Helsinki University of Technology Institute of Mathematics Research Reports

Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja

Espoo 2005 A486

A MATHEMATICAL MODEL FOR ELECTRICAL IMPEDANCE PROCESS TOMOGRAPHY

Hanna Katriina Pikkarainen

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Helsinki University of Technology Institute of Mathematics Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2005 A486

A MATHEMATICAL MODEL FOR ELECTRICAL IMPEDANCE PROCESS TOMOGRAPHY

Hanna Katriina Pikkarainen

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Department of Engineering Physics and Mathematics for public examination and debate in Auditorium E at Helsinki University of Technology (Espoo, Finland) on 13th of May, 2005, at 12 o'clock noon.

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics

Hanna Katriina Pikkarainen: *A Mathematical Model for Electrical Impedance Process Tomography*; Helsinki University of Technology Institute of Mathematics Research Reports A486 (2005).

Abstract: We consider the process tomography problem of the following kind: based on electromagnetic measurements on the surface of a pipe, describe the concentration distribution of a given substance in a fluid moving in the pipeline. We view the problem as a state estimation problem. The concentration distribution is treated as a stochastic process satisfying a stochastic differential equation. This is referred to as the state evolution equation. The measurements are described in terms of an observation equation containing the measurement noise. The time evolution is modelled by a stochastic convection-diffusion equation. The measurement situation is represented by the most realistic model for electrical impedance tomography, the complete electrode model. In this thesis, we give the mathematical formulation of the state evolution and observation equations and then we derive the discrete infinite dimensional state estimation system. Since our motive is to monitor the flow in the pipeline in real time, we are dealing with a filtering problem in which the estimator is based on the current history of the measurement process. For computational reasons we present a discretized state estimation system where the discretization error is taken into account. The discretized filtering problem is solved by the Bayesian filtering method.

AMS subject classifications: 62M20 (Primary); 93E10, 60H15, 35J25 (Secondary)

Keywords: statistical inversion theory, nonstationary inverse problem, state estimation, Bayesian filtering, process tomography, electrical impedance tomography, complete electrode model

hanna.pikkarainen@tkk.fi

ISBN 951-22-7651-8
ISSN 0784-3143

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, 02015 TKK, Finland
email: math@hut.fi http://www.math.hut.fi/

Acknowledgements

This work has been carried out at the Institute of Mathematics at Helsinki University of Technology during the years 2000–2005. I am thankful for Professor Olavi Nevanlinna, the chairman of the institute, for arranging an excellent working environment.

I express my gratitude to Professor Erkki Somersalo, the supervisor and instructor of this thesis, for proposing me the subject of my dissertation. I have learnt a lot from different fields of mathematics.

I consider myself privileged for having Professors Luis Alvarez and Yaroslav Kurylev as the pre-examiners of my thesis. I am grateful to them for reading the manuscript with great accuracy and making invaluable comments. Professor Alvarez pointed out that the Feynman–Kac Formula is a useful tool for solving partial differential equations. Professor Kurylev's intense interest in my thesis has been extremely encouraging. I also thank Professor Markku Lehtinen for being my opponent.

Several people have helped me during this project. I owe my thanks to each and everyone of them. Professor Esko Valkeila has always found time for my questions about stochastics. He introduced me to the book of Da Prato and Zabczyk which is the main reference of Chapter 4. Mr Lasse Leskelä has been my tutor in stochastics. Professor Matti Lassas suggested how to extend the results concerning the complete electrode model (CEM) in bounded domains to unbounded domains. Doctors Ville Havu and Jarmo Malinen verified details of the results concerning CEM in unbounded domains and gave many useful comments and suggestions. Mr Kenrick Bingham has always had time to discuss with me about issues that have troubled me at least in mathematics, computers and the English language. Doctor Ville Havu and Docent Samuli Siltanen have looked through the manuscript and helped me in improving the language. With Professor Jari Kaipio and Mr Aku Seppänen I have had interesting discussions concerning the numerical implementation of the electrical impedance process tomography problem. I also want to thank all my colleagues at the Institute of Mathematics for an inspiring research atmosphere and all the people in the Finnish Inverse Problems Society for creating a great social atmosphere.

I would like to thank the Vilho, Yrjö and Kalle Väisälä Foundation, the Magnus Ehrnrooth Foundation, the Finnish Konkordia Fund, the Finnish Cultural Foundation and the Foundation of Technology for financial support. I also have been a participant of the Graduate School of Inverse Problems during this project.

Last but not least, I wish to thank my family and friends for being there for me. My warmest thanks belong to Tapio. He has loved me even though every now and then I have been totally unbearable.

Espoo, 19th April, 2005 Hanna Katriina Pikkarainen

Contents

Notations

$:=$, $=:$	definition
$\propto$	proportional to
$\sim$	identically distributed
$\mathbb{N}$	the set of natural numbers $\{1, 2, \ldots\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$
$\mathbb{R}$	the set of real numbers
$\mathbb{R}_+$	the set of positive real numbers
$\mathbb{C}$	the complex plane
$a \wedge b$	the minimum of real numbers a and b
$\operatorname{sgn} a$	the sign of a real number a
i	the imaginary unit
$\lvert z \rvert$	the absolute value of a complex number z
$\arg z$	the argument of a complex number z
$\mathrm{Re} z$	the real part of a complex number z
$\mathrm{Im} z$	the imaginary part of a complex number z
$\bar{z}$	the complex conjugate of a complex number z
$\mathbb{R}^n$	the n-dimensional Euclidean space
$\mathbb{R}^{n \times m}$	the set of $n \times m$ real matrices
A^T	the transpose of a matrix A
$\det A$	the determinant of a matrix A
$\bar{A}$	the closure of a set A
A^c	the complement of a set A
∂A	the boundary of a set A
χ_A	the characteristic (or indicator) function of a set A
$[a, b]$	the interval $\{x \in \mathbb{R} : a \leq x \leq b\}$
(a, b)	the interval $\{x \in \mathbb{R} : a < x < b\}$
$[a, b)$	the interval $\{x \in \mathbb{R} : a \leq x < b\}$
$(a, b]$	the interval $\{x \in \mathbb{R} : a < x \leq b\}$
$B(a, r)$	an open ball with radius r and centre at a
$\rho(\cdot, \cdot)$	a metric
$\lVert \cdot \rVert$	a norm
$(\cdot, \cdot)$	an inner product
$\mathrm{span}(A)$	the linear span of a set A
$\dim E$	the dimension of a space E
$B(E, F)$	the space of bounded linear operators from E to F

E'	the dual space of a norm space E
$\langle\cdot,\cdot\rangle$	a dual operation
$\mathcal{D}(A)$	the domain of an operator A
$\mathcal{R}(A)$	the range of an operator A
$\mathrm{Ker}(A)$	the null space of an operator A
A'	the Banach adjoint of an operator A
A^*	the Hilbert adjoint of an operator A
l^∞	the space of bounded sequences
$\mathrm{supp}\, f$	the support of a function f
$\frac{d}{dt}$	the derivative with respect to t
∂_t	the partial derivative with respect to t
∂_j	the j^{th} partial derivative
∇	$(\partial_1,\ldots,\partial_n)^T$
$C_0^2(D)$	the space of twice continuously differentiable functions in D with compact support
$C_0^\infty(D)$	the space of infinitely many times continuously differentiable functions in D with compact support
$UC(D)$	the space of uniformly continuous bounded functions in D
$UC^1(D)$	the space of uniformly continuously differentiable bounded functions in D
$C^{1,2}(\mathbb{R}\times\mathbb{R}^n)$	the space of functions in $\mathbb{R}\times\mathbb{R}^n$ which are continuously differentiable w.r.t. the first component and twice continuously differentiable w.r.t. the second component
m	the Lebesgue measure in $\mathbb{R}^n$
$L^p(D)$	the space of functions in D whose p^{th} power is integrable
$L^p_{\text{loc}}(D)$	the space of functions in D whose p^{th} power is locally integrable
$L^\infty(D)$	the space of essentially bounded functions in D
$W^{k,p}(D)$	the Sobolev space
$H^s(D)$	the Sobolev space where $p=2$
$\mathfrak{F}$	the Fourier transform
$\mathfrak{F}^{-1}$	the inverse Fourier transform
$\hat{f}$	the Fourier transform of a function f
$*$	the convolution
C^k	k times continuously differentiable
C^∞	infinitely many times continuously differentiable
$\frac{\partial}{\partial\nu}$	$\nu\cdot\nabla$
dS	a surface measure
$L^p(0,T;E)$	$L^p([0,T],\mathcal{B}([0,T]),m\vert_{[0,T]};E)$, see Appendix B
$L^p(D;\mathbb{R}^n)$	$L^p(D,\mathcal{B}(D),m\vert_D;\mathbb{R}^n)$, see Appendix B
$L^\infty(0,T;E)$	the space of essentially bounded functions from $[0,T]$ to E
$C([0,T];E)$	the space of continuous functions from $[0,T]$ to E
$C^\infty((0,\infty);E)$	the space of infinitely many times continuously differentiable functions from $(0,\infty)$ to E

Chapter 1

Introduction

In practical measurements of physical quantities we have directly observable quantities and others that cannot be observed. If some of the unobservable quantities are of our primary interest, we are dealing with an inverse problem. In that case, we need to discover how to compute the values of the quantities of primary interest from the observed values of the observable quantities, the measured data. The interdependence of the quantities in the measurement setting is described through mathematical models. For solving the inverse problem we have to be able to analyse mathematically the model of the measurement process. If we have some prior information about the quantities of primary interest, it is beneficial to use statistical approach to inverse problems. In *statistical inversion theory* it is assumed that all quantities included in the model are represented by random variables. The randomness describes our degree of knowledge concerning their realizations. Our information about their values is coded into their distributions. Therefore the randomness is due to the lack of information, not to the intrinsic randomness of the quantities in the measurement setting. The statistical inversion theory is based on the *Bayes formula*. The prior information of the quantities of primary interest is presented in the form of a *prior distribution*. The *likelihood function* is given by the model for the measurement process. The solution to the inverse problem is the *posterior distribution* of the random variables of interest after performing the measurements. By the Bayes formula the posterior distribution is proportional to the product of the prior distribution and likelihood function.

In several applications one encounters a situation in which measurements that constitute the data of an inverse problem are done in a nonstationary environment. More precisely, it may happen that the physical quantities that are in the focus of our primary interest are time dependent and the measured data depends on these quantities at different time instants. Inverse problems of this type are called *nonstationary inverse problems*. They are often viewed as a *state estimation problem*. Then the quantities in the measurement setting are treated as stochastic processes. Usually, the time evolution of the quantities of primary interest, the state of the system, is described by a stochastic differential equation referred to as the *state evolution equation*. The measurements are modeled by an *observation equation* containing the measurement noise. The solution to the state estimation problem is the conditional expectation of the quantities of primary interest with respect to the measured data. If our motive is, for instance, to have a real-time monitoring of the quantities of primary interest, we are dealing with a *filtering problem* in which the estimator is based on the current history of the measurement process.

Often in state estimation approach the time variable is assumed to be discrete and the space variable to be finite dimensional. This is convenient from the practical point of view. Observations are usually done at discrete time instants and the computation requires space discretization. Hence discrete state evolution and observation equations are needed. They may be derived from the continuous ones, especially if the state evolution and observation equations are linear. In many applications, it is assumed that the discretized version of the discrete infinite dimensional state estimation problem represents the reality. Nevertheless, discretization causes always an error, which should be included into the state estimation system. If we analyse the continuous infinite dimensional state evolution and observation equations, we may be able to present the distribution of the discretization error. The discretized filtering problem can be solved by the *Bayesian filtering method*. The discretized state evolution equation is used to find the prior distribution and the likelihood function is given by the discretized observation equation. The solution to the filtering problem is the posterior distribution given by the Bayes formula. As an example of nonstationary inverse problem we examine the electrical impedance process tomography problem.

1.1 Electrical Impedance Process Tomography

In this thesis we consider the process tomography problem of imaging the concentration distribution of a given substance in a fluid moving in a pipeline based on electromagnetic measurements on the surface of the pipe. In electrical impedance tomography (EIT) electric currents are applied to electrodes on the surface of an object and the resulting voltages are measured using the same electrodes (Figure 1.1). The conductivity distribution inside the object is reconstructed based on the voltage measurements. The relation between the conductivity and concentration depends on the process and is usually non-linear. At least for strong electrolytes and multiphase mixtures such relations are studied and discussed in the literature [7, 12]. In traditional EIT it is assumed that the object remains stationary during the measurement process. A complete set of measurements, also called a *frame*, consists of all possible linearly independent injected current patterns and the corresponding set of voltage measurements. In process tomography we cannot in general assume that the target remains unaltered during a full set of measurements. Thus conventional reconstruction methods [4, 5, 6, 46, 47, 49] cannot be used. The time evolution needs to be modeled properly. We view the problem as a state estimation problem. The concentration distribution is treated as a stochastic process that satisfies a stochastic differential equation referred to as the state evolution equation. The measurements are described in terms of an observation equation containing the measurement noise. Our goal is to have a real-time monitoring for the flow in a pipeline. For that reason the computational time has to be minimized. Therefore, we use a simple model, the convection–diffusion equation, for the flow. It allows numerical implementation using FEM techniques. Since we cannot be sure that other features such as turbulence of the flow do not appear, we use stochastic modelling. The measurement situation is represented by the most realistic model for EIT, the *complete electrode model*. The measurements are done in a part of the boundary

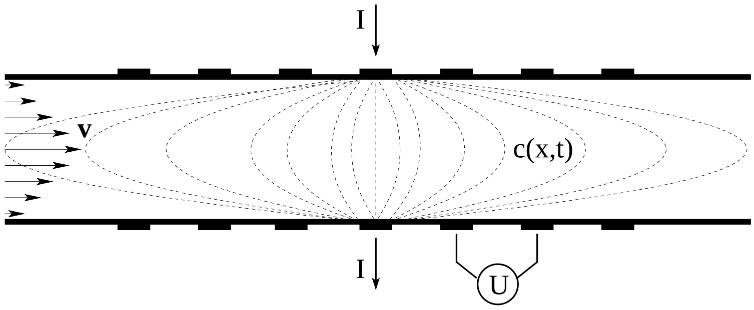

Figure 1.1: EIT in process tomography

of the pipe. We get enough information for an accurate computation only from a segment of the pipe. It would be natural to choose the domain of the model to be the segment of the pipe. If the domain is restricted to be a segment of the pipe, we have to use some boundary conditions in the input and output end of the segment. The choice of boundary conditions has an effect on the solution. The most commonly used boundary conditions do not represent the actual circumstances in the pipe. Therefore, we do not truncate the domain but instead assume that the pipe is infinitely long. With the assumption we derive the discrete infinite dimensional state estimation system.

This problem has been considered in the articles [43, 45, 41] and in the proceedings papers [40, 44, 42, 38]. Those articles and proceeding papers concentrate on the numerical implementation of the problem. An experimental evaluation is presented in the proceeding paper [39]. In those articles and proceeding papers the discretized state estimation system is assumed to model the real measurement process. The discretization error is omitted. In this thesis the main interest is in the mathematical formulation of the state evolution and observation equations and presenting a discretized state estimation system in which the discretization error is taken into account. Preliminary results have been published in proceedings papers [33, 34] written by the author.

1.2 Overview of this Thesis

The main purpose of this thesis is to present the state estimation system corresponding to electrical impedance process tomography and to perform discretization in such a manner that the discretization error is taken into account. We combine the theory of partial differential equations and stochastic analysis in infinite dimensions to solve the stochastic convection–diffusion equation. Since only few researchers interested in inverse problems are familiar with both branches of mathematics, we present well-known results concerning both fields. This thesis is rather self-contained even though it is assumed that the reader has a firm background in mathematics. The Lebesgue integration theory of scalar valued functions and stochastic analysis in $\mathbb{R}^n$ are supposed to be known. The reader should also be acquainted with the principles of functional analysis and theory of partial differential equations. Chapters 2–4 introduce the theory needed to solve the stochastic convection–diffusion equation. In Chapter 2 we discuss the concept of analytic semigroups and sectorial operators.

We use analytic semigroups generated by sectorial operators to solve initial value problems. Elliptic partial differential operators are studied in terms of sectoriality in Chapter 3. Chapter 4 considers stochastic analysis in infinite dimensional spaces. As a consequence we are able to solve linear stochastic differential equations. The existence and uniqueness of the solution to the complete electrode model in unbounded domains are proved in Chapter 5. Finally, in Chapter 6 we return to the electrical impedance process tomography problem. We present the continuous infinite dimensional state estimation system concerning the problem. A discretized state estimation system and the evolution and observation updating formulas of the Bayesian filtering are also introduced.

In this thesis there are four appendixes which contain theory needed in Chapters 2–4. In Appendix A basic properties of the resolvent set and operator used in Chapter 2 are introduced. The Bochner integration theory for Banach space valued functions is handled in Appendix B. The analytic semigroup generated by a sectorial operator is defined as an integral of an operator valued function along a curve in the complex plane. In Appendix C we apply the Bochner integration theory and show that the Cauchy integral theorem and formula are valid for holomorphic operator valued functions. The covariance operator of a Gaussian measure in a Hilbert space is a nuclear operator. Proper integrands of the stochastic integral with respect to a Hilbert space valued Wiener process are processes with values in the space of Hilbert-Schmidt operators. In Appendix D we present basic properties of Hilbert-Schmidt and nuclear operators.

In the beginning of each chapter we comment on the references used in that chapter and related literature. We do not refer to the literature concerning single results since the proofs of almost all theorems, propositions, lemmas etc. are included in the thesis. Often the proofs contain more details than those which can be found from the literature. In Chapter 4 there are few lemmas which we could not find from the literature in the required form. However, the proofs have only slight differences between those introduced in the literature. New results are presented in Chapters 5 and 6. All details in the proofs of the results concerning the complete electrode model in unbounded domains in Chapter 5 are made by the author. The main results of this thesis are presented in Section 6.3, which is entirely based on the author's individual work.

Chapter 2

Analytic Semigroups

In this chapter we introduce some properties of analytic semigroups generated by unbounded operators. We shall use analytic semigroups to find solutions to initial value problems. The theory of semigroups can be found among others in the books of Davies [8], Goldstein [15], Hille and Phillips [16], Lunardi [28], Pazy [31] and Tanabe [50, 51].

Let $(E, \|\cdot\|_E)$ be a Banach space. We denote by $B(E)$ the space of bounded linear operators from E to E equipped with the operator norm

$$\|A\|_{B(E)} := \sup\{\|Ax\|_E : x \in E,\ \|x\|_E \leq 1\}$$

for all $A \in B(E)$. An operator family $\{T(t)\}_{t\geq 0} \subset B(E)$ is called a *semigroup* if

(i) $T(t)T(s) = T(s+t)$ for all $s, t \geq 0$ and

(ii) $T(0) = I$.

The linear operator $A : \mathcal{D}(A) \to E$ defined by

$$\mathcal{D}(A) := \left\{ x \in E : \exists \lim_{t\to 0^+} \frac{T(t)x - x}{t} \right\},$$
$$Ax := \lim_{t\to 0^+} \frac{T(t)x - x}{t} \quad \text{if } x \in \mathcal{D}(A),$$

is called the *infinitesimal generator* of the semigroup $\{T(t)\}_{t\geq 0}$. A semigroup $\{T(t)\}_{t\geq 0}$ is said to be *strongly continuous* if for all $x \in E$ the function $t \mapsto T(t)x$ is continuous in the interval $[0, \infty)$. It is said to be *analytic* if the function $t \mapsto T(t)$ can be extended to be an analytic function from a sector

$$\{z \in \mathbb{C} : z \neq 0,\ |\arg z| \leq \beta\} \tag{2.1}$$

with some $\beta \in (0, \pi)$ to the space $B(E)$, i.e., for every disc $B(a, r)$ in Sector (2.1) there exists a series

$$\sum_{n=0}^{\infty} A_n (z-a)^n$$

where $A_n \in B(E)$ which converges in $B(E)$ to $T(z)$ for all $z \in B(a, r)$.

