } L^2(\Omega, \mathcal{F}, P; U)\textit{)},$$

where β_k, $k \in \mathbb{N}$, are independent real-valued random variables with $P \circ {\beta_k}^{-1} = N(0,1)$ for all $k \in \mathbb{N}$ with $\lambda_k > 0$. The series converges in $L^2(\Omega, \mathcal{F}, P; U)$.

Proof.

1. Let X be a Gaussian random variable with mean m and covariance Q. Below we set $\langle \, , \, \rangle := \langle \, , \, \rangle_U$.

   Then $X = \sum_{k \in \mathbb{N}} \langle X, e_k \rangle e_k$ in U where $\langle X, e_k \rangle$ is normally distributed with mean $\langle m, e_k \rangle$ and variance λ_k, $k \in \mathbb{N}$, by Proposition 2.1.4. If we now define

   $$\beta_k : \begin{cases} = \frac{\langle X, e_k \rangle - \langle m, e_k \rangle}{\sqrt{\lambda_k}} & \text{if } k \in \mathbb{N} \text{ with } \lambda_k > 0 \\ \equiv 0 \ \in \mathbb{R} & \text{else,} \end{cases}$$

   then we get that $P \circ \beta_k^{-1} = N(0,1)$ and $X = \sum_{k \in \mathbb{N}} \sqrt{\lambda_k} \beta_k e_k + m$. To prove the independence of β_k, $k \in \mathbb{N}$, we take an arbitrary $n \in \mathbb{N}$ and $a_k \in \mathbb{R}$, $1 \leqslant k \leqslant n$, and obtain that for $c := -\sum_{k=1,\, \lambda_k \neq 0}^{n} \frac{a_k}{\sqrt{\lambda_k}} \langle m, e_k \rangle$

   $$\sum_{k=1}^{n} a_k \beta_k = \sum_{\substack{k=1, \\ \lambda_k \neq 0}}^{n} \frac{a_k}{\sqrt{\lambda_k}} \langle X, e_k \rangle + c = \left\langle X, \sum_{\substack{k=1, \\ \lambda_k \neq 0}}^{n} \frac{a_k}{\sqrt{\lambda_k}} e_k \right\rangle + c$$

   which is normally distributed since X is a Gaussian random variable. Therefore we have that β_k, $k \in \mathbb{N}$, form a Gaussian family. Hence, to get the independence, we only have to check that the covariance of β_i and β_j, $i, j \in \mathbb{N}$, $i \neq j$, with $\lambda_i \neq 0 \neq \lambda_j$, is equal to zero. But this is clear since

   $$\begin{aligned} E(\beta_i \beta_j) &= \frac{1}{\sqrt{\lambda_i \lambda_j}} E\big(\langle X - m, e_i \rangle \langle X - m, e_j \rangle\big) = \frac{1}{\sqrt{\lambda_i \lambda_j}} \langle Q e_i, e_j \rangle \\ &= \frac{\lambda_i}{\sqrt{\lambda_i \lambda_j}} \langle e_i, e_j \rangle = 0 \end{aligned}$$

   for $i \neq j$.

   Besides, the series $\sum_{k=1}^{n} \sqrt{\lambda_k} \beta_k e_k$, $n \in \mathbb{N}$, converges in $L^2(\Omega, \mathcal{F}, P; U)$ since the space is complete and

   $$E\Big(\Big\| \sum_{k=m}^{n} \sqrt{\lambda_k} \beta_k e_k \Big\|^2\Big) = \sum_{k=m}^{n} \lambda_k E\big(|\beta_k|^2\big) = \sum_{k=m}^{n} \lambda_k.$$

   Since $\sum_{k \in \mathbb{N}} \lambda_k = \operatorname{tr} Q < \infty$ this expression becomes arbitrarily small for m and n large enough.