2.1 Sectorial Operators

This section is based on the beginning of Chapter 2 in the book of Lunardi [28]. Basic properties of the resolvent set $\rho(A)$ and resolvent operator $R(\lambda, A)$ which will be used in this section are presented in Appendix A.

Let $(E, \|\cdot\|_E)$ be a Banach space and $A : \mathcal{D}(A) \subseteq E \to E$ a linear operator with not necessarily dense domain $\mathcal{D}(A)$. If A is a bounded operator and $\mathcal{D}(A) = E$, we can define the operator e^{tA} by the series

$$e^{tA} := \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \tag{2.2}$$

for all $t > 0$. In addition, we denote $e^{0A} := I$. Then the operator family $\{e^{tA}\}_{t\geq 0}$ has properties

(i) $e^{tA} \in B(E)$ for all $t \geq 0$,

(ii) $e^{tA}e^{sA} = e^{(s+t)A}$ for all $s, t \geq 0$,

(iii) $e^{0A} = I$,

(iv) the function $z \mapsto e^{zA}$ is holomorphic in the whole complex plane and

(v) $\lim_{t\to 0^+} \frac{e^{tA}x - x}{t} = Ax$ for all $x \in E$.

Hence a bounded linear operator A defined in the whole E generates a strongly continuous analytic semigroup $\{e^{tA}\}_{t\geq 0}$.

If A is unbounded, Series (2.2) does not make sense. Under some specific assumptions an unbounded linear operator generates an analytic semigroup.

Definition 2.1. *A linear operator A is sectorial if there exist constants $\omega \in \mathbb{R}$, $\theta \in (\pi/2, \pi)$ and $M > 0$ such that*

(i) $S_{\omega,\theta} := \{\lambda \in \mathbb{C} : \lambda \neq \omega,\ |\arg(\lambda - \omega)| < \theta\} \subset \rho(A)$ *and*

(ii) $\|R(\lambda, A)\|_{B(E)} \leq \frac{M}{|\lambda - \omega|}$ *for all* $\lambda \in S_{\omega,\theta}$.

Let A be a sectorial operator with the constants ω, θ and M. Since the resolvent set of A is not empty, A is closed. According to the conditions (i) and (ii) in Definition 2.1 we can define a bounded linear operator $U(t)$ in the space E as a uniform Bochner integral

$$U(t) := \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} e^{t\lambda} R(\lambda, A)\, d\lambda \tag{2.3}$$

for all $t > 0$ where $r > 0$, $\eta \in (\pi/2, \theta)$ and $\gamma_{r,\eta}$ is the curve

$$\{\lambda \in \mathbb{C} : |\arg \lambda| = \eta,\ |\lambda| \geq r\} \cup \{\lambda \in \mathbb{C} : |\arg \lambda| \leq \eta,\ |\lambda| = r\}$$

oriented counterclockwise. In addition, we define

$$U(0)x := x \tag{2.4}$$

for all $x \in E$. By Proposition A.2 the function $\lambda \mapsto e^{t\lambda} R(\lambda, A)$ is holomorphic in the domain $S_{\omega,\theta}$. Since $\omega + \gamma_{r,\eta} \subset S_{\omega,\theta}$ for all r and η, the operator $U(t)$ does not depend on the choice of r and η. Details concerning the definition of the integral can be found from Appendix C.

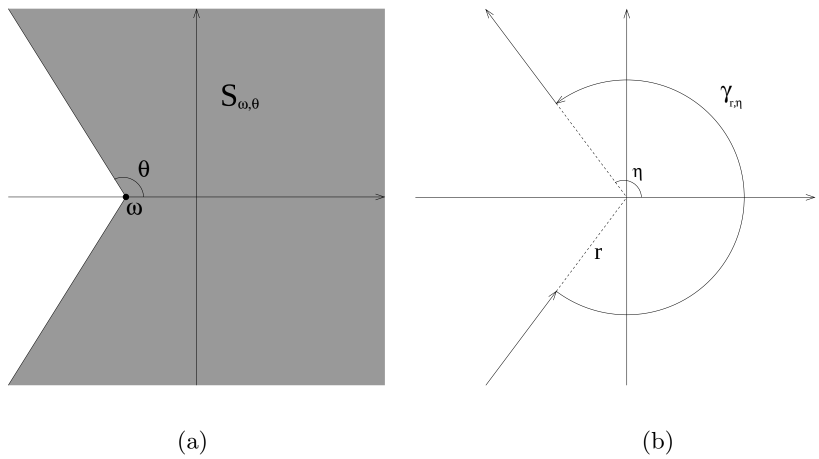

Figure 2.1: (a) The set $S_{\omega,\theta}$ and (b) the integration path $\gamma_{r,\eta}$

In the following proposition we state the main properties of the operator family $\{U(t)\}_{t\geq 0}$ defined by Formulas (2.3) and (2.4).

Proposition 2.2. *Let A be a sectorial operator with the constants ω, θ and M and the operator family $\{U(t)\}_{t\geq 0}$ defined by Formulas (2.3) and (2.4). Then the following statements are valid.*

(i) *$U(t)x \in \mathcal{D}(A^k)$ for all $k \in \mathbb{N}$, $t > 0$ and $x \in E$. If $x \in \mathcal{D}(A^k)$ for $k \in \mathbb{N}$, then $A^k U(t)x = U(t)A^k x$ for all $t \geq 0$.*

(ii) *$U(t)U(s) = U(s+t)$ for all $s, t \geq 0$.*

(iii) *There exist constants M_0, M_1, M_2, ... such that*

$$\|U(t)\|_{B(E)} \leq M_0 e^{\omega t} \quad \textit{and} \quad \|t^k (A - \omega I)^k U(t)\|_{B(E)} \leq M_k e^{\omega t}$$

for all $k \in \mathbb{N}$ and $t > 0$. In particular, for all $k \in \mathbb{N}$ there exists a constant $C_k > 0$ such that

$$\|t^k A^k U(t)\|_{B(E)} \leq C_k e^{(\omega+1)t}$$

for all $t > 0$.

(iv) *The function $t \mapsto U(t)$ belongs to the space $C^\infty((0,\infty); B(E))$ and*

$$\frac{d^k}{dt^k} U(t) = A^k U(t)$$

for all $t > 0$. In addition, the function $t \mapsto U(t)$ has an analytic extension in the sector

$$S := \left\{ z \in \mathbb{C} : z \neq 0,\ |\arg z| < \theta - \frac{\pi}{2} \right\}.$$

Proof. (i) Let $\lambda \in \rho(A)$. By the definition of the resolvent operator

$$AR(\lambda, A) = \lambda R(\lambda, A) - I$$

on E. Since A is sectorial and $R(\lambda, A)x \in \mathcal{D}(A)$ for all $\lambda \in \rho(A)$ and $x \in E$,

$$\begin{aligned} AU(t)x &= \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} e^{t\lambda} AR(\lambda, A)x \, d\lambda \\ &= \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda e^{t\lambda} R(\lambda, A)x \, d\lambda - \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} e^{t\lambda} x \, d\lambda \\ &= \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda e^{t\lambda} R(\lambda, A)x \, d\lambda \end{aligned}$$

for all $t > 0$ because the function $\lambda \mapsto e^{t\lambda}$ is holomorphic, $\eta > \pi/2$ and therefore

$$\int_{\omega+\gamma_{r,\eta}} e^{t\lambda} \, d\lambda = 0.$$

Using induction we are able to prove that

$$A^k U(t)x = \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda^k e^{t\lambda} R(\lambda, A)x \, d\lambda$$

for all $k \in \mathbb{N}$, $t > 0$ and $x \in E$. Since A is sectorial, the integral is well defined for all $k \in \mathbb{N}$. The calculation above proves the beginning of the induction. Let us assume that

$$A^k U(t)x = \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda^k e^{t\lambda} R(\lambda, A)x \, d\lambda$$

for all $k \leq n$, $t > 0$ and $x \in E$. Since A is sectorial and $R(\lambda, A)x \in \mathcal{D}(A)$ for all $\lambda \in \rho(A)$ and $x \in E$,

$$\begin{aligned} A^{n+1}U(t)x = AA^n U(t)x &= \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda^n e^{t\lambda} AR(\lambda, A)x \, d\lambda \\ &= \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda^{n+1} e^{t\lambda} R(\lambda, A)x \, d\lambda - \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda^n e^{t\lambda} x \, d\lambda \\ &= \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda^{n+1} e^{t\lambda} R(\lambda, A)x \, d\lambda \end{aligned}$$

for all $t > 0$ because the function $\lambda \mapsto \lambda^n e^{t\lambda}$ is holomorphic, $\eta > \pi/2$ and thus

$$\int_{\omega+\gamma_{r,\eta}} \lambda^n e^{t\lambda} \, d\lambda = 0.$$

Hence $U(t)x \in \mathcal{D}(A^k)$ and

$$A^k U(t)x = \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} \lambda^k e^{t\lambda} R(\lambda, A)x \, d\lambda$$

for all $k \in \mathbb{N}$, $t > 0$ and $x \in E$.

We show that $A^k U(t)x = U(t)A^k x$ for all $k \in \mathbb{N}$, $t > 0$ and $x \in \mathcal{D}(A^k)$ by using the induction. Let $t > 0$ and $x \in \mathcal{D}(A)$. Since A is sectorial and $AR(\lambda, A) = R(\lambda, A)A$ on $\mathcal{D}(A)$ for all $\lambda \in \rho(A)$,

$$AU(t)x = \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} e^{t\lambda} AR(\lambda, A)x \, d\lambda = \frac{1}{2\pi i} \int_{\omega+\gamma_{r,\eta}} e^{t\lambda} R(\lambda, A)Ax \, d\lambda = U(t)Ax.$$

We assume that $A^kU(t)x = U(t)A^kx$ for all $k \leq n$, $t > 0$ and $x \in \mathcal{D}(A^k)$. Then for all $t > 0$ and $x \in \mathcal{D}(A^{n+1})$

$$A^{n+1}U(t)x = AA^nU(t)x = AU(t)A^nx = U(t)A^{n+1}x.$$

Since $U(0) = I$, the statement is valid also for $t = 0$.

(ii) We introduce the operator

$$\begin{aligned} B : \mathcal{D}(A) &\to E \\ x &\mapsto Bx := Ax - \omega x. \end{aligned}$$

Then the resolvent set of B contains the sector $S_{0,\theta}$ and $R(\lambda, B) = R(\lambda + \omega, A)$ for all $\lambda \in S_{0,\theta}$. Thus for all $\lambda \in S_{0,\theta}$

$$\|R(\lambda, B)\|_{B(E)} = \|R(\lambda + \omega, A)\|_{B(E)} \leq \frac{M}{|\lambda|}.$$

Hence B is sectorial. By changing the variables $\kappa = \lambda + \omega$

$$\begin{aligned} U_B(t) &= \frac{1}{2\pi i} \int_{\gamma_{r,\eta}} e^{t\lambda} R(\lambda, B)\, d\lambda = \frac{1}{2\pi i} \int_{\gamma_{r,\eta}} e^{t\lambda} R(\lambda + \omega, A)\, d\lambda \\ &= \frac{1}{2\pi i} \int_{\omega + \gamma_{r,\eta}} e^{t(\kappa - \omega)} R(\kappa, A)\, d\kappa = e^{-\omega t} U_A(t) \end{aligned}$$

for all $t > 0$. Thus $U_B(t) = e^{-\omega t} U_A(t)$ for all $t \geq 0$.

Let $s, t > 0$ and $\pi/2 < \eta' < \eta < \theta$. Then

$$\begin{aligned} U_B(t)U_B(s) &= \left(\frac{1}{2\pi i}\right)^2 \int_{\gamma_{r,\eta}} e^{t\lambda} R(\lambda, B)\, d\lambda \int_{\gamma_{2r,\eta'}} e^{s\mu} R(\mu, B)\, d\mu \\ &= \left(\frac{1}{2\pi i}\right)^2 \int_{\gamma_{r,\eta} \times \gamma_{2r,\eta'}} e^{t\lambda + s\mu} R(\lambda, B) R(\mu, B)\, d\lambda d\mu \\ &= \left(\frac{1}{2\pi i}\right)^2 \int_{\gamma_{r,\eta} \times \gamma_{2r,\eta'}} e^{t\lambda + s\mu} \frac{R(\lambda, B) - R(\mu, B)}{\mu - \lambda}\, d\lambda d\mu \end{aligned}$$

by the resolvent identity. Since

$$\int_{\gamma_{2r,\eta'}} \frac{e^{s\mu}}{\mu - \lambda}\, d\mu = 2\pi i e^{s\lambda}$$

when $\lambda \in \gamma_{r,\eta}$, and

$$\int_{\gamma_{r,\eta}} \frac{e^{t\lambda}}{\mu - \lambda}\, d\lambda = 0$$

when $\mu \in \gamma_{2r,\eta'}$, the operator family $\{U_B(t)\}_{t>0}$ has the semigroup property for all $s, t > 0$, i.e.,

$$\begin{aligned} U_B(t)U_B(s) = &\left(\frac{1}{2\pi i}\right)^2 \int_{\gamma_{r,\eta}} e^{t\lambda} R(\lambda, B) \int_{\gamma_{2r,\eta'}} \frac{e^{s\mu}}{\mu - \lambda}\, d\mu\, d\lambda + \\ &- \left(\frac{1}{2\pi i}\right)^2 \int_{\gamma_{2r,\eta'}} e^{s\mu} R(\mu, B) \int_{\gamma_{r,\eta}} \frac{e^{t\lambda}}{\mu - \lambda}\, d\lambda\, d\mu \\ = &\frac{1}{2\pi i} \int_{\gamma_{r,\eta}} e^{(s+t)\lambda} R(\lambda, B)\, d\lambda = U_B(s + t). \end{aligned}$$

Hence

$$U_A(t)U_A(s) = e^{\omega(s+t)}U_B(t)U_B(s) = e^{\omega(s+t)}U_B(s+t) = U_A(s+t)$$

for all $s, t > 0$. Since $U_A(0) = I$, the operator family $\{U_A(t)\}_{t\geq 0}$ has the semigroup property, i.e., $U_A(t)U_A(s) = U_A(s+t)$ for all $s, t \geq 0$.

(iii) Let $t > 0$. By changing the variables $\xi = t\lambda$

$$U_B(t) = \frac{1}{2\pi i t}\int_{\gamma_{tr,\eta}} e^{\xi}R\left(\frac{\xi}{t}, B\right)\, d\xi = \frac{1}{2\pi i t}\int_{\gamma_{r,\eta}} e^{\xi}R\left(\frac{\xi}{t}, B\right)\, d\xi$$

since the integral does not depend on the choice of r and η. Thus for all $t > 0$

$$\begin{aligned}
&\|U_B(t)\|_{B(E)} \\
&\leq \frac{1}{2\pi t}\Bigg[\int_r^\infty e^{\rho\cos\eta}\left(\left\|R\left(\frac{\rho e^{-i\eta}}{t}, B\right)\right\|_{B(E)} + \left\|R\left(\frac{\rho e^{i\eta}}{t}, B\right)\right\|_{B(E)}\right)\, d\rho + \\
&\qquad + \int_{-\eta}^{\eta} r e^{r\cos\varphi}\left\|R\left(\frac{re^{i\varphi}}{t}, B\right)\right\|_{B(E)}\, d\varphi\Bigg] \\
&\leq \frac{M}{2\pi}\left[2\int_r^\infty \rho^{-1}e^{\rho\cos\eta}\, d\rho + \int_{-\eta}^{\eta} e^{r\cos\varphi}\, d\varphi\right] \leq M_0
\end{aligned}$$

since $\pi/2 < \eta < \theta < \pi$. Hence $\|U_A(t)\|_{B(E)} \leq M_0 e^{\omega t}$ for all $t > 0$.

Due to the statement (i) $U_B(t)x$ belongs to $\mathcal{D}(B) = \mathcal{D}(A)$ for all $t > 0$ and $x \in E$ and

$$BU_B(t) = \frac{1}{2\pi i}\int_{\gamma_{r,\eta}} \lambda e^{t\lambda}R(\lambda, B)\, d\lambda.$$

Let $t > 0$. By changing the variables $\xi = t\lambda$

$$BU_B(t) = \frac{1}{2\pi i t^2}\int_{\gamma_{tr,\eta}} \xi e^{\xi}R\left(\frac{\xi}{t}, B\right)\, d\xi = \frac{1}{2\pi i t^2}\int_{\gamma_{r,\eta}} \xi e^{\xi}R\left(\frac{\xi}{t}, B\right)\, d\xi.$$

Thus for all $t > 0$

$$\begin{aligned}
&\|BU_B(t)\|_{B(E)} \\
&\leq \frac{1}{2\pi t^2}\Bigg[\int_r^\infty \rho e^{\rho\cos\eta}\left(\left\|R\left(\frac{\rho e^{-i\eta}}{t}, B\right)\right\|_{B(E)} + \left\|R\left(\frac{\rho e^{i\eta}}{t}, B\right)\right\|_{B(E)}\right)\, d\rho + \\
&\qquad + \int_{-\eta}^{\eta} r^2 e^{r\cos\varphi}\left\|R\left(\frac{re^{i\varphi}}{t}, B\right)\right\|_{B(E)}\, d\varphi\Bigg] \\
&\leq \frac{M}{2\pi t}\left[2\int_r^\infty e^{\rho\cos\eta}\, d\rho + \int_{-\eta}^{\eta} r e^{r\cos\varphi}\, d\varphi\right] \leq \frac{M_1}{t}
\end{aligned}$$

since $\pi/2 < \eta < \theta < \pi$. Hence $\|t(A - \omega I)U_A(t)\|_{B(E)} \leq M_1 e^{\omega t}$ for all $t > 0$.

From the equality $BU_B(t) = U_B(t)B$ on $\mathcal{D}(B)$ it follows that

$$B^k U_B(t) = B^k\left(U_B\left(\frac{t}{k}\right)\right)^k = \left(BU_B\left(\frac{t}{k}\right)\right)^k$$

for all $k \in \mathbb{N}$ and $t > 0$. So for all $t > 0$

$$\|B^k U_B(t)\|_{B(E)} \le \left\| BU_B\left(\frac{t}{k}\right)\right\|_{B(E)}^k \le \left(\frac{M_1 k}{t}\right)^k \le (M_1 e)^k k! t^{-k}.$$

Hence $\|t^k (A - \omega I)^k U_A(t)\|_{B(E)} \le M_k e^{\omega t}$ for all $t > 0$ where $M_k = (M_1 e)^k k!$.

By using the induction we are able to prove that for all $k \in \mathbb{N}$ and $t > 0$

$$\|t^k A^k U_A(t)\|_{B(E)} \le \tilde{C}_k (1 + t + \cdots + t^k) e^{\omega t} \le C_k e^{(\omega+1)t}.$$

Let $t > 0$. The beginning of the induction is shown by

$$\begin{aligned}\|tAU_A(t)\|_{B(E)} &\le \|t(A - \omega I)U_A(t)\|_{B(E)} + |\omega| t \|U_A(t)\|_{B(E)} \\ &\le M_1 e^{\omega t} + |\omega| t M_0 e^{\omega t} \\ &\le \tilde{C}_1 (1 + t) e^{\omega t} \le C_1 e^{(\omega+1)t}.\end{aligned}$$

Let us assume that

$$\|t^k A^k U_A(t)\|_{B(E)} \le \tilde{C}_k (1 + t + \cdots + t^k) e^{\omega t} \le C_k e^{(\omega+1)t}$$

for all $k < n$ and $t > 0$. Then for all $t > 0$

$$\begin{aligned}\|t^n A^n U_A(t)\|_{B(E)} &\le \|t^n (A - \omega I)^n U_A(t)\|_{B(E)} + \sum_{l=0}^{n-1} \tbinom{n}{l} |\omega|^{n-l} t^{n-l} \|t^l A^l U_A(t)\|_{B(E)} \\ &\le M_n e^{\omega t} + \sum_{l=0}^{n-1} \tbinom{n}{l} |\omega|^{n-l} \tilde{C}_l (t^{n-l} + \cdots + t^n) e^{\omega t} \\ &\le \tilde{C}_n (1 + t + \cdots + t^n) e^{\omega t} \le C_n e^{(\omega+1)t}.\end{aligned}$$

Hence $\|t^k A^k U_A(t)\|_{B(E)} \le C_k e^{(\omega+1)t}$ for all $k \in \mathbb{N}$ and $t > 0$.

(iv) For all $t > 0$

$$\frac{d}{dt} U(t) = \frac{1}{2\pi i} \int_{\omega + \gamma_{r,\eta}} \lambda e^{t\lambda} R(\lambda, A)\, d\lambda = AU(t).$$

Hence

$$\frac{d^k}{dt^k} U(t) = A^k U(t)$$

for all $k \in \mathbb{N}$ and $t > 0$. By the statement (iii) the function $t \mapsto U(t)$ belongs to the space $C^\infty((0, \infty); B(E))$.

Let $0 < \varepsilon < \theta - \pi/2$ and choose $\eta = \theta - \varepsilon$. Since A is sectorial, the operator valued function

$$z \mapsto U(z) = \frac{1}{2\pi i} \int_{\omega + \gamma_{r,\eta}} e^{z\lambda} R(\lambda, A)\, d\lambda$$

is well defined and holomorphic in the sector

$$S_\varepsilon = \left\{ z \in \mathbb{C} : z \neq 0,\ |\arg z| < \theta - \frac{\pi}{2} - \varepsilon \right\}.$$

The union of the sectors S_ε for all $0 < \varepsilon < \theta - \pi/2$ is S. $\square$

Corollary 2.3. *The operator family $\{U(t)\}_{t\geq 0}$ defined by Formulas (2.3) and (2.4) is an analytic semigroup.*

In the following proposition we study, how the analytic semigroup $\{U(t)\}_{t\geq 0}$ behaves at the origin.

Proposition 2.4. *Let A be a sectorial operator with the constants ω, θ and M and the analytic semigroup $\{U(t)\}_{t\geq 0}$ defined by Formulas (2.3) and (2.4). Then the following statements are valid.*

(i) If $x \in \overline{\mathcal{D}(A)}$,

$$\lim_{t\to 0^+} U(t)x = x.$$

Conversely, if there exists

$$y = \lim_{t\to 0^+} U(t)x,$$

$x \in \overline{\mathcal{D}(A)}$ and $y = x$.

(ii) For all $x \in E$ and $t \geq 0$ the integral $\int_0^t U(s)x\, ds$ belongs to the set $\mathcal{D}(A)$ and

$$A\int_0^t U(s)x\, ds = U(t)x - x.$$

If, in addition, the function $s \mapsto AU(s)x$ belongs to the space $L^1(0, t; E)$,

$$U(t)x - x = \int_0^t AU(s)x\, ds.$$

(iii) If $x \in \mathcal{D}(A)$ and $Ax \in \overline{\mathcal{D}(A)}$,

$$\lim_{t\to 0^+} \frac{U(t)x - x}{t} = Ax.$$

Conversely, if there exists

$$z = \lim_{t\to 0^+} \frac{U(t)x - x}{t},$$

$x \in \mathcal{D}(A)$ and $z = Ax \in \overline{\mathcal{D}(A)}$.

Proof. (i) Let $\xi > \omega$ and $0 < r < \xi - \omega$. For every $x \in \mathcal{D}(A)$ we denote $y := \xi x - Ax$. Then by the resolvent identity,

$$\begin{aligned}
U(t)x &= U(t)R(\xi, A)y \\
&= \frac{1}{2\pi i}\int_{\omega+\gamma_{r,\eta}} e^{t\lambda} R(\lambda, A)R(\xi, A)y\, d\lambda \\
&= \frac{1}{2\pi i}\int_{\omega+\gamma_{r,\eta}} e^{t\lambda} \frac{R(\lambda, A) - R(\xi, A)}{\xi - \lambda} y\, d\lambda \\
&= \frac{1}{2\pi i}\int_{\omega+\gamma_{r,\eta}} e^{t\lambda} \frac{R(\lambda, A)}{\xi - \lambda} y\, d\lambda - \frac{1}{2\pi i}\int_{\omega+\gamma_{r,\eta}} \frac{e^{t\lambda}}{\xi - \lambda} R(\xi, A)x\, d\lambda \\
&= \frac{1}{2\pi i}\int_{\omega+\gamma_{r,\eta}} e^{t\lambda} \frac{R(\lambda, A)}{\xi - \lambda} y\, d\lambda
\end{aligned}$$

since

$$\int_{\omega+\gamma_{r,\eta}} \frac{e^{t\lambda}}{\xi-\lambda}\, d\lambda = 0$$

when $\xi > \omega$. Hence by Theorems B.20 and C.2,

$$\begin{aligned}\lim_{t\to 0^+} U(t)x &= \lim_{t\to 0^+} \frac{1}{2\pi i}\int_{\omega+\gamma_{r,\eta}} e^{t\lambda}\frac{R(\lambda,A)}{\xi-\lambda} y\, d\lambda \\ &= \frac{1}{2\pi i}\int_{\omega+\gamma_{r,\eta}} \frac{R(\lambda,A)}{\xi-\lambda} y\, d\lambda \\ &= R(\xi,A)y = x\end{aligned}$$

for each $x \in \mathcal{D}(A)$. Since $\mathcal{D}(A)$ is dense in $\overline{\mathcal{D}(A)}$ and $U(t)$ is continuous in $(0,\infty)$, $\lim_{t\to 0^+} U(t)x = x$ for all $x \in \overline{\mathcal{D}(A)}$.

Conversely, if $y = \lim_{t\to 0^+} U(t)x$, then $y \in \overline{\mathcal{D}(A)}$ because $U(t)x \in \mathcal{D}(A)$ for all $t > 0$ and $x \in E$. Moreover for $\xi \in \rho(A)$

$$R(\xi,A)y = \lim_{t\to 0^+} R(\xi,A)U(t)x = \lim_{t\to 0^+} U(t)R(\xi,A)x = R(\xi,A)x$$

since $R(\xi,A)R(\lambda,A) = R(\lambda,A)R(\xi,A)$ for all $\lambda,\xi \in \rho(A)$ and $R(\xi,A)x \in \mathcal{D}(A)$ for all $x \in E$. Therefore $y = x$.

(ii) Let $t > 0$, $x \in E$ and $\xi \in \rho(A)$. Then for every $\varepsilon \in (0,t)$

$$\begin{aligned}\int_\varepsilon^t U(s)x\, ds &= \int_\varepsilon^t (\xi - A)R(\xi,A)U(s)x\, ds \\ &= \xi\int_\varepsilon^t R(\xi,A)U(s)x\, ds - \int_\varepsilon^t R(\xi,A)AU(s)x\, ds \\ &= \xi\int_\varepsilon^t R(\xi,A)U(s)x\, ds - \int_\varepsilon^t \frac{d}{ds}(R(\xi,A)U(s)x)\, ds \\ &= \xi\int_\varepsilon^t R(\xi,A)U(s)x\, ds - R(\xi,A)U(t)x + R(\xi,A)U(\varepsilon)x \\ &= \xi\int_\varepsilon^t R(\xi,A)U(s)x\, ds - R(\xi,A)U(t)x + U(\varepsilon)R(\xi,A)x.\end{aligned}$$

The integral is well defined since $U(t)$ is continuous in $(0,\infty)$ and $\|U(t)\|_{B(E)} \leq \max\left(1, M_0, M_0e^{\omega t}\right)$ for all $t \geq 0$. Since $R(\xi,A)x \in \mathcal{D}(A)$, by Theorem B.20 and the statement (i),

$$\int_0^t U(s)x\, ds = \xi R(\xi,A)\int_0^t U(s)x\, ds - R(\xi,A)(U(t)x - x).$$

Hence $\int_0^t U(s)x\, ds$ belongs to $\mathcal{D}(A)$ and

$$A\int_0^t U(s)x\, ds = U(t)x - x$$

for all $t \geq 0$ and $x \in E$.

(iii) Let $t > 0$, $x \in \mathcal{D}(A)$ and $Ax \in \overline{\mathcal{D}(A)}$. Then

$$\frac{U(t)x - x}{t} = \frac{1}{t}A\int_0^t U(s)x\, ds = \frac{1}{t}\int_0^t AU(s)x\, ds = \frac{1}{t}\int_0^t U(s)Ax\, ds$$

since A is sectorial and the function $s \mapsto U(s)Ax$ is continuous in $[0,t]$ by the statement (i). Thus

$$\lim_{t\to 0^+} \frac{U(t)x - x}{t} = \lim_{t\to 0^+} \frac{1}{t}\int_0^t U(s)Ax\, ds = U(0)Ax = Ax$$

by the continuity.

Conversely, if there exists

$$z = \lim_{t\to 0^+} \frac{U(t)x - x}{t},$$

$\lim_{t\to 0^+} U(t)x = x$. Thus $x \in \overline{\mathcal{D}(A)}$ and therefore $z \in \overline{\mathcal{D}(A)}$. For every $\xi \in \rho(A)$

$$R(\xi, A)z = \lim_{t\to 0^+} R(\xi, A)\frac{U(t)x - x}{t}.$$

By the statement (ii),

$$R(\xi, A)z = \lim_{t\to 0^+} \frac{1}{t} R(\xi, A)A \int_0^t U(s)x\, ds = \lim_{t\to 0^+} (\xi R(\xi, A) - I)\frac{1}{t}\int_0^t U(s)x\, ds.$$

Since $x \in \overline{\mathcal{D}(A)}$, the function $s \mapsto U(s)x$ is continuous near $s = 0$. Hence $R(\xi, A)z = \xi R(\xi, A)x - x$. Therefore $x \in \mathcal{D}(A)$ and $z = Ax$. □

Corollary 2.3 and Proposition 2.4 motivate the following definition.

Definition 2.5. *Let $A : \mathcal{D}(A) \subset E \to E$ be a sectorial operator. The operator family $\{U(t)\}_{t\geq 0}$ defined by Formulas (2.3) and (2.4) is said to be the analytic semigroup generated by the operator A.*

Often in literature the analytic semigroup $\{U(t)\}_{t\geq 0}$ generated by a sectorial operator A is denoted by $\{e^{tA}\}_{t\geq 0}$. It can be seen as an extension of the exponent function to unbounded sectorial operators. We prefer the notation $\{U(t)\}_{t\geq 0}$.

If the operator A is sectorial with the constants ω, θ and M, the analytic semigroup $\{U(t)\}_{t>0}$ defined by Formulas (2.3) and (2.4) is analytic in the sector

$$\left\{ z \in \mathbb{C} : z \neq 0,\ |\arg z| < \theta - \frac{\pi}{2} \right\}.$$

Hence it is strongly continuous if and only if $\lim_{t\to 0^+} U(t)x = x$ for all $x \in E$. According the statement (i) of Proposition 2.4 $\lim_{t\to 0^+} U(t)x = x$ if and only if $x \in \overline{\mathcal{D}(A)}$. Thus the analytic semigroup $\{U(t)\}_{t\geq 0}$ is strongly continuous if and only if the domain $\mathcal{D}(A)$ is dense in E.

In Chapter 3 we shall need the following proposition. It gives sufficient conditions for a linear operator to be sectorial.

Proposition 2.6. *Let $A : \mathcal{D}(A) \subset E \to E$ be a linear operator such that the resolvent set $\rho(A)$ contains a half plane $\{\lambda \in \mathbb{C} : \mathrm{Re}\lambda \geq \omega\}$ and the resolvent $R(\lambda, A)$ satisfies*

$$\|\lambda R(\lambda, A)\|_{B(E)} \leq M \tag{2.5}$$

if $\mathrm{Re}\lambda \geq \omega$ with $\omega \geq 0$ and $M > 0$. Then A is sectorial.

Proof. According to Proposition A.2 for every $r > 0$ the resolvent set $\rho(A)$ contains the open ball centered at $\omega \pm ir$ with radius $|\omega \pm ir|/M$. The union of such balls and the half plane $\{\lambda \in \mathbb{C} : \mathrm{Re}\lambda \geq \omega\}$ contains the sector

$$\{\lambda \in \mathbb{C} : |\arg(\lambda - \omega)| < \pi - \arctan M\} \tag{2.6}$$

since λ belongs to the open ball centered at $\omega + i\mathrm{Im}\lambda$ with radius $|\omega + i\mathrm{Im}\lambda|/M$ by

$$\begin{aligned} |\omega + i\mathrm{Im}\lambda - \lambda| &= \frac{|\mathrm{Im}\lambda|}{\tan(\pi - |\arg(\lambda - \omega)|)} = \frac{|\mathrm{Im}\lambda|}{-\tan|\arg(\lambda - \omega)|} \\ &< \frac{|\mathrm{Im}\lambda|}{-\tan(\pi - \arctan M)} = \frac{|\mathrm{Im}\lambda|}{M} \leq \frac{|\omega + i\mathrm{Im}\lambda|}{M} \end{aligned}$$

if λ belongs to Sector (2.6) and $\mathrm{Re}\lambda < \omega$. Hence the resolvent set contains a sector.

We need to prove that the norm of the resolvent operator has an upper bound of the form required in Definition 2.1 in some sector. Let λ belong to the sector

$$\{\lambda \in \mathbb{C} : |\arg(\lambda - \omega)| < \pi - \arctan 2M\} \tag{2.7}$$

and $\mathrm{Re}\lambda < \omega$. Then $\lambda = \omega \pm ir - \theta r/M$ with $r > 0$ and $0 < \theta < 1/2$. By Formula (A.1),

$$\begin{aligned} \|R(\lambda, A)\|_{B(E)} &= \left\| \sum_{n=0}^{\infty} (-1)^n (\lambda - (\omega \pm ir))^n R^{n+1}(\omega \pm ir, A) \right\|_{B(E)} \\ &\leq \sum_{n=0}^{\infty} |\lambda - (\omega \pm ir)|^n \|R(\omega \pm ir, A)\|_{B(E)}^{n+1} \\ &\leq \sum_{n=0}^{\infty} \left(\frac{\theta r}{M}\right)^n \left(\frac{M}{|\omega \pm ir|}\right)^{n+1} = \sum_{n=0}^{\infty} \frac{\theta^n r^n M}{(\omega^2 + r^2)^{\frac{n+1}{2}}} \\ &\leq \frac{M}{r} \sum_{n=0}^{\infty} \theta^n = \frac{M}{r} \frac{1}{1-\theta} < \frac{2M}{r}. \end{aligned}$$

Since $\lambda = \omega \pm ir - \theta r/M$ where $r > 0$ and $0 < \theta < 1/2$,

$$|\lambda - \omega| = \left| -\frac{\theta r}{M} \pm ir \right| = r\sqrt{1 + \frac{\theta^2}{M^2}} < r\sqrt{1 + \frac{1}{4M^2}}.$$

Hence

$$r > \left(1 + \frac{1}{4M^2}\right)^{-\frac{1}{2}} |\lambda - \omega|.$$

Thus

$$\|R(\lambda, A)\|_{B(E)} < \frac{2M}{|\lambda - \omega|} \left(1 + \frac{1}{4M^2}\right)^{\frac{1}{2}}$$

for all λ such that λ belongs to Sector (2.7) and $\mathrm{Re}\lambda < \omega$. Furthermore, for all λ with $\mathrm{Re}\lambda \geq \omega$

$$\|R(\lambda, A)\|_{B(E)} \leq \frac{M}{|\lambda|} \leq \frac{M}{|\lambda - \omega|}.$$

Hence A is sectorial. $\square$

2.2 Homogeneous Initial Value Problems

Let $(E, \|\cdot\|_E)$ be a Banach space. Let $A : \mathcal{D}(A) \subseteq E \to E$ be a linear operator with not necessarily dense domain $\mathcal{D}(A)$. We are dealing with a solution to the homogeneous initial value problem

$$\begin{cases} u'(t) = Au(t), \\ u(0) = u_0 \end{cases} \tag{2.8}$$

in the space E with $t > 0$ and an arbitrary $u_0 \in E$.

Definition 2.7. *A function $u : [0, \infty) \to E$ is a (classical) solution to the initial value problem (2.8) on $[0, \infty)$, if u is continuous on $[0, \infty)$, continuously differentiable on $(0, \infty)$, $u(t) \in \mathcal{D}(A)$ for $0 < t < \infty$ and Equations (2.8) are satisfied on $[0, \infty)$.*

If the operator A is sectorial and the initial value u_0 belongs to $\overline{\mathcal{D}(A)}$, by the statement (iv) of Proposition 2.2 and the statement (i) of Proposition 2.4 a solution to the initial value problem (2.8) is given by the formula $u(t) = U(t)u_0$ for all $t > 0$ where $\{U(t)\}_{t \geq 0}$ is the analytic semigroup generated by the operator A. Let $u(t)$ be a solution to the initial value problem (2.8). Then $u(t) \in \mathcal{D}(A)$ for all $t > 0$ and the E-valued function $g(s) = U(t-s)u(s)$ is differentiable for $0 < s < t$. Hence

$$\frac{dg}{ds} = -AU(t-s)u(s) + U(t-s)u'(s) = 0$$

for all $0 < s < t$. By integrating from 0 to t we get

$$u(t) = U(t)u_0$$

for all $t > 0$.

Theorem 2.8. *If $U(t)$ is the analytic semigroup generated by a sectorial operator A and $u_0 \in \overline{\mathcal{D}(A)}$, the unique solution to the initial value problem (2.8) is $u(t) = U(t)u_0$ for all $t > 0$.*

2.3 Nonhomogeneous Initial Value Problems

This section is based on Section 4.2 and 4.3 in the book of Pazy [31].

Let $(E, \|\cdot\|_E)$ be a Banach space. Let $A : \mathcal{D}(A) \subseteq E \to E$ be a linear operator with dense domain $\mathcal{D}(A)$. We are considering the solution to the nonhomogeneous initial value problem

$$\begin{cases} u'(t) = Au(t) + f(t) \\ u(0) = u_0 \end{cases} \tag{2.9}$$

in the space E with $0 < t < T$, a known function $f : [0, T) \to E$ and an arbitrary $u_0 \in E$.

Definition 2.9. *A function $u : [0, T) \to E$ is a (classical) solution to the initial value problem (2.9) on $[0, T)$, if u is continuous on $[0, T)$, continuously differentiable on $(0, T)$, $u(t) \in \mathcal{D}(A)$ for $0 < t < T$ and Equations (2.9) are satisfied on $[0, T)$.*

We assume that A is sectorial with the constants ω, θ and M. Then the corresponding homogeneous problem has the unique solution for every $u_0 \in E$, namely $u(t) = U(t)u_0$ for all $t > 0$ where $U(t)$ is the analytic semigroup generated by A. Let $u(t)$ be a solution to the initial value problem (2.9). Then the E-valued function $g(s) = U(t-s)u(s)$ is differentiable for $0 < s < t$ and

$$\frac{dg}{ds} = -AU(t-s)u(s) + U(t-s)u'(s) = U(t-s)f(s).$$

If $f \in L^1(0,T;E)$, then $U(t-s)f(s)$ is Bochner integrable and by integrating from 0 to t we get

$$u(t) = U(t)u_0 + \int_0^t U(t-s)f(s)\, ds \tag{2.10}$$

for $0 \le t \le T$.

Theorem 2.10. *If $f \in L^1(0,T;E)$, for every $u_0 \in E$ the initial value problem (2.9) has at most one solution. If it has a solution, the solution is given by Formula (2.10).*

For every $f \in L^1(0,T;E)$ the right-hand side of (2.10) is a continuous function on $[0,T]$ since $U(t)$ is strongly continuous semigroup and there exists $M_0 > 0$ such that $\|U(t)\|_{B(E)} \le \max\left\{1, M_0, M_0e^{\omega T}\right\}$ for all $0 \le t \le T$. It is natural to consider Function (2.10) as a generalized solution to the initial value problem (2.9) even if it is not differentiable and does not satisfy the equation in the sense of Definition 2.9.

Definition 2.11. *Let $U(t)$ be the analytic semigroup generated by a densely defined sectorial operator A. Let $u_0 \in E$ and $f \in L^1(0,T;E)$. The function $u \in C([0,T];E)$ given by Formula (2.10) for $0 \le t \le T$ is the weak solution to the initial value problem (2.9) on $[0,T]$.*

Chapter 3

Sectorial Elliptic Operators

In this chapter we present a family of sectorial elliptic second order differential operators. The theory introduced in Chapter 2 can be applied to them to solve parabolic partial differential equations. This chapter is based on Section 3.1 and especially on Subsection 3.1.1 in the book of Lunardi [28]. Elliptic differential operator of the order $2m$, $m \geq 1$, has been handled among others in the books of Pazy [31] and Tanabe [50].

Let $n \geq 1$ and D be either $\mathbb{R}^n$ or an open subset of $\mathbb{R}^n$ with uniformly C^2-smooth boundary ∂D. We examine a second order differential operator

$$\mathcal{A}(x, \partial) = \sum_{i,j=1}^{n} a_{ij}(x)\partial_i\partial_j + \sum_{i=1}^{n} b_i(x)\partial_i + c(x) \tag{3.1}$$

with real uniformly continuous and bounded coefficient functions a_{ij}, b_i and c for all $i, j = 1, \ldots, n$. We assume that the matrix $[a_{ij}(x)]_{i,j=1}^n$ is symmetric for all $x \in \bar{D}$ and

$$\mathcal{A}_0(x, \xi) := \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \mu|\xi|^2 \tag{3.2}$$

for all $x \in \bar{D}$ and $\xi \in \mathbb{R}^n$ with some $\mu > 0$. Hence the differential operator $\mathcal{A}(x, \partial)$ is *elliptic*, i.e., $\mathcal{A}_0(x, \xi) \neq 0$ for all $x \in \bar{D}$ and $\xi \in \mathbb{R}^n \setminus \{0\}$. In addition, if $D \neq \mathbb{R}^n$, we consider a first order differential operator

$$\mathcal{B}(x, \partial) = \sum_{i=1}^{n} \beta_i(x)\partial_i + \gamma(x) \tag{3.3}$$

acting on the boundary ∂D. We assume that the coefficient functions β_i and γ are real uniformly continuously differentiable and bounded, i.e., belong to the space $UC^1(\bar{D})$ for all $i = 1, \ldots, n$ and that *the uniform nontangentiality condition*

$$\inf_{x \in \partial D} \left| \sum_{i=1}^{n} \beta_i(x)\nu_i(x) \right| > 0 \tag{3.4}$$

where $\nu(x) = (\nu_1(x), \ldots, \nu_n(x))$ is the exterior unit normal vector to ∂D at a point $x \in \partial D$ is valid. We are interested in realizations of the operator $\mathcal{A}(x, \partial)$ (with homogeneous boundary condition $\mathcal{B}(x, \partial)u = 0$ on ∂D if $D \neq \mathbb{R}^n$) in the space $L^p(D)$ with $1 < p < \infty$. As domains of the realizations we have the Sobolev space $W^{2,p}(D)$ or its subspace.

3.1 The Agmon-Douglis-Nirenberg Estimates

Our main purpose is to prove that the realizations of the operator $\mathcal{A}(x, \partial)$ (with homogeneous boundary condition if $D \neq \mathbb{R}^n$) are sectorial. The fundamental tools are the Agmon-Douglis-Nirenberg *a priori* estimates for elliptic problems in the whole $\mathbb{R}^n$ and regular domains of $\mathbb{R}^n$ when $n \geq 2$. The estimates are valid for differential operators of the type (3.1) with complex coefficient functions and under ellipticity assumptions

(i) $\left|\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j\right| \geq \mu|\xi|^2$ for all $x \in \bar{D}$, $\xi \in \mathbb{R}^n$ with some $\mu > 0$ and

(ii) if $\xi, \eta \in \mathbb{R}^n$ are linearly independent, for all $x \in \bar{D}$ the polynomial $\tau \mapsto P(\tau) = \sum_{i,j=1}^n a_{ij}(x)(\xi_i + \tau\eta_i)(\xi_j + \tau\eta_j)$ has a unique root such that its imaginary part is positive,

i.e., for differential operators which are *uniformly* and *properly elliptic*, respectively. If $n \geq 3$, then *the root condition* (ii) is not needed since all uniformly elliptic operators are properly elliptic [26, Proposition 1.2, p. 110]. Since we are also interested in the two dimensional case, both conditions have to be assumed. The following theorem formulates the Agmon-Douglis-Nirenberg *a priori* estimates.

Theorem 3.1 (The Agmon-Douglis-Nirenberg a Priori Estimates).

(i) Let a_{ij}, b_i, $c : \mathbb{R}^n \to \mathbb{C}$ be uniformly continuous and bounded functions for all $i, j = 1, \ldots, n$. Let $\mathcal{A}(x, \partial)$ be defined by Formula (3.1) and be uniformly and properly elliptic. Then for all $1 < p < \infty$ there exists such a constant $c_p > 0$ that for each $u \in W^{2,p}(\mathbb{R}^n)$

$$\|u\|_{W^{2,p}(\mathbb{R}^n)} \leq c_p \left(\|u\|_{L^p(\mathbb{R}^n)} + \|\mathcal{A}(\cdot, \partial)u\|_{L^p(\mathbb{R}^n)}\right). \tag{3.5}$$

(ii) Let D be an open set in $\mathbb{R}^n$ with uniformly C^2-smooth boundary and a_{ij}, b_i, $c : \bar{D} \to \mathbb{C}$ uniformly continuous and bounded functions for all $i, j = 1, \ldots, n$. Let $\mathcal{A}(x, \partial)$ be defined by Formula (3.1) and be uniformly and properly elliptic. Assume that β_i, $\gamma : \bar{D} \to \mathbb{C}$ belong to the space $UC^1(\bar{D})$ for all $i = 1, \ldots, n$. Let $\mathcal{B}(x, \partial)$ be defined by Formula (3.3) and satisfy the uniform nontangentiality condition (3.4). Then for all $1 < p < \infty$ there exists such a constant $c_p > 0$ that for each $u \in W^{2,p}(D)$

$$\|u\|_{W^{2,p}(D)} \leq c_p \left(\|u\|_{L^p(D)} + \|\mathcal{A}(\cdot, \partial)u\|_{L^p(D)} + \|g_1\|_{W^{1,p}(D)}\right) \tag{3.6}$$

where g_1 is any $W^{1,p}$-extension of $g = \mathcal{B}(\cdot, \partial)u|_{\partial D}$ to the whole D.

Proof. If the domain D is bounded, see [1, Theorem 15.2 pp. 704–706]. If the domain D is unbounded, see [2, Theorem 12.1 p. 653]. □

The reason why we have to consider complex valued coefficient functions is that in Section 3.2 we shall use Estimates (3.5) and (3.6) of the Agmon-Douglis-Nirenberg theorem 3.1 for the operator $\mathcal{A}_\theta(x, t, \partial) := \mathcal{A}(x, \partial) + e^{i\theta}\partial_t^2$ where $t \in \mathbb{R}$, $\theta \in [-\pi/2, \pi/2]$ and $x \in \bar{D}$. In the following lemma we show that the operator $\mathcal{A}_\theta(x, t, \partial)$ satisfies the Agmon-Douglis-Nirenberg assumptions if $\theta \in [-\pi/2, \pi/2]$.

Lemma 3.2. *The operator $\mathcal{A}_\theta(x,t,\partial)$ satisfies the Agmon-Douglis-Nirenberg assumptions if $t \in \mathbb{R}$, $\theta \in [-\pi/2, \pi/2]$ and $x \in \bar{D}$ and the operator $\mathcal{A}(x,\partial)$ has real uniformly continuous and bounded coefficient functions and it satisfies the ellipticity condition (3.2).*

Proof. The domain of the operator $\mathcal{A}_\theta(x,t,\partial)$ is $D \times \mathbb{R} \subset \mathbb{R}^{n+1}$. The terms of the second degree are $\sum_{i,j=1}^n a_{ij}\partial_i\partial_j + e^{i\theta}\partial_t^2$. Let $x \in \bar{D}$ and $\xi \in \mathbb{R}^{n+1}$. We denote $\xi := (\xi', \xi_{n+1})$. Then

$$\begin{aligned}&\left|\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j + e^{i\theta}\xi_{n+1}^2\right|^2\\ &= \left(\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j\right)^2 + 2\xi_{n+1}^2\cos\theta\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j + \xi_{n+1}^4.\end{aligned}$$

Since $\theta \in [-\pi/2, \pi/2]$, the values of the cosine function are non-negative. According to the ellipticity condition (3.2) the sum $\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j$ is positive. Thus

$$\begin{aligned}\left|\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j + e^{i\theta}\xi_{n+1}^2\right|^2 &\geq \mu^2|\xi'|^4 + \xi_{n+1}^4\\ &\geq \min(\mu,1)^2(|\xi'|^4 + \xi_{n+1}^4)\\ &\geq \frac{\min(\mu,1)^2}{2}|\xi|^4.\end{aligned}$$

Hence the operator $\mathcal{A}_\theta(x,t,\partial)$ is uniformly elliptic with the constant $\min(\mu,1)/\sqrt{2}$.

If $n \geq 2$, then $n+1 \geq 3$. Thus the uniform ellipticity implies the proper ellipticity. We still have to prove that the root condition is valid also if $n = 1$. Then the operator $\mathcal{A}_\theta(x,t,\partial)$ is of the form

$$\mathcal{A}_\theta(x,t,\partial) = a(x)\partial_x^2 + e^{i\theta}\partial_t^2 + b(x)\partial_x + c(x).$$

Since $\mathcal{A}(x,\partial)$ satisfies the ellipticity condition (3.2), we know that $a(x) \geq \mu > 0$ for all $x \in \bar{D}$. Let $\xi, \eta \in \mathbb{R}^2$ be linearly independent and $x \in \bar{D}$. The polynomial $P(\tau)$ is

$$P(\tau) = \left(a(x)\eta_1^2 + e^{i\theta}\eta_2^2\right)\tau^2 + 2\left(a(x)\xi_1\eta_1 + e^{i\theta}\xi_2\eta_2\right)\tau + a(x)\xi_1^2 + e^{i\theta}\xi_2^2.$$

The discriminant of the second order equation $P(\tau) = 0$ is

$$-4e^{i\theta}a(x)(\eta_1\xi_2 - \eta_2\xi_1)^2 = -4e^{i\theta}a(x)(\det[\eta,\xi])^2.$$

Thus the roots of the polynomial $P(\tau)$ are

$$\tau = \frac{\left(a(x)\eta_1^2 + e^{-i\theta}\eta_2^2\right)\left[-\left(a(x)\xi_1\eta_1 + e^{i\theta}\xi_2\eta_2\right) + i|\det[\eta,\xi]|\sqrt{e^{i\theta}a(x)}\right]}{a(x)^2\eta_1^4 + 2a(x)(\cos\theta)\eta_1^2\eta_2^2 + \eta_2^4}.$$

The denominator is always positive. Let us study separately the imaginary parts of the terms of the numerator. For the first term we get

$$\operatorname{Im}\left[-\left(a(x)\eta_1^2 + e^{-i\theta}\eta_2^2\right)\left(a(x)\xi_1\eta_1 + e^{i\theta}\xi_2\eta_2\right)\right] = -a(x)(\sin\theta)\eta_1\eta_2\det[\eta,\xi].$$

The imaginary part of the second term is

$$\begin{aligned}
&\operatorname{Im}\left[i\left(a(x)\eta_1^2+e^{-i\theta}\eta_2^2\right)|\det[\eta,\xi]|\sqrt{e^{i\theta}a(x)}\right]\\
&=\sqrt{a(x)}|\det[\eta,\xi]|\left[a(x)\cos\left(\frac{\theta}{2}+n\pi\right)\eta_1^2+\cos\left(\frac{\theta}{2}-n\pi\right)\eta_2^2\right]
\end{aligned}$$

where $n=0,1$. Thus the imaginary parts of the roots are

$$\begin{aligned}
\operatorname{Im}\tau_1 =& \frac{\sqrt{a(x)}\cos\left(\frac{\theta}{2}\right)|\det[\eta,\xi]|}{a(x)^2\eta_1^4+2a(x)(\cos\theta)\eta_1^2\eta_2^2+\eta_2^4}\times\\
&\times\left\{\left[\sqrt{a(x)}\eta_1-\operatorname{sgn}(\det[\eta,\xi])\left(\sin\frac{\theta}{2}\right)\eta_2\right]^2+\left(\cos\frac{\theta}{2}\right)^2\eta_2^2\right\}
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{Im}\tau_2 =& -\frac{\sqrt{a(x)}\cos\left(\frac{\theta}{2}\right)|\det[\eta,\xi]|}{a(x)^2\eta_1^4+2a(x)(\cos\theta)\eta_1^2\eta_2^2+\eta_2^4}\times\\
&\times\left\{\left[\sqrt{a(x)}\eta_1+\operatorname{sgn}(\det[\eta,\xi])\left(\sin\frac{\theta}{2}\right)\eta_2\right]^2+\left(\cos\frac{\theta}{2}\right)^2\eta_2^2\right\}.
\end{aligned}$$

Since $\theta\in[-\pi/2,\pi/2]$ and the vectors ξ and η are linearly independent, $\cos(\theta/2)\in[1/\sqrt{2},1]$ and $\det[\eta,\xi]\neq 0$. Hence $\operatorname{Im}\tau_1>0$ and $\operatorname{Im}\tau_2<0$. Thus the root condition is valid for the operator $\mathcal{A}_\theta(x,t,\partial)$ if $n=1$. □

3.2 Sectoriality

Let $D=\mathbb{R}^n$. We define the operator $A_0:\mathcal{D}(A_0)\to L^p(\mathbb{R}^n)$ by

$$\begin{aligned}
&\mathcal{D}(A_0):=W^{2,p}(\mathbb{R}^n),\\
&A_0u:=\mathcal{A}(\cdot,\partial)u \quad \text{if } u\in\mathcal{D}(A_0)
\end{aligned}$$

where the operator $\mathcal{A}(x,\partial)$ is defined by Formula (3.1). The operator A_0 is said to be a realization of the operator $\mathcal{A}(x,\partial)$ in $L^p(\mathbb{R}^n)$. The domain $\mathcal{D}(A_0)$ is dense in $L^p(\mathbb{R}^n)$.

Let D be an open subset of $\mathbb{R}^n$ with uniformly C^2-smooth boundary. We define the operator $A_1:\mathcal{D}(A_1)\to L^p(D)$ by

$$\begin{aligned}
&\mathcal{D}(A_1):=\{u\in W^{2,p}(D):\mathcal{B}(\cdot,\partial)u=0\text{ on }\partial D\},\\
&A_1u:=\mathcal{A}(\cdot,\partial)u \quad \text{if } u\in\mathcal{D}(A_1)
\end{aligned}$$

where the operator $\mathcal{B}(x,\partial)$ is defined by Formula (3.3). The operator A_1 is called a realization of the operator $\mathcal{A}(x,\partial)$ in $L^p(D)$ with homogeneous Robin boundary condition. We note that the domain $\mathcal{D}(A_1)$ is dense in $L^p(D)$.

In the following theorem it is shown that the resolvent sets of the realizations A_0 and A_1 contain complex half planes. The assumptions for the operators $\mathcal{A}(x,\partial)$ and $\mathcal{B}(x,\partial)$ are those stated in the beginning of this chapter.

Theorem 3.3. *Let* $1 < p < \infty$.

(i) There exists $\omega_0 \in \mathbb{R}$ *such that* $\rho(A_0) \supset \{\lambda \in \mathbb{C} : \mathrm{Re}\lambda \geq \omega_0\}$.

(ii) Let $D \subset \mathbb{R}^n$ *be an open set with uniformly* C^2*-smooth boundary. Then there exists* $\omega_1 \in \mathbb{R}$ *such that* $\rho(A_1) \supset \{\lambda \in \mathbb{C} : \mathrm{Re}\lambda \geq \omega_1\}$. *If* D *is bounded, the constant* ω_1 *does not depend on* p.

Proof. See [13, Theorem 4.1. p. 160]. According to Lemma 3.2 the condition (AN; θ) is valid when $\theta \in [-\pi/2, \pi/2]$. Hence the resolvent set contains a half plane. $\square$

We want to prove that A_0 and A_1 are sectorial. According to Proposition 2.6 we need bounds of the type (2.5) for the norms of the resolvents of the operators A_0 and A_1. In the following theorem we present the needed estimates.

Theorem 3.4. *Let* $1 < p < \infty$.

(i) There exist $\omega_p \geq \omega_0$ *and* $M_p > 0$ *such that for all* $u \in \mathcal{D}(A_0)$

$$|\lambda| \|u\|_{L^p(\mathbb{R}^n)} \leq M_p \|(\lambda - A_0)u\|_{L^p(\mathbb{R}^n)}$$

if $\mathrm{Re}\lambda \geq \omega_p$.

(ii) Let $D \subset \mathbb{R}^n$ *be an open set with uniformly* C^2*-smooth boundary. Then there exist* $\omega_p \geq \omega_1$ *and* $M_p > 0$ *such that for all* $u \in \mathcal{D}(A_1)$

$$|\lambda| \|u\|_{L^p(D)} \leq M_p \|(\lambda - A_1)u\|_{L^p(D)}$$

if $\mathrm{Re}\lambda \geq \omega_p$.

Proof. Let D be either $\mathbb{R}^n$ or an open subset of $\mathbb{R}^n$ with uniformly C^2-smooth boundary and $\theta \in [-\pi/2, \pi/2]$. We study the operator of $n+1$ variables

$$\mathcal{A}_\theta(x, t, \partial) := \mathcal{A}(x, \partial) + e^{i\theta}\partial_t^2$$

where $t \in \mathbb{R}$ and $x \in \bar{D}$. According to Lemma 3.2 the operator $\mathcal{A}_\theta(x, t, \partial)$ satisfies the Agmon-Douglis-Nirenberg assumptions. Let $\zeta \in C_0^\infty(\mathbb{R})$ be such that $\mathrm{supp}\,\zeta \subset [-1, 1]$ and $\zeta \equiv 1$ on the interval $[-1/2, 1/2]$. For every $u \in W^{2,p}(D)$ and $r > 0$ we set $v_r(x, t) := \zeta(t)e^{irt}u(x)$ for all $t \in \mathbb{R}$ and $x \in D$. Then for all $t \in \mathbb{R}$ and $x \in D$

$$\begin{aligned}
&\mathcal{A}_\theta(x, t, \partial)v_r(x, t) \\
&= \zeta(t)e^{irt}\left(\mathcal{A}(x, \partial) - r^2 e^{i\theta}\right)u(x) + e^{i(\theta + rt)}\left(\zeta''(t) + 2ir\zeta'(t)\right)u(x).
\end{aligned}$$

Let us assume that $D = \mathbb{R}^n$ and $u \in \mathcal{D}(A_0)$. We use Estimate (3.5) for the function v_r. Then

$$\begin{aligned}
\|v_r\|_{W^{2,p}(\mathbb{R}^{n+1})} &\leq c_p\left(\|v_r\|_{L^p(\mathbb{R}^{n+1})} + \|\mathcal{A}_\theta(\cdot, \cdot, \partial)v_r\|_{L^p(\mathbb{R}^{n+1})}\right) \\
&\leq c_p\Bigg(\left\|\zeta e^{irt}u\right\|_{L^p(\mathbb{R}^{n+1})} + \left\|\zeta e^{irt}\left(A_0 - r^2 e^{i\theta}\right)u\right\|_{L^p(\mathbb{R}^{n+1})} + \\
&\qquad + \left\|e^{i(\theta + rt)}(\zeta'' + 2ir\zeta')u\right\|_{L^p(\mathbb{R}^{n+1})}\Bigg).
\end{aligned}$$

The variables x and t can be separated. Thus

$$\|v_r\|_{W^{2,p}(\mathbb{R}^{n+1})} \le c_p \Big[\left(\|\zeta\|_{L^p(\mathbb{R})} + 2r\|\zeta'\|_{L^p(\mathbb{R})} + \|\zeta''\|_{L^p(\mathbb{R})}\right) \|u\|_{L^p(\mathbb{R}^n)} + \\ + \|\zeta\|_{L^p(\mathbb{R})} \left\| \left(A_0 - r^2 e^{i\theta}\right) u \right\|_{L^p(\mathbb{R}^n)} \Big].$$

We denote

$$c_p' := 2c_p \max\left\{\|\zeta\|_{L^p(\mathbb{R})},\ \|\zeta'\|_{L^p(\mathbb{R})},\ \|\zeta''\|_{L^p(\mathbb{R})}\right\}.$$

Then

$$\|v_r\|_{W^{2,p}(\mathbb{R}^{n+1})} \le c_p' \left[(1+r)\|u\|_{L^p(\mathbb{R}^n)} + \left\| \left(A_0 - r^2 e^{i\theta}\right) u \right\|_{L^p(\mathbb{R}^n)}\right].$$

On the other hand, since $\zeta \equiv 1$ on the interval $[-1/2, 1/2]$,

$$\begin{aligned}
&\|v_r\|^p_{W^{2,p}\left(\mathbb{R}^n \times \left(-\frac{1}{2},\frac{1}{2}\right)\right)} \\
&= \int_{\mathbb{R}^n \times \left(-\frac{1}{2},\frac{1}{2}\right)} \sum_{|\alpha|\le 2} |\partial^\alpha (u(x) e^{irt})|^p \, dxdt \\
&= \int_{\mathbb{R}^n} \left[(1 + r^p + r^{2p})|u(x)|^p + (1 + 2r^p) \sum_{j=1}^n |\partial_j u(x)|^p + \sum_{j,k=1}^n |\partial_j \partial_k u(x)|^p \right] dx \\
&\ge r^{2p} \|u\|^p_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

Hence

$$r^2 \|u\|_{L^p(\mathbb{R}^n)} \le c_p' \left[(1+r)\|u\|_{L^p(\mathbb{R}^n)} + \left\| \left(A_0 - r^2 e^{i\theta}\right) u \right\|_{L^p(\mathbb{R}^n)}\right].$$

We choose r so large that

$$c_p'(1+r) \le \frac{r^2}{2}$$

and denote $\lambda := r^2 e^{i\theta}$. Then

$$|\lambda| \|u\|_{L^p(\mathbb{R}^n)} \le 2c_p' \|(\lambda - A_0)u\|_{L^p(\mathbb{R}^n)}.$$

By choosing

$$\omega_p := \max\left\{2c_p'\left(c_p' + 1 + \sqrt{c_p'^2 + 2c_p'}\right), \omega_0\right\}$$

and $M_p := 2c_p'$ the statement (i) is proved.

The statement (ii) is shown in the same way. Instead of Estimate (3.5) we use Estimate (3.6). We assume that $u \in \mathcal{D}(A_1)$. Hence the proper extension of the boundary value $\mathcal{B}(\cdot, \partial) v_r|_{\partial D \times \mathbb{R}}$ to $D \times \mathbb{R}$ is the zero function. The only changes in the proof are that $\mathbb{R}^n$ is replaced by D, A_0 by A_1 and ω_0 by ω_1. □

Corollary 3.5. *The operators A_0 and A_1 are sectorial.*

Proof. We should show that the operators A_0 and A_1 satisfy the assumptions of Proposition 2.6. According to their definitions the operators A_0 and A_1 are linear. By Theorem 3.3 the resolvent sets of the operators A_0 and A_1 contain a half plane $\{\lambda \in \mathbb{C} : \mathrm{Re}\lambda \ge \omega_p\}$ where $\omega_p \ge 0$. Let D be either $\mathbb{R}^n$ or an open subset of $\mathbb{R}^n$

with uniformly C^2-smooth boundary. According to Theorem 3.4 for all $u \in \mathcal{D}(A_i)$, $i = 0, 1$,

$$|\lambda| \|u\|_{L^p(D)} \leq M_p \|(\lambda - A_i)u\|_{L^p(D)}$$

if $\mathrm{Re}\lambda \geq \omega_p$. By setting $u = R(\lambda, A_i)v$ we get

$$\|\lambda R(\lambda, A_i)v\|_{L^p(D)} \leq M_p \|v\|_{L^p(D)}$$

if $\mathrm{Re}\lambda \geq \omega_p$. Hence

$$\|\lambda R(\lambda, A_i)\|_{B(L^p(D))} \leq M_p$$

if $\mathrm{Re}\lambda \geq \omega_p$, i.e., a bound of the type (2.5) is valid for the operator A_i, $i = 0, 1$. Hence the operators A_0 and A_1 satisfy the assumptions of Proposition 2.6 with the constants ω_p and M_p and therefore they are sectorial. $\square$

Chapter 4

Stochastic Analysis in Infinite Dimensions

In this chapter we introduce the stochastic integral of operator valued stochastic processes with respect to the Hilbert space valued Wiener process. We present the concepts of the stochastic analysis in Banach and Hilbert spaces. The conditional expectation, Gaussian measures, martingales and the Wiener process are defined in this setting. The stochastic integral and its properties, especially the Ito formula, are the main purpose of this chapter. As an application we are able to solve linear stochastic initial value problems. This chapter is based on Chapters 1–5 in the book of Da Prato and Zabczyk [35]. We have included more detailed proofs for some theorems than those presented in [35]. In addition, we have corrected several misprints. Theorems 4.9 (partly) and 4.39, Propositions 4.25 and 4.40 and Lemmas 4.26, 4.33, 4.42 and 4.45 are used but not stated in [35]. Lemmas 4.13, 4.34 and 4.35 are Hilbert space versions of known results in $\mathbb{R}^n$. We could not find them from the literature. However, the proofs have only slight differences between those in $\mathbb{R}^n$. The definition of the weak solution to a linear stochastic initial value problem is different than the one in [35]. We have imitated the definition of the weak solution to deterministic nonhomogeneous initial value problems in Section 2.3. We assume that the reader is familiar with the Lebesgue integration theory of scalar valued functions and stochastic analysis in $\mathbb{R}^n$. The Bochner integration theory for vector valued functions is presented in Appendix B and the theory concerning nuclear and Hilbert-Schmidt operators in Appendix D.

4.1 Probability space

Let Ω be a set. A collection $\mathcal{F}$ of subsets of Ω is said to be a *σ-algebra* in Ω if $\mathcal{F}$ has the following properties

(i) $\Omega \in \mathcal{F}$,

(ii) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$,

(iii) if $A = \cup_{n=1}^{\infty} A_n$ and $A_n \in \mathcal{F}$ for all $n \in \mathbb{N}$, then $A \in \mathcal{F}$.

If $\mathcal{F}$ is a σ-algebra in Ω, then $(\Omega, \mathcal{F})$ is called a *measurable space* and the members of $\mathcal{F}$ are called the *measurable sets* in Ω. Let $\mathcal{K}$ be a collection of subsets of Ω. The smallest σ-algebra on Ω which contains $\mathcal{K}$ is denoted by $\sigma(\mathcal{K})$ and is called the *σ-algebra generated by* $\mathcal{K}$. Let E be a topological space. Then the *Borel σ-algebra* of E is the smallest σ-algebra containing all open subsets of E. It is denoted by $\mathcal{B}(E)$ and the elements of $\mathcal{B}(E)$ are called the *Borel sets* of E.

A collection $\mathcal{K}$ of subsets of Ω is said to be a *π-system* if $\emptyset \in \mathcal{K}$ and $A \cap B \in \mathcal{K}$ for all $A, B \in \mathcal{K}$. The following proposition is often used for proving that a given set is measurable.

Proposition 4.1. *Let $\mathcal{K}$ be a π-system and $\mathcal{G}$ the smallest family of subsets of Ω such that*

(i) $\mathcal{K} \subset \mathcal{G}$,

(ii) if $A \in \mathcal{G}$, *then* $A^c \in \mathcal{G}$,

(iii) if $A_n \in \mathcal{G}$ *for all* $n \in \mathbb{N}$ *and* $A_m \cap A_n = \emptyset$ *for all* $m \neq n$, *then* $\cup_{n=1}^{\infty} A_n \in \mathcal{G}$.

Then $\mathcal{G} = \sigma(\mathcal{K})$.

Proof. Let $A \in \mathcal{G}$. We define $\mathcal{G}_A := \{B \in \mathcal{G} : A \cap B \in \mathcal{G}\}$. Since

$$A \cap B^c = A \cap (A \cap B)^c = [A^c \cup (A \cap B)]^c$$

and $A^c \cap (A \cap B) = \emptyset$, then $A \cap B^c \in \mathcal{G}$ if $B \in \mathcal{G}_A$. Hence $B^c \in \mathcal{G}_A$ if $B \in \mathcal{G}_A$. If $B_n \in \mathcal{G}_A$ for all $n \in \mathbb{N}$ and $B_m \cap B_n = \emptyset$ for all $m \neq n$,

$$A \cap \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A \cap B_n \in \mathcal{G}$$

since $(A \cap B_m) \cap (A \cap B_n) = \emptyset$ for all $m \neq n$. Thus $\mathcal{G}_A$ satisfies the conditions (ii) and (iii). Since $\mathcal{K}$ is a π-system, $\mathcal{K} \subset \mathcal{G}_A$ for all $A \in \mathcal{K}$. Since $\mathcal{G}$ is the smallest family satisfying the conditions (i), (ii) and (iii), $\mathcal{G}_A = \mathcal{G}$ for all $A \in \mathcal{K}$. Hence if $B \in \mathcal{K}$, then $A \cap B \in \mathcal{G}$ for all $A \in \mathcal{G}$. Thus $\mathcal{K} \subset \mathcal{G}_A$ for all $A \in \mathcal{G}$ and consequently $\mathcal{G}_A = \mathcal{G}$ for all $A \in \mathcal{G}$. Therefore $\mathcal{G}$ is a π-system.

Let $A_n \in \mathcal{G}$ for all $n \in \mathbb{N}$. We define

$$\begin{aligned}
&B_1 := A_1, \\
&B_2 := A_2 \setminus B_1 = A_2 \cap B_1^c, \\
&B_3 := A_3 \setminus \cup_{i=1}^{2} B_i = A_3 \cap B_1^c \cap B_2^c, \\
&\quad \vdots \\
&B_n := A_n \setminus \cup_{i=1}^{n-1} B_i = A_n \cap B_1^c \cap \ldots \cap B_{n-1}^c.
\end{aligned}$$

Since $\mathcal{G}$ is a π-system and satisfies the condition (ii), $B_n \in \mathcal{G}$ for all $n \in \mathbb{N}$. Furthermore, $B_m \cap B_n = \emptyset$ for all $m \neq n$. Thus by the condition (iii),

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathcal{G}.$$

Hence $\mathcal{G}$ is a σ-algebra. Therefore $\sigma(\mathcal{K}) \subset \mathcal{G}$ because $\mathcal{K} \subset \mathcal{G}$. Since $\sigma(\mathcal{K})$ satisfies the conditions (i), (ii) and (iii), $\mathcal{G} \subset \sigma(\mathcal{K})$. Thus $\mathcal{G} = \sigma(\mathcal{K})$. $\square$

Let $(\Omega, \mathcal{F})$ be a measurable space. A function $\mu : \mathcal{F} \to [0, \infty]$ is a *positive measure* if $\mu(A) < \infty$ at least for one $A \in \mathcal{F}$ and μ is σ-additive, i.e., if $\{A_i\}_{i=1}^{\infty}$ is a disjoint collection of measurable sets,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

The triplet $(\Omega, \mathcal{F}, \mu)$ is called a *measure space*. If $(\Omega, \mathcal{F}, \mu)$ is a measure space, we define the *completion* of $\mathcal{F}$ by

$$\bar{\mathcal{F}} := \{A \subseteq \Omega : \text{there exist } B, C \in \mathcal{F} \text{ such that } B \subseteq A \subseteq C \text{ and } \mu(B) = \mu(C)\}.$$

Then $\bar{\mathcal{F}}$ is a σ-algebra. If $\mathcal{F} = \bar{\mathcal{F}}$, the measure space $(\Omega, \mathcal{F}, \mu)$ is said to be *complete*. A function $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a *probability measure* if $\mathbb{P}$ is a positive measure and $\mathbb{P}(\Omega) = 1$. The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space*.

4.2 Random Variables

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(E, \mathcal{G})$ a measurable space.

Definition 4.2. *A function $X : \Omega \to E$ such that the set $\{\omega \in \Omega : X(\omega) \in A\}$ belongs to $\mathcal{F}$ for each $A \in \mathcal{G}$ is called a measurable function or a random variable from $(\Omega, \mathcal{F})$ to $(E, \mathcal{G})$.*

If E is a topological vector space, an E-valued random variable is a function $X : \Omega \to E$ which is measurable from $(\Omega, \mathcal{F})$ to $(E, \mathcal{B}(E))$. A random variable is called *simple* if it is of the form

$$X(\omega) = \sum_{k=1}^{n} x_k \chi_{A_k}(\omega)$$

for all $\omega \in \Omega$ where $n \in \mathbb{N}$, $x_k \in E$ and $A_k \in \mathcal{F}$ are disjoint for all $k = 1, \ldots, n$ and

$$\chi_{A_k}(\omega) := \begin{cases} 1 & \text{if } \omega \in A_k, \\ 0 & \text{if } \omega \notin A_k. \end{cases}$$

Hence a simple random variable takes only a finite number of values.

Lemma 4.3. *Let (E, ρ) be a separable metric space and X an E-valued random variable. Then there exists a sequence $\{X_n\}_{n=1}^{\infty}$ of simple E-valued random variables such that for all $\omega \in \Omega$ the sequence $\{\rho(X(\omega), X_n(\omega))\}_{n=1}^{\infty}$ is monotonically decreasing to zero.*

Proof. Let $\{e_j\}_{j=1}^{\infty}$ be a countable dense subset of E. For $\omega \in \Omega$ and $n \in \mathbb{N}$ we define

$$\begin{aligned} \rho_n(\omega) &:= \min\{\rho(X(\omega), e_j),\ j = 1, \ldots, n\}, \\ k_n(\omega) &:= \min\{j \leq n : \rho_n(\omega) = \rho(X(\omega), e_j)\}, \\ X_n(\omega) &:= e_{k_n(\omega)}. \end{aligned}$$

Then X_n is a simple random variable since $X_n(\Omega) \subset \{e_1, \dots, e_n\}$, for all $i = 1, \dots, n$

$$\begin{aligned}
&\{\omega \in \Omega : X_n(\omega) = e_i\} \\
&= \bigcap_{j=1}^{i-1} \{\omega \in \Omega : \rho_n(\omega) \neq \rho(X(\omega), e_j)\} \cap \{\omega \in \Omega : \rho_n(\omega) = \rho(X(\omega), e_i)\}
\end{aligned}$$

and ρ_n is a random variable from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Moreover, by the density of $\{e_j\}_{j=1}^\infty$ the sequence $\{\rho_n(\omega)\}_{n=1}^\infty$ is monotonically decreasing to zero for each $\omega \in \Omega$. Since $\rho_n(\omega) = \rho(X(\omega), X_n(\omega))$, the conclusion follows. □

If X is a random variable from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(E, \mathcal{G})$, we denote the image of the probability measure $\mathbb{P}$ under the function X by $\mathcal{L}(X)$, i.e.,

$$\mathcal{L}(X)(A) := \mathbb{P}(\omega \in \Omega : X(\omega) \in A)$$

for all $A \in \mathcal{G}$. The probability measure $\mathcal{L}(X)$ is called the *distribution* or the *law* of the random variable X.

Let $\{X_i\}_{i \in I}$ be a family of functions from Ω to E. Then the smallest σ-algebra $\sigma(X_i : i \in I)$ on Ω such that all functions X_i are measurable from $(\Omega, \sigma(X_i : i \in I))$ to $(E, \mathcal{G})$ is called the σ-algebra generated by $\{X_i\}_{i \in I}$. Let $\mathcal{K}$ be a collection of subsets of E. If $\{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}$ for each $A \in \mathcal{K}$, then X is a measurable function from $(\Omega, \mathcal{F})$ to $(E, \sigma(\mathcal{K}))$ since the family of all sets $A \in \sigma(\mathcal{K})$ for which $\{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}$ is a σ-algebra.

Lemma 4.4. *Let $(H, (\cdot,\cdot)_H)$ be a separable Hilbert space and F a linearly dense subset of H. If X is a function from Ω to H such that for each $f \in F$ the function*

$$(X(\cdot), f)_H : \Omega \to \mathbb{C}$$

is measurable, X is a random variable from $(\Omega, \mathcal{F})$ to $(H, \mathcal{B}(H))$.

Proof. Since $\operatorname{span}(F)$ is dense in H, for every $h \in H$ there exists $\{f_n\}_{n=1}^\infty \subset \operatorname{span}(F)$ such that $f_n \to h$ in H as $n \to \infty$. Since $(X(\cdot), f)_H : \Omega \to \mathbb{C}$ is measurable for every $f \in F$, it is also measurable for every $f \in \operatorname{span}(F)$. Let $h \in H$ and $\{f_n\}_{n=1}^\infty \subset \operatorname{span}(F)$ be an approximating sequence of h. Then for all $\omega \in \Omega$

$$|(X(\omega), f_n)_H - (X(\omega), h)_H| \leq \|X(\omega)\|_H \|f_n - h\|_H \longrightarrow 0$$

as $n \to \infty$. Thus $(X(\cdot), h)_H : \Omega \to \mathbb{C}$ is measurable for every $h \in H$. Since by Corollary B.7 in a separable Banach space the weak measurability is equivalent to the measurability, X is measurable. □

Let $(E, \|\cdot\|_E)$ be a separable Banach space. An E-valued random variable X is said to be *Bochner integrable* or simply *integrable* if

$$\int_\Omega \|X(\omega)\|_E \, \mathbb{P}(d\omega) < \infty.$$

Then the integral of X is defined by

$$\mathbb{E}(X) := \int_\Omega X(\omega) \, \mathbb{P}(d\omega) := \lim_{n\to\infty} \int_\Omega X_n(\omega) \, \mathbb{P}(d\omega) := \lim_{n\to\infty} \sum_{k=1}^{m_n} x_k^n \mathbb{P}(A_k^n)$$

where $\{X_n\}_{n=1}^{\infty}$ is a sequence of simple random variables $X_n = \sum_{k=1}^{m_n} x_n^k \chi_{A_n^k}$ converging pointwise to X and satisfying

$$\int_{\Omega} \|X_m - X_n\|_E \, d\mathbb{P} \longrightarrow 0$$

as $m, n \to \infty$. We denote by $L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ the set of all equivalence classes of E-valued integrable random variables with respect to the equivalence relation $X(\omega) = Y(\omega)$ for almost all $\omega \in \Omega$, i.e., almost surely. The space $L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ equipped with the norm $\|X\|_1 := \mathbb{E}\|X\|_E$ is a Banach space. In a similar way one can define $L^p(\Omega, \mathcal{F}, \mathbb{P}; E)$ for each $p > 1$ with the norm $\|X\|_p = (\mathbb{E}\|X\|_E^p)^{1/p}$. The theory concerning Banach space valued functions and the Bochner integral is presented in Appendix B.

4.2.1 Operator Valued Random Variables

Let $(U, (\cdot, \cdot)_U)$ and $(H, (\cdot, \cdot)_H)$ be separable Hilbert spaces. We denote by $B(U, H)$ the collection of all bounded linear operator from U to H. If $U = H$, we use the notation $B(H) := B(H, H)$. The set $B(U, H)$ is a vector space and equipped with the operator norm

$$\|T\|_{B(U,H)} := \sup\{\|Tx\|_H : x \in U, \ \|x\|_U \leq 1\}$$

for all $T \in B(U, H)$ it is a Banach space. However, if U and H are both infinite dimensional, $B(U, H)$ is not separable. The non-separability of $B(U, H)$ has several consequences. First of all the corresponding Borel σ-algebra $\mathcal{B}(B(U, H))$ is very rich to the extent that very simple $B(U, H)$-valued functions turn out to be non-measurable. The lack of separability of $B(U, H)$ implies also that Bochner's definition of the integrability cannot be directly applied to the $B(U, H)$-valued functions.

Definition 4.5. *A function $\Phi : \Omega \to B(U, H)$ is said to be strongly measurable if for each $u \in U$ the function $\Phi(\cdot)u$ is measurable as a function from $(\Omega, \mathcal{F})$ to $(H, \mathcal{B}(H))$.*

Let $\mathcal{B}(U, H)$ be the smallest σ-algebra on $B(U, H)$ containing all sets of the form $\{\Psi \in B(U, H) : \Psi u \in A\}$ for each $u \in U$ and $A \in \mathcal{B}(H)$. Then a strongly measurable function $\Phi : \Omega \to B(U, H)$ is a measurable function from $(\Omega, \mathcal{F})$ to $(B(U, H), \mathcal{B}(U, H))$. The elements of $\mathcal{B}(U, H)$ are called *strongly measurable sets.*

Definition 4.6. *A strongly measurable function $\Phi : \Omega \to B(U, H)$ is said to be strongly Bochner integrable if for each $u \in U$ the function $\Phi(\cdot)u$ is Bochner integrable. Then there exists an operator $\Psi \in B(U, H)$ such that*

$$\int_{\Omega} \Phi(\omega)u \, \mathbb{P}(d\omega) = \Psi u$$

for each $u \in U$.

The operator Ψ is then denoted by

$$\Psi := \int_{\Omega} \Phi(\omega) \, \mathbb{P}(d\omega)$$

and called the *strong Bochner integral* of Φ. More about the Bochner integration theory of operator valued functions can be found in Appendix B.

For all $a, b \in H$ the bounded linear operator $a \otimes b$ is defined by $(a \otimes b)h := a(h, b)_H$ for each $h \in H$. If $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$, then $(X(\omega) - \mathbb{E}X) \otimes (X(\omega) - \mathbb{E}X)$ is a bounded linear operator in H for almost all $\omega \in \Omega$. Since X is measurable,

$$((X(\omega) - \mathbb{E}X) \otimes (X(\omega) - \mathbb{E}X))h = (X(\omega) - \mathbb{E}X)(h, X(\omega) - \mathbb{E}X)_H$$

is measurable as a function from $(\Omega, \mathcal{F})$ to $(H, \mathcal{B}(H))$ for all $h \in H$. Since $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$,

$$\begin{aligned}
&\int_\Omega \|((X(\omega) - \mathbb{E}X) \otimes (X(\omega) - \mathbb{E}X))h\|_H \, \mathbb{P}(d\omega) \\
&= \int_\Omega \|X(\omega) - \mathbb{E}X\|_H |(h, X(\omega) - \mathbb{E}X)_H| \, \mathbb{P}(d\omega) \\
&\leq \|h\|_H \|X - \mathbb{E}X\|_2^2 = \|h\|_H \left(\|X\|_2^2 - \|\mathbb{E}X\|_H^2\right)
\end{aligned}$$

for all $h \in H$. Hence $(X - \mathbb{E}X) \otimes (X - \mathbb{E}X)$ is strongly Bochner integrable and the bounded linear operator

$$\mathrm{Cov}(X) := \mathbb{E}[(X - \mathbb{E}X) \otimes (X - \mathbb{E}X)]$$

from H to itself is well defined. The operator $\mathrm{Cov}(X)$ is called the *covariance operator* of X. The operator $\mathrm{Cov}(X)$ is a non-negative self-adjoint operator, because

$$(\mathrm{Cov}(X)h, h)_H = \mathbb{E}\left[|(h, X - \mathbb{E}X)_H|^2\right] \geq 0$$

and

$$\begin{aligned}
(\mathrm{Cov}(X)h, g)_H &= (\mathbb{E}\left[(X - \mathbb{E}X)(h, X - \mathbb{E}X)_H\right], g)_H \\
&= \mathbb{E}\left[(h, X - \mathbb{E}X)_H (X - \mathbb{E}X, g)_H\right] \\
&= \overline{\mathbb{E}\left[(g, X - \mathbb{E}X)_H (X - \mathbb{E}X, h)_H\right]} \\
&= \overline{(\mathrm{Cov}(X)g, h)_H} = (h, \mathrm{Cov}(X)g)_H
\end{aligned}$$

for all $g, h \in H$. Let $\{e_k\}_{k=1}^\infty$ be an orthonormal basis in H. Then

$$\begin{aligned}
\mathrm{Tr}\ \mathrm{Cov}(X) &= \sum_{k=1}^\infty (\mathrm{Cov}(X)e_k, e_k)_H = \sum_{k=1}^\infty \mathbb{E}\left[|(X(\omega) - \mathbb{E}X, e_k)_H|^2\right] \\
&= \mathbb{E}\|X - \mathbb{E}X\|_H^2 = \|X\|_2^2 - \|\mathbb{E}X\|_H^2.
\end{aligned}$$

Thus by Proposition D.14 the covariance operator $\mathrm{Cov}(X)$ is nuclear.

If $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$, we may define the *correlation operator*

$$\mathrm{Cor}(X, Y) := \mathbb{E}[(X - \mathbb{E}X) \otimes (Y - \mathbb{E}Y)]$$

of X and Y as a strong Bochner integral similarly. Then $\mathrm{Cor}(X, Y)$ is a bounded linear operator from H to itself.

4.2.2 Conditional Expectation and Independence

The conditional expectation of scalar valued random variables is assumed to be known. The books of Neveu [30] or Williams [53] can be used as a reference.

Theorem 4.7. *Let $(E, \|\cdot\|_E)$ be a separable Banach space. A random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{B}(E))$ is assumed to be Bochner integrable and $\mathcal{G}$ to be a σ-algebra contained in $\mathcal{F}$. Then there exists an integrable E-valued random variable Z which is measurable with respect to $\mathcal{G}$ such that*

$$\int_A X \, d\mathbb{P} = \int_A Z \, d\mathbb{P}$$

for all $A \in \mathcal{G}$. Furthermore, Z is unique up to a set of probability zero.

The random variable Z is denoted by $\mathbb{E}(X|\mathcal{G})$ and called the *conditional expectation* of X given $\mathcal{G}$.

Proof. First we show the existence of a conditional expectation. Let X be a simple random variable. We define $Z := \sum_{k=1}^n x_k \mathbb{E}(\chi_{A_k}|\mathcal{G})$ where $\mathbb{E}(\chi_{A_k}|\mathcal{G})$ represents the classical notion of the conditional expectation of the characteristic function χ_{A_k} given $\mathcal{G}$. Since χ_{A_k} is non-negative, $\mathbb{E}(\chi_{A_k}|\mathcal{G}) \geq 0$ almost surely [53, Theorem 9.7]. Thus

$$\begin{aligned}
\mathbb{E}\|Z\|_E \leq \mathbb{E}\left(\sum_{k=1}^n \|x_k\|_E \mathbb{E}(\chi_{A_k}|\mathcal{G})\right) &= \sum_{k=1}^n \|x_k\|_E \mathbb{E}(\mathbb{E}(\chi_{A_k}|\mathcal{G})) \\
= \sum_{k=1}^n \|x_k\|_E \mathbb{E}(\chi_{A_k}) &= \sum_{k=1}^n \|x_k\|_E \mathbb{P}(A_k) = \mathbb{E}\|X\|_E.
\end{aligned}$$

Hence Z is integrable E-valued $\mathcal{G}$-measurable function. Furthermore,

$$\int_A X \, d\mathbb{P} = \sum_{k=1}^n x_k \int_A \chi_{A_k} \, d\mathbb{P} = \sum_{k=1}^n x_k \int_A \mathbb{E}(\chi_{A_k}|\mathcal{G}) \, d\mathbb{P} = \int_A Z \, d\mathbb{P}$$

for all $A \in \mathcal{G}$. Thus there exists a conditional expectation for a simple random variable.

Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$. By Theorem B.12 there exists a sequence $\{X_n\}_{n=1}^\infty$ of simple random variables such that X_n converges pointwise to X and $\mathbb{E}\|X_m - X_n\|_E \to 0$ as $m, n \to \infty$. Furthermore, $\mathbb{E}\|X - X_n\|_E \to 0$ as $n \to \infty$. Let $Z_n = \mathbb{E}(X_n|\mathcal{G})$ for all $n \in \mathbb{N}$. Then

$$\mathbb{E}\|Z_m - Z_n\|_E \leq \mathbb{E}\|X_m - X_n\|_E \longrightarrow 0$$

as $m, n \to \infty$. Hence $\{Z_n\}_{n=1}^\infty$ is a Cauchy sequence in $L^1(\Omega, \mathcal{G}, \mathbb{P}; E)$. Therefore there exists $Z \in L^1(\Omega, \mathcal{G}, \mathbb{P}; E)$ such that $\mathbb{E}\|Z - Z_n\|_E \to 0$ as $n \to \infty$. Hence for each $A \in \mathcal{G}$

$$\int_A X \, d\mathbb{P} = \lim_{n\to\infty} \int_A X_n \, d\mathbb{P} = \lim_{n\to\infty} \int_A Z_n \, d\mathbb{P} = \int_A Z \, d\mathbb{P}$$

since $X_n \to X$ in $L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ and $Z_n \to Z$ in $L^1(\Omega, \mathcal{G}, \mathbb{P}; E)$ as $n \to \infty$ and $Z_n = \mathbb{E}(X_n|\mathcal{G})$ for all $n \in \mathbb{N}$. Thus there exists a conditional expectation.

We still have to prove the uniqueness of the conditional expectation. We need the following lemma. In the proof of the lemma the uniqueness of the conditional expectation is not expected.

Lemma 4.8. *Let $(E, \|\cdot\|_E)$ be a separable Banach space. A random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{B}(E))$ is assumed to be Bochner integrable and $\mathcal{G}$ to be a σ-algebra contained in $\mathcal{F}$. If φ a continuous linear functional on E,*

$$\langle \mathbb{E}(X|\mathcal{G}), \varphi\rangle = \mathbb{E}(\langle X, \varphi\rangle|\mathcal{G})$$

almost surely.

Proof. Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ and $Z = \mathbb{E}(X|\mathcal{G})$. Then for all $A \in \mathcal{G}$

$$\int_A \langle X, \varphi\rangle \, d\mathbb{P} = \left\langle \int_A X \, d\mathbb{P}, \varphi \right\rangle = \left\langle \int_A Z \, d\mathbb{P}, \varphi \right\rangle = \int_A \langle Z, \varphi\rangle \, d\mathbb{P}.$$

Hence $\langle Z, \varphi\rangle = \mathbb{E}(\langle X, \varphi\rangle|\mathcal{G})$ almost surely since the conditional expectation of scalar valued random variable is unique up to a set of probability zero. □

We assume that there exist two random variables Z and $\tilde{Z}$ such that they have the properties of the conditional expectation and $\mathbb{P}(Z - \tilde{Z} \neq 0) > 0$. The separability of E implies that for some $a \in E$

$$\mathbb{P}\left(\|Z - \tilde{Z} - a\|_E < \frac{1}{3}\|a\|_E \right) > 0.$$

By the Hahn-Banach theorem there exists a continuous linear functional φ on E such that $\langle a, \varphi\rangle = \|a\|_E$ and $\|\varphi\|_{E'} = 1$. Hence

$$\mathbb{P}\left(\left| \langle Z - \tilde{Z}, \varphi\rangle - \|a\|_E \right| < \frac{1}{3}\|a\|_E \right) > 0$$

since

$$\left| \langle Z - \tilde{Z}, \varphi\rangle - \|a\|_E \right| = |\langle Z - \tilde{Z} - a, \varphi\rangle| \leq \|Z - \tilde{Z} - a\|_E.$$

Thus

$$\mathbb{P}\left(\langle Z - \tilde{Z}, \varphi\rangle \geq \frac{2}{3}\|a\|_E \right) > 0.$$

But $\langle Z, \varphi\rangle = \mathbb{E}(\langle X, \varphi\rangle|\mathcal{G}) = \langle \tilde{Z}, \varphi\rangle$ almost surely by Lemma 4.8. The obtained contradiction implies the uniqueness up to a set of probability zero. □

Let $\{\mathcal{F}_i\}_{i\in I}$ be a family of sub-σ-algebras of $\mathcal{F}$. These σ-algebras are said to be *independent* if for every finite subset $J \subset I$ and every family $\{A_i\}_{i\in J}$ such that $A_i \in \mathcal{F}_i$ for each $i \in J$

$$\mathbb{P}\left(\bigcap_{i\in J} A_i \right) = \prod_{i\in J} \mathbb{P}(A_i).$$

Random variables $\{X_i\}_{i\in I}$ are *independent* if the σ-algebras $\{\sigma(X_i)\}_{i\in I}$ are independent.

In the following theorem we have gathered properties of the conditional expectation.

Theorem 4.9. *Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be separable Banach spaces. Random variables $X, Y : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{B}(E))$ are assumed to be Bochner integrable and $\mathcal{G}$ to be a σ-algebra contained in $\mathcal{F}$. Then*

(i) $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$,

(ii) if X is $\mathcal{G}$-measurable, $\mathbb{E}(X|\mathcal{G}) = X$ almost surely,

(iii) $\mathbb{E}(\alpha X + \beta Y|\mathcal{G}) = \alpha\mathbb{E}(X|\mathcal{G}) + \beta\mathbb{E}(Y|\mathcal{G})$ almost surely for all $\alpha, \beta \in \mathbb{C}$,

(iv) if $\mathcal{H}$ is a sub-σ-algebra of $\mathcal{G}$, then $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{H}) = \mathbb{E}(X|\mathcal{H})$ almost surely,

(v) $\|\mathbb{E}(X|\mathcal{G})\|_E \le \mathbb{E}(\|X\|_E|\mathcal{G})$ almost surely,

(vi) if $\Phi \in B(E, F)$, then $\mathbb{E}(\Phi X|\mathcal{G}) = \Phi\mathbb{E}(X|\mathcal{G})$ almost surely,

(vii) if Z is a bounded scalar valued $\mathcal{G}$-measurable function, $\mathbb{E}(ZX|\mathcal{G}) = Z\mathbb{E}(X|\mathcal{G})$ almost surely,

(viii) if X is independent of $\mathcal{G}$, then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$ almost surely.

Proof. The statements (i)-(iv) are obvious consequences of the definition of the conditional expectation.

(v) Let X be a simple random variable. Then

$$\|\mathbb{E}(X|\mathcal{G})\|_E \le \sum_{k=1}^{n} \|x_k\|_E \mathbb{E}(\chi_{A_k}|\mathcal{G}) = \mathbb{E}(\|X\|_E|\mathcal{G}).$$

Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ and $\{X_n\}_{n=1}^{\infty}$ be a sequence defined in Theorem B.12. Then $X_n \to X$ in $L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ as $n \to \infty$. Therefore $\|X_n\|_E \to \|X\|_E$ in $L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ as $n \to \infty$. Similarly as in the proof of Theorem 4.7 we get $\mathbb{E}(X_n|\mathcal{G}) \to \mathbb{E}(X|\mathcal{G})$ in $L^1(\Omega, \mathcal{G}, \mathbb{P}; E)$ and $\mathbb{E}(\|X_n\|_E|\mathcal{G}) \to \mathbb{E}(\|X\|_E|\mathcal{G})$ in $L^1(\Omega, \mathcal{G}, \mathbb{P}; \mathbb{R})$ as $n \to \infty$. Since by Theorem B.16 every convergent sequence in L^1 has a subsequence which converges pointwise almost surely,

$$\|\mathbb{E}(X|\mathcal{G})\|_E = \lim_{k\to\infty} \|\mathbb{E}(X_{n_k}|\mathcal{G})\|_E \le \lim_{k\to\infty} \mathbb{E}(\|X_{n_k}\|_E|\mathcal{G}) = \mathbb{E}(\|X\|_E|\mathcal{G})$$

almost surely.

(vi) Let $\Phi \in B(E, F)$ and $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$. Then for all $A \in \mathcal{G}$

$$\int_A \Phi X \, d\mathbb{P} = \Phi \int_A X \, d\mathbb{P} = \Phi \int_A \mathbb{E}(X|\mathcal{G}) \, d\mathbb{P} = \int_A \Phi\mathbb{E}(X|\mathcal{G}) \, d\mathbb{P}.$$

Hence $\mathbb{E}(\Phi X|\mathcal{G}) = \Phi\mathbb{E}(X|\mathcal{G})$ almost surely.

(vii) Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ and Z be a bounded scalar valued $\mathcal{G}$-measurable function. Let φ be a continuous linear functional on E. Then by the statement (v),

$$\langle \mathbb{E}(ZX|\mathcal{G}), \varphi\rangle = \mathbb{E}(Z\langle X, \varphi\rangle|\mathcal{G}) = Z\mathbb{E}(\langle X, \varphi\rangle|\mathcal{G}) = \langle Z\mathbb{E}(X|\mathcal{G}), \varphi\rangle$$

almost surely [53, Theorem 9.7]. Thus $\mathbb{E}(ZX|\mathcal{G}) = Z\mathbb{E}(X|\mathcal{G})$ almost surely.

(viii) Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ be independent of $\mathcal{G}$ and φ a continuous linear functional on E. Then $\langle X, \varphi \rangle$ is independent of $\mathcal{G}$ [11, Theorem 4.1.1]. Hence by the statement (v),

$$\langle \mathbb{E}(X|\mathcal{G}), \varphi \rangle = \mathbb{E}(\langle X, \varphi \rangle | \mathcal{G}) = \mathbb{E}(\langle X, \varphi \rangle) = \langle \mathbb{E}(X), \varphi \rangle$$

almost surely [53, Theorem 9.7]. Thus $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$ almost surely. □

The following proposition is a useful tool in the construction of the stochastic integral in Section 4.5.

Proposition 4.10. *Let $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$ be measurable spaces and $\psi : E_1 \times E_2 \to \mathbb{R}$ a bounded measurable function. Let X_1 and X_2 be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$, respectively, and $\mathcal{G}$ a sub-σ-algebra of $\mathcal{F}$. If X_1 is $\mathcal{G}$-measurable and X_2 is independent of $\mathcal{G}$,*

$$\mathbb{E}(\psi(X_1, X_2)|\mathcal{G}) = \hat{\psi}(X_1)$$

almost surely where $\hat{\psi}(x_1) = \mathbb{E}(\psi(x_1, X_2))$ for $x_1 \in E_1$.

Proof. We assume first that $\psi(x_1, x_2) = \psi_1(x_1)\psi_2(x_2)$ for all $(x_1, x_2) \in E_1 \times E_2$ where $\psi_1 : E_1 \to \mathbb{R}$ and $\psi_2 : E_2 \to \mathbb{R}$ are bounded measurable functions. Then $\psi_1(X_1)$ is $\mathcal{G}$-measurable and $\psi_2(X_2)$ is independent of $\mathcal{G}$ [11, Theorem 4.1.1]. Hence

$$\mathbb{E}(\psi(X_1, X_2)|\mathcal{G}) = \mathbb{E}(\psi_1(X_1)\psi_2(X_2)|\mathcal{G}) = \psi_1(X_1)\mathbb{E}(\psi_2(X_2)|\mathcal{G}) = \psi_1(X_1)\mathbb{E}(\psi_2(X_2))$$

almost surely [53, Theorem 9.7]. So the claim is true in this case.

The set $\mathcal{K} := \{A_1 \times A_2 : A_1 \in \mathcal{E}_1,\ A_2 \in \mathcal{E}_2\}$ is a π-system and $\sigma(\mathcal{K}) = \mathcal{E}_1 \times \mathcal{E}_2$. Furthermore, $\mathbb{E}(\chi_{A_1 \times A_2}(X_1, X_2)|\mathcal{G}) = \hat{\chi}_{A_1 \times A_2}(X_1)$ almost surely for all $A_1 \times A_2 \in \mathcal{K}$ where

$$\hat{\chi}_{A_1 \times A_2}(x_1) = \chi_{A_1}(x_1)\mathbb{E}(\chi_{A_2}(X_2)) = \mathbb{E}(\chi_{A_1 \times A_2}(x_1, X_2))$$

for all $x_1 \in E_1$. Let

$$\mathcal{H} := \{A \in \mathcal{E}_1 \times \mathcal{E}_2 : \mathbb{E}(\chi_A(X_1, X_2)|\mathcal{G}) = \hat{\chi}_A(X_1) \text{ almost surely}\}.$$

Then $\mathcal{K} \subset \mathcal{H}$. Since $\chi_{A^c} = 1 - \chi_A$ and $\chi_{\cup_{n=1}^\infty A_n} = \sum_{n=1}^\infty \chi_{A_n}$ for all disjoint $A_n \in \mathcal{E}_1 \times \mathcal{E}_2$, the set $\mathcal{H}$ satisfies the assumptions of Proposition 4.1 [53, Theorem 9.7]. Thus $\mathcal{H} = \mathcal{E}_1 \times \mathcal{E}_2$. Therefore the claim is true for all simple functions ψ.

In the general case according to Lemma 4.3 there exists a sequence $\{\psi_n\}_{n=1}^\infty$ of simple functions such that $|\psi_n(x_1, x_2) - \psi(x_1, x_2)| \downarrow 0$ as $n \to \infty$ for all $(x_1, x_2) \in E_1 \times E_2$. Hence $|\psi_n(X_1(\omega), X_2(\omega)) - \psi(X_1(\omega), X_2(\omega))| \downarrow 0$ and $|\psi_n(x_1, X_2(\omega)) - \psi(x_1, X_2(\omega))| \downarrow 0$ as $n \to \infty$ for all $\omega \in \Omega$ and $x_1 \in E_1$. By Lebesgue's monotone convergence theorem $\psi_n(X_1, X_2) \to \psi(X_1, X_2)$ and $\psi_n(x_1, X_2) \to \psi(x_1, X_2)$ in $L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ as $n \to \infty$ for all $x_1 \in E_1$. Therefore, similarly as in the proof of Theorem 4.7 we get $\mathbb{E}(\psi_n(X_1, X_2)|\mathcal{G}) \to \mathbb{E}(\psi(X_1, X_2)|\mathcal{G})$ in $L^1(\Omega, \mathcal{G}, \mathbb{P}; \mathbb{R})$ and $\hat{\psi}_n(x_1) \to \hat{\psi}(x_1)$ as $n \to \infty$ for all $x_1 \in E_1$. Since $\mathbb{E}(\psi_n(X_1, X_2)|\mathcal{G}) = \hat{\psi}_n(X_1)$ almost surely and by Theorem B.16 every convergent sequence in L^1 has a subsequence which converges pointwise almost surely, $\mathbb{E}(\psi(X_1, X_2)|\mathcal{G}) = \hat{\psi}(X_1)$ almost surely. □

The following lemma is used in the proof of the Ito formula in Subsection 4.5.2.

Lemma 4.11. *If* $\{\eta_j\}_{j=1}^k$ *is a sequence of real valued random variables with finite second moments and* $\{\mathcal{G}_j\}_{j=1}^k$ *is an increasing sequence of* σ*-algebras such that* η_j *is measurable with respect to* $\mathcal{G}_i$ *for all* $1 \le j < i \le k$*,*

$$\mathbb{E}\left(\sum_{j=1}^k \eta_j - \sum_{j=1}^k \mathbb{E}(\eta_j|\mathcal{G}_j)\right)^2 = \sum_{j=1}^k \left(\mathbb{E}\eta_j^2 - \mathbb{E}(\mathbb{E}(\eta_j|\mathcal{G}_j))^2\right).$$

Proof. Since $\eta_j \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ for all $j = 1, \dots, k$,

$$\begin{aligned}
\mathbb{E}(\eta_j \mathbb{E}(\eta_j|\mathcal{G}_j)) &= \mathbb{E}(\mathbb{E}(\eta_j \mathbb{E}(\eta_j|\mathcal{G}_j)|\mathcal{G}_j)) = \mathbb{E}(\mathbb{E}(\eta_j|\mathcal{G}_j))^2, \\
\mathbb{E}(\eta_i \mathbb{E}(\eta_j|\mathcal{G}_j)) &= \mathbb{E}(\mathbb{E}(\eta_i \eta_j|\mathcal{G}_j)) = \mathbb{E}(\eta_i \eta_j), \\
\mathbb{E}(\eta_j \mathbb{E}(\eta_i|\mathcal{G}_i)) &= \mathbb{E}(\mathbb{E}(\eta_j \mathbb{E}(\eta_i|\mathcal{G}_i)|\mathcal{G}_j)) = \mathbb{E}(\mathbb{E}(\eta_i|\mathcal{G}_i)\mathbb{E}(\eta_j|\mathcal{G}_j))
\end{aligned}$$

if $i < j$ [53, Theorem 9.7]. Thus

$$\begin{aligned}
&\mathbb{E}\left(\sum_{j=1}^k \eta_j - \sum_{j=1}^k \mathbb{E}(\eta_j|\mathcal{G}_j)\right)^2 \\
&= \mathbb{E}\sum_{j=1}^k (\eta_j - \mathbb{E}(\eta_j|\mathcal{G}_j))^2 + 2\mathbb{E}\sum_{j=1}^k \sum_{i<j} (\eta_i - \mathbb{E}(\eta_i|\mathcal{G}_i))(\eta_j - \mathbb{E}(\eta_j|\mathcal{G}_j)) \\
&= \mathbb{E}\sum_{j=1}^k (\eta_j^2 - 2\eta_j \mathbb{E}(\eta_j|\mathcal{G}_j) + \mathbb{E}(\eta_j|\mathcal{G}_j)^2) + \\
&\quad + 2\mathbb{E}\sum_{j=1}^k \sum_{i<j} (\eta_i \eta_j - \eta_i \mathbb{E}(\eta_j|\mathcal{G}_j) - \eta_j \mathbb{E}(\eta_i|\mathcal{G}_i) + \mathbb{E}(\eta_i|\mathcal{G}_i)\mathbb{E}(\eta_j|\mathcal{G}_j)) \\
&= \sum_{j=1}^k \left(\mathbb{E}\eta_j^2 - \mathbb{E}(\mathbb{E}(\eta_j|\mathcal{G}_j))^2\right).
\end{aligned}$$

Hence the statement is proved. □

4.3 Probability Measures

Let $(E, \|\cdot\|_E)$ be a real Banach space. A subset of E of the form

$$\{x \in E : (\langle x, y_1' \rangle, \langle x, y_2' \rangle, \dots, \langle x, y_n' \rangle) \in A\}$$

where $n \in \mathbb{N}$, $y_i' \in E'$ for $i = 1, \dots, n$ and $A \in \mathcal{B}(\mathbb{R}^n)$ is called *cylindrical.* Cylindrical sets form a π-system. If two probability measures are identical on cylindrical sets, they are equal on $\mathcal{B}(E)$ by Proposition 4.1. If μ is a probability measure on $(E, \mathcal{B}(E))$, the function φ_μ on E'

$$\varphi_\mu(y') := \int_E e^{i\langle x, y' \rangle}\, \mu(dx)$$

for all $y' \in E'$ is called the *characteristic function* of μ. Instead of φ_μ we also denote it by $\hat{\mu}$. If $(H, (\cdot,\cdot)_H)$ is a real Hilbert space, φ_μ is regarded as a function on H and

$$\varphi_\mu(h) := \int_E e^{i(x,h)_H}\, \mu(dx)$$

for all $h \in H$. Probability measures on a real Banach space are uniquely determined by their characteristic functions.

Lemma 4.12. *If μ and ν are probability measures on $(E, \mathcal{B}(E))$ such that*

$$\varphi_\mu(y') = \varphi_\nu(y')$$

for all $y' \in E'$, then $\mu = \nu$.

Proof. The claim is true if $E = \mathbb{R}^n$ [3, pp. 333–334]. In the general case we fix $n \in \mathbb{N}$, $y_1', \ldots, y_n' \in E'$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. By the hypothesis

$$\int_E e^{i\lambda_1\langle x,y_1'\rangle + \ldots + i\lambda_n\langle x,y_n'\rangle}\, \mu(dx) = \int_E e^{i\lambda_1\langle x,y_1'\rangle + \ldots + i\lambda_n\langle x,y_n'\rangle}\, \nu(dx). \tag{4.1}$$

Identity (4.1) implies that the $\mathbb{R}^n$-valued mapping $x \mapsto (\langle x, y_1'\rangle, \ldots, \langle x, y_n'\rangle)$ maps the measures μ and ν onto measures $\tilde{\mu}$ and $\tilde{\nu}$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ with identical characteristic functions where $\tilde{\mu}(A) = \mu(x \in E : (\langle x, y_1'\rangle, \ldots, \langle x, y_n'\rangle) \in A)$ for all $A \in \mathcal{B}(\mathbb{R}^n)$. Hence the measures $\tilde{\mu}$ and $\tilde{\nu}$ are identical. But this implies that the measures μ and ν are equal on all cylindrical sets. Therefore $\mu = \nu$ on $\mathcal{B}(E)$. $\square$

Next we present some lemmas concerning probability measures that will be used later in this chapter. By the uniqueness of the characteristic function we are able to prove the following lemma. The version in $\mathbb{R}^n$ can be found from the book of Karatzas and Shreve [20, Lemma 2.6.13].

Lemma 4.13. *Let $(H, (\cdot,\cdot)_H)$ be a real separable Hilbert space and X an H-valued random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $\mathcal{G}$ is a sub-σ-algebra of $\mathcal{F}$ and that for each $\omega \in \Omega$ there exists a function $\varphi(\cdot;\omega) : H \to \mathbb{C}$ such that*

$$\varphi(h;\omega) = \mathbb{E}\left[e^{i(X,h)_H} \middle| \mathcal{G}\right](\omega)$$

for all $h \in H$ and almost all $\omega \in \Omega$. If $\varphi(\cdot;\omega)$ is the characteristic function of some probability measure μ_ω on $(H, \mathcal{B}(H))$ for each $\omega \in \Omega$, i.e.,

$$\varphi(h;\omega) = \int_H e^{i(x,h)_H} \mu_\omega(dx)$$

for all $h \in H$,

$$\mathbb{P}[X \in B|\mathcal{G}](\omega) = \mu_\omega(B)$$

for each $B \in \mathcal{B}(H)$ and almost all $\omega \in \Omega$.

Proof. Since for all $A \in \mathcal{G}$ and $B \in \mathcal{B}(H)$

$$\begin{aligned}\mathbb{E}[\mathbb{P}(X \in B|\mathcal{G})\chi_A] &= \mathbb{E}[\mathbb{E}(\chi_{\{X\in B\}}|\mathcal{G})\chi_A] = \mathbb{E}(\chi_{\{X\in B\}}\chi_A) \\ &= \mathbb{P}(\{X \in B\} \cap A) = \mathcal{L}(X)(B \cap X(A))\end{aligned}$$

and $\mathbb{P}[X \in \cdot\, |\mathcal{G}](\omega)$ is a probability measure on $(H, \mathcal{B}(H))$ for all almost $\omega \in \Omega$ [53, Theorem 9.7],

$$\int_A \int_B \mathbb{P}[X \in dx|\mathcal{G}](\omega)\, \mathbb{P}(d\omega) = \int_{B \cap X(A)} \mathcal{L}(X)(dx)$$

for all $A \in \mathcal{G}$ and $B \in \mathcal{B}(H)$. Thus

$$\begin{aligned} \int_A \int_H e^{i(x,h)_H}\, \mathbb{P}[X \in dx|\mathcal{G}](\omega)\, \mathbb{P}(d\omega) &= \int_{X(A)} e^{i(x,h)_H}\, \mathcal{L}(X)(dx) \\ &= \int_A e^{i(X,h)_H}\, d\mathbb{P} \end{aligned}$$

for all $A \in \mathcal{G}$ and $h \in H$. Hence

$$\mathbb{E}\left[e^{i(X,h)_H}\big|\mathcal{G}\right](\omega) = \int_H e^{i(x,h)_H}\, \mathbb{P}[X \in dx|\mathcal{G}](\omega) \tag{4.2}$$

for all $h \in H$ and almost all $\omega \in \Omega$. The set of ω for which Equality (4.2) fails may depend on h. We can choose a countable dense subset D of H and an event $\tilde{\Omega} \in \mathcal{F}$ with $\mathbb{P}(\tilde{\Omega}) = 1$ such that Equality (4.2) holds for every $\omega \in \tilde{\Omega}$ and $h \in D$. The continuity in h of both side of Equality (4.2) allows us to conclude its validity for every $\omega \in \tilde{\Omega}$ and $h \in H$ [53, Theorem 9.7]. Since a probability measure on $(H, \mathcal{B}(H))$ is uniquely determined by its characteristic function, $\mu_\omega = \mathbb{P}[X \in \cdot\, |\mathcal{G}](\omega)$ for almost all $\omega \in \Omega$. Thus the result follows. □

In Subsection 4.5.1 we shall need the previous lemma to show properties of the stochastic integral. The following corollary of Lemma 4.13 is used to prove some properties of the Hilbert space valued Wiener process in Subsection 4.4.3.

Corollary 4.14. *Let $(H, (\cdot,\cdot)_H)$ be a real separable Hilbert space and X an H-valued random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $\mathcal{G}_i$, $i = 1, 2$, are sub-σ-algebras of $\mathcal{F}$ and*

$$\mathbb{E}\left(e^{i(X,h)_H}\big|\mathcal{G}_1\right) = \mathbb{E}\left(e^{i(X,h)_H}\big|\mathcal{G}_2\right)$$

almost surely for all $h \in H$. If for each $\omega \in \Omega$ there exists a function $\varphi(\cdot;\omega) : H \to \mathbb{C}$ such that

$$\varphi(h;\omega) = \mathbb{E}\left[e^{i(X,h)_H}\big|\mathcal{G}_1\right](\omega)$$

for all $h \in H$ and almost all $\omega \in \Omega$ and $\varphi(\cdot;\omega)$ is the characteristic function of some probability measure μ_ω on $(H, \mathcal{B}(H))$ for each $\omega \in \Omega$, i.e.,

$$\varphi(h;\omega) = \int_H e^{i(x,h)_H} \mu_\omega(dx)$$

for all $h \in H$,

$$\mathbb{E}(f(X)|\mathcal{G}_1) = \mathbb{E}(f(X)|\mathcal{G}_2)$$

almost surely for all measurable function f from H to $\mathbb{C}$ such that f is integrable with respect to the measure $\mathcal{L}(X)$.

Proof. By the proof of Lemma 4.13,

$$\mathbb{E}\left[f(X)\middle|\mathcal{G}_i\right](\omega) = \int_H f(x)\, \mathbb{P}[X \in dx|\mathcal{G}_i](\omega)$$

for all $f \in L^1(H, \mathcal{B}(H), \mathcal{L}(X); \mathbb{C})$, almost all $\omega \in \Omega$ and $i = 1, 2$. Since according to Lemma 4.13,

$$\mathbb{P}[X \in B|\mathcal{G}_1](\omega) = \mu_\omega(B) = \mathbb{P}[X \in B|\mathcal{G}_2](\omega)$$

for each $B \in \mathcal{B}(H)$ and almost all $\omega \in \Omega$,

$$\mathbb{E}\left(f(X)\middle|\mathcal{G}_1\right) = \mathbb{E}\left(f(X)\middle|\mathcal{G}_2\right)$$

almost surely for all $f \in L^1(H, \mathcal{B}(H), \mathcal{L}(X); \mathbb{C})$. □

In Subsection 4.3.1 we use the following lemma to define the characteristic function of a Gaussian measure on a real Hilbert space.

Lemma 4.15. *Let $(H, (\cdot,\cdot)_H)$ be a real Hilbert space and ν a probability measure on $(H, \mathcal{B}(H))$. If for some $k \in \mathbb{N}$*

$$\int_H |(x,h)_H|^k\, \nu(dx) < \infty$$

for all $h \in H$, the transformation from H^k to $\mathbb{R}$

$$(h_1, \dots, h_k) \mapsto \int_H (x, h_1)_H \cdots (x, h_k)_H\, \nu(dx)$$

is a bounded symmetric k-linear form.

Proof. The transformation is obviously symmetric and k-linear. We define for each $n \in \mathbb{N}$ the set U_n by

$$U_n := \left\{ h \in H : \int_H |(x,h)_H|^k\, \nu(dx) \leq n \right\}.$$

By the hypothesis $H = \cup_{n=1}^\infty U_n$. Since H is a Hilbert space, by Baire's category theorem there exist $n_0 \in \mathbb{N}$, $h_0 \in U_{n_0}$ and $r_0 > 0$ such that $B(h_0, r_0) \subset U_{n_0}$. Hence

$$\int_H |(x, h_0 + y)_H|^k\, \nu(dx) \leq n_0$$

for all $y \in B(0, r_0)$. But for each $y \in B(0, r_0)$

$$\begin{aligned}\int_H |(x,y)_H|^k\, \nu(dx) &\leq 2^{k-1} \int_H |(x, h_0 + y)_H|^k\, \nu(dx) + 2^{k-1} \int_H |(x, h_0)_H|^k\, \nu(dx) \\ &\leq 2^k n_0.\end{aligned}$$

For all $h \in H$ different for zero $y = \frac{r_0 h}{2\|h\|_H} \in B(0, r_0)$. Hence

$$\int_H |(x,h)_H|^k\, \nu(dx) \leq 2^{2k} n_0 \|h\|_H^k r_0^{-k}$$

for all $h \in H$. By the generalized Hölder inequality the transformation is bounded since

$$\begin{aligned}
&\left| \int_H (x,h_1)_H \cdots (x,h_k)_H \, \nu(dx) \right| \\
&\leq \left(\int_H |(x,h_1)_H|^k \, \nu(dx) \right)^{\frac{1}{k}} \cdots \left(\int_H |(x,h_k)_H|^k \, \nu(dx) \right)^{\frac{1}{k}} \\
&\leq 2^{2k} n_0 r_0^{-k} \|h_1\|_H \cdots \|h_k\|_H
\end{aligned}$$

for all $(h_1, \dots, h_k) \in H^k$. □

4.3.1 Gaussian Measures

Let Q be a positive definite symmetric $n \times n$ matrix and m a vector in $\mathbb{R}^n$. The function

$$\pi_{m,Q}(x) := \frac{1}{(2\pi \det Q)^{n/2}} \; e^{-\frac{1}{2}(x-m)^T Q^{-1}(x-m)}$$

for all $x \in \mathbb{R}^n$ is the density of a probability measure on $\mathbb{R}^n$ called the *non-degenerate Gaussian distribution* and denoted by $\mathcal{N}(m,Q)$. Its characteristic function is of the form

$$\widehat{\mathcal{N}}(m,Q)(\lambda) = e^{im^T\lambda} e^{-\frac{1}{2}\lambda^T Q \lambda} \tag{4.3}$$

for each $\lambda \in \mathbb{R}^n$. The general *Gaussian distribution* on $\mathbb{R}^n$ is an image of a non-degenerate Gaussian distribution under a linear mapping. Its characteristic function is of the form (4.3) and is determined by $m \in \mathbb{R}^n$ and a non-negative symmetric matrix Q. If $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is a random variable and $\mathcal{L}(X) = \mathcal{N}(m,Q)$, then X is said to be a *Gaussian random variable* and $\mathbb{E}X = m$ and $\mathrm{Cov}(X) = Q$. If $m = 0$, the measure $\mathcal{N}(0,Q)$ is symmetric in the sense that it associates the same value on the sets which are symmetric with respect to the origin.

Let $(E, \|\cdot\|_E)$ be a real Banach space. A probability measure μ on $(E, \mathcal{B}(E))$ is said to be a *Gaussian measure* if and only if the law of an arbitrary continuous linear functional considered as a random variable on $(E, \mathcal{B}(E), \mu)$ is a Gaussian measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. If the law of each $\varphi \in E'$ is in addition symmetric (zero mean) Gaussian distribution on $\mathbb{R}$, then μ is called a *symmetric Gaussian measure.* A random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{B}(E))$ is *Gaussian* if its law $\mathcal{L}(X)$ is a Gaussian measure on $(E, \mathcal{B}(E))$. Hence X is a Gaussian random variable if and only if $\langle X, \varphi \rangle$ is a real valued Gaussian random variable for all $\varphi \in E'$ since

$$\begin{aligned}
\mathcal{L}(\langle \cdot, \varphi \rangle)(A) &= \mathcal{L}(X)(x \in E : \langle x, \varphi \rangle \in A) \\
&= \mathbb{P}(\omega \in \Omega : X(\omega) \in \{x \in E : \langle x, \varphi \rangle \in A\}) \\
&= \mathbb{P}(\omega \in \Omega : \langle X(\omega), \varphi \rangle \in A)
\end{aligned}$$

for all $A \in \mathcal{B}(\mathbb{R})$.

Let $(H, (\cdot,\cdot)_H)$ be a real Hilbert space. A probability measure μ on $(H, \mathcal{B}(H))$ is *Gaussian* if for each $h \in H$ there exist $m_h \in \mathbb{R}$ and $q_h \geq 0$ such that

$$\mathcal{L}((\cdot, h)_H)(A) = \mu(x \in H : (x,h)_H \in A) = \mathcal{N}(m_h, q_h)(A)$$

for all $A \in \mathcal{B}(\mathbb{R})$. If μ is a Gaussian measure, the functionals

$$\begin{aligned} H \to \mathbb{R}, &\qquad h \mapsto \int_H (x,h)_H \, \mu(dx), \\ H \times H \to \mathbb{R}, &\qquad (h_1,h_2) \mapsto \int_H (x,h_1)_H (x,h_2)_H \, \mu(dx) \end{aligned}$$

are well defined since

$$\int_H (x,h)_H \, \mu(dx) = m_h$$

and

$$\begin{aligned} \left| \int_H (x,h_1)_H (x,h_2)_H \, \mu(dx) \right| &\le \left(\int_H (x,h_1)_H^2 \, \mu(dx) \right)^{\frac{1}{2}} \left(\int_H (x,h_2)_H^2 \, \mu(dx) \right)^{\frac{1}{2}} \\ &= \left(q_{h_1} + m_{h_1}^2 \right)^{\frac{1}{2}} \left(q_{h_2} + m_{h_2}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

The first functional is linear and the second one is bilinear. According to Lemma 4.15 they are also bounded and symmetric. By the Riesz representation theorem and the Lax-Milgram lemma there exist an element $m \in H$ and a bounded linear operator Q such that

$$\int_H (x,h)_H \, \mu(dx) = (h,m)_H$$

for all $h \in H$ and

$$\int_H (x,h_1)_H (x,h_2)_H \, \mu(dx) - (h_1,m)_H (h_2,m)_H = (Qh_1,h_2)_H$$

for all $h_1, h_2 \in H$. The operator Q is non-negative and self-adjoint since

$$\begin{aligned} (Qh,h)_H &= \int_H (x,h)_H^2 \, \mu(dx) - \left(\int_H (x,h)_H \, \mu(dx) \right)^2 \\ &\ge \int_H (x,h)_H^2 \, \mu(dx) - \int_H (x,h)_H^2 \, \mu(dx) = 0 \end{aligned}$$

for all $h \in H$ by Jensen's inequality and

$$(Qh_1,h_2)_H = (Qh_2,h_1)_H = (h_1,Qh_2)_H$$

for all $h_1, h_2 \in H$. The element m is called the *mean* and the operator Q the *covariance operator* of μ. A Gaussian measure μ on H with mean m and covariance Q has the characteristic function

$$\hat{\mu}(h) = \int_H e^{i(x,h)_H} \, \mu(dx) = \int_{\mathbb{R}} e^{it} \mathcal{L}((\cdot,h)_H)(dt) = e^{i(h,m)_H - \frac{1}{2}(Qh,h)_H}$$

for all $h \in H$. Therefore μ is uniquely determined by m and Q. It is also denoted by $\mathcal{N}(m,Q)$.

If $(H,(\cdot,\cdot)_H)$ is a real separable Hilbert space, the covariance operator of a Gaussian measure is nuclear.

Proposition 4.16. *Let $(H,(\cdot,\cdot)_H)$ be a real separable Hilbert space and μ a Gaussian measure with mean 0 and covariance Q. Then Q is a trace class operator.*

Proof. We consider the characteristic function of the measure μ

$$\hat{\mu}(h) = \int_H e^{i(x,h)_H}\ \mu(dx) = e^{-\frac{1}{2}(Qh,h)_H}$$

for all $h \in H$. Since $\hat{\mu}(h) \in \mathbb{R}$, for all $h \in H$ and each $c > 0$

$$\begin{aligned} 1 - \hat{\mu}(h) &= \int_H (1 - \cos(x,h)_H)\ \mu(dx) \\ &\le \frac{1}{2}\int_{\|x\|_H \le c} (x,h)_H^2\ \mu(dx) + 2\mu(x \in H : \|x\|_H > c) \\ &= \frac{1}{2}(Q_c h, h)_H + 2\mu(x \in H : \|x\|_H > c) \end{aligned}$$

where Q_c is the bounded linear operator defined by

$$(Q_c h_1, h_2)_H = \int_{\|x\|_H \le c} (x, h_1)_H (x, h_2)_H\ \mu(dx)$$

for all $h_1, h_2 \in H$. Let $h \in H$ be such that $(Q_c h, h)_H \le 1$. Then

$$\begin{aligned} e^{-\frac{1}{2}(Qh,h)_H} &\ge 1 - \frac{1}{2}(Q_c h, h)_H - 2\mu(x \in H : \|x\|_H > c) \\ &\ge \frac{1}{2} - 2\mu(x \in H : \|x\|_H > c). \end{aligned}$$

We choose c such that

$$\mu(x \in H : \|x\|_H > c) < \frac{1}{4}.$$

Then

$$(Qh, h)_H \le -2\log\left(\frac{1}{2} - 2\mu(x \in H : \|x\|_H > c)\right) := \beta.$$

Let $h \in H$. We denote $\alpha_h := (Q_c h, h)_H$. Then $\alpha_h \ge 0$. Hence for all $h \in H$ such that $\alpha_h \neq 0$

$$\left(Q_c \frac{h}{\sqrt{\alpha_h}}, \frac{h}{\sqrt{\alpha_h}}\right)_H = 1.$$

Thus

$$\left(Q \frac{h}{\sqrt{\alpha_h}}, \frac{h}{\sqrt{\alpha_h}}\right)_H \le \beta$$

for all $h \in H$ such that $\alpha_h \neq 0$. If $(Q_c h, h)_H = 0$ for some $h \in H$, then $(Qh, h)_H = 0$. Therefore $(Qh, h)_H \le \beta (Q_c h, h)_H$ for all $h \in H$. The operator Q_c is a trace class operator by Proposition D.14 since Q_c is non-negative and self-adjoint and

$$\operatorname{Tr} Q_c = \sum_{n=1}^{\infty} (Q_c e_n, e_n)_H = \int_{\|x\|_H \le c} \sum_{n=1}^{\infty} (x, e_n)_H^2\ \mu(dx) = \int_{\|x\|_H \le c} \|x\|_H^2\ \mu(dx) \le c^2$$

where $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis in H. Thus Q is a trace class operator because

$$\operatorname{Tr} Q = \sum_{n=1}^{\infty} (Q e_n, e_n)_H \le \sum_{n=1}^{\infty} \beta (Q_c e_n, e_n)_H = \beta \operatorname{Tr} Q_c$$

and Q is non-negative and self-adjoint. □

The following proposition shows that there exist Gaussian measures in a real separable Hilbert space.

Proposition 4.17. *Let $(H,(\cdot,\cdot)_H)$ be a real separable Hilbert space, $m \in H$ and Q be a positive self-adjoint trace class operator in H with $\operatorname{Ker} Q = \{0\}$. Then there exists a Gaussian measure with mean m and covariance Q.*

Proof. Since Q is nuclear, by Proposition D.9 it is compact. Since Q is a compact self-adjoint operator with $\operatorname{Ker} Q = \{0\}$, the normalized eigenvectors $\{e_k\}_{k=1}^\infty$ form an orthonormal basis in H [14, Theorem 5.1, Observation 6.1.b, pp. 113–116]. We denote by $\{\lambda_k\}_{k=1}^\infty$ the corresponding set of eigenvalues of Q. Then $\operatorname{Tr} Q = \sum_{k=1}^\infty \lambda_k < \infty$ by Proposition D.14. Let $\{\xi_k\}_{k=1}^\infty$ be a sequence of independent real $\mathcal{N}(0,1)$-random variables on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ [3, Theorem 20.4]. We set

$$X := m + \sum_{k=1}^\infty \sqrt{\lambda_k}\xi_k e_k. \tag{4.4}$$

Let $i,j \in \mathbb{N}$ and $j \geq i$. Then

$$\begin{aligned}
&\mathbb{E}\left\|\sum_{k=1}^{j} \sqrt{\lambda_k}\xi_k e_k - \sum_{k=1}^{i} \sqrt{\lambda_k}\xi_k e_k\right\|_H^2 \\
&= \mathbb{E}\left(\sum_{k=i+1}^{j} \sqrt{\lambda_k}\xi_k e_k, \sum_{k=i+1}^{j} \sqrt{\lambda_k}\xi_k e_k\right)_H \\
&= \mathbb{E}\sum_{k=i+1}^{j} \lambda_k\xi_k^2 = \sum_{k=i+1}^{j} \lambda_k \leq \sum_{k=i+1}^{\infty} \lambda_k \longrightarrow 0
\end{aligned}$$

as $i \to \infty$. Hence the series on the right hand side of Definition (4.4) converges in $L^2(\Omega,\mathcal{F},\mathbb{P};H)$. Therefore $X \in L^2(\Omega,\mathcal{F},\mathbb{P};H)$.

We prove that the law $\mathcal{L}(X)$ of the random variable X is Gaussian, i.e., for $h \in H$

$$\mathcal{L}((\cdot,h)_H)(A) = \mathcal{L}(X)(x \in H : (x,h)_H \in A) = \mathcal{N}(c_h, q_h)(A)$$

for all $A \in \mathcal{B}(\mathbb{R})$ with some $c_h \in \mathbb{R}$ and $q_h > 0$. Let $h \in H$ be fixed. We show that $(\cdot,h)_H$ is a Gaussian random variable from $(H,\mathcal{B}(H),\mathcal{L}(X))$ to $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. We use the characteristic function. Let $\tau \in \mathbb{R}$. Then

$$\varphi_{\mathcal{L}((\cdot,h)_H)}(\tau) = \int_{\mathbb{R}} e^{i\tau t}\, \mathcal{L}((\cdot,h)_H)(dt) = \int_H e^{i\tau(x,h)_H}\, \mathcal{L}(X)(dx) = \int_\Omega e^{i\tau(X,h)_H}\, d\mathbb{P}.$$

By Lebesgue's dominated convergence theorem,

$$\begin{aligned}
\varphi_{\mathcal{L}((\cdot,h)_H)}(\tau) &= e^{i\tau(m,h)_H}\mathbb{E}\left[\exp\left(\sum_{k=1}^\infty i\tau\sqrt{\lambda_k}\xi_k(e_k,h)_H\right)\right] \\
&= e^{i\tau(m,h)_H}\mathbb{E}\prod_{k=1}^\infty \exp\left(i\tau\sqrt{\lambda_k}\xi_k(e_k,h)_H\right) \\
&= e^{i\tau(m,h)_H}\lim_{n\to\infty}\mathbb{E}\prod_{k=1}^n \exp\left(i\tau\sqrt{\lambda_k}\xi_k(e_k,h)_H\right).
\end{aligned}$$

Since ξ_k are independent and $\xi_k \sim \mathcal{N}(0,1)$,

$$\begin{aligned}
\varphi_{\mathcal{L}((\cdot,h)_H)}(\tau) &= e^{i\tau(m,h)_H} \lim_{n\to\infty} \prod_{k=1}^{n} \mathbb{E}\left[\exp\left(i\tau\sqrt{\lambda_k}(e_k,h)_H\xi_k\right)\right] \\
&= e^{i\tau(m,h)_H} \prod_{k=1}^{\infty} \exp\left(-\frac{1}{2}\lambda_k(e_k,h)_H^2\tau^2\right) \\
&= \exp\left(i\tau(m,h)_H - \frac{1}{2}\sum_{k=1}^{\infty}\lambda_k(e_k,h)_H^2\tau^2\right) \\
&= e^{i\tau(m,h)_H - \frac{1}{2}\tau^2(Qh,h)_H}.
\end{aligned}$$

Thus $\mathcal{L}((\cdot,h)_H) = \mathcal{N}((m,h)_H,(Qh,h)_H)$. Hence $(\cdot,h)_H$ is a Gaussian random variable for all $h \in H$. Therefore $\mathcal{L}(X)$ is a Gaussian measure in H. Furthermore, for all $h \in H$

$$\int_H (x,h)_H\; \mathcal{L}(X)(dx) = \int_\Omega (X,h)_H\; d\mathbb{P} = \left(\int_\Omega X\; d\mathbb{P}, h\right)_H = (\mathbb{E}(X),h)_H.$$

Since $\mathbb{E}X = m$, the mean of $\mathcal{L}(X)$ is m. Since for all $h_1,h_2 \in H$

$$\begin{aligned}
&\int_H (x,h_1)_H(x,h_2)_H\; \mathcal{L}(X)(dx) - (h_1,m)_H(h_2,m)_H \\
&= \int_\Omega (X,h_1)_H(X,h_2)_H\; d\mathbb{P} - (h_1,\mathbb{E}(X))_H(h_2,\mathbb{E}(X))_H \\
&= \int_\Omega (h_1, X-\mathbb{E}(X))_H(h_2,X-\mathbb{E}(X))_H\; d\mathbb{P} \\
&= (\mathrm{Cov}(X)h_1,h_2)_H
\end{aligned}$$

and by Lebesgue's dominated convergence theorem,

$$\begin{aligned}
(\mathrm{Cov}(X)h_1,h_2)_H &= \mathbb{E}\left(\sum_{k=1}^{\infty}\sqrt{\lambda_k}\xi_k(h_1,e_k)_H\sum_{k=1}^{\infty}\sqrt{\lambda_k}\xi_k(h_2,e_k)_H\right) \\
&= \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\sqrt{\lambda_k}\sqrt{\lambda_l}(h_1,e_k)_H(h_2,e_l)_H\mathbb{E}(\xi_k\xi_l) \\
&= \sum_{k=1}^{\infty}\lambda_k(h_1,e_k)_H(h_2,e_k)_H \\
&= (Qh_1,h_2)_H,
\end{aligned}$$

the covariance of $\mathcal{L}(X)$ is Q. □

Remark 4.18 If $\mathcal{L}(X)$ is a Gaussian measure with mean m and covariance Q, then $\mathbb{E}X = m$ and $\mathrm{Cov}\, X = Q$.

4.4 Stochastic Processes

Let $(E, \|\cdot\|_E)$ be a separable Banach space, $(\Omega,\mathcal{F},\mathbb{P})$ a probability space and I an interval in $\mathbb{R}$. A family $X = \{X(t)\}_{t\in I}$ of E-valued random variables defined on Ω

is called a *stochastic process*. The definition of stochastic processes does not assume anything about the behaviour of processes with respect to the index t. However, it is appropriate to be interested in X as a function of t, as well. We set $X(t,\omega) := X(t)(\omega)$ for all $t \in I$ and $\omega \in \Omega$. The function $X(\cdot,\omega)$ for a fixed $\omega \in \Omega$ is called a *trajectory* of X. In the following definition it has been gathered some measurability and continuity properties of stochastic processes with respect to the index t.

Definition 4.19. *Let X be an E-valued stochastic process. Then*

(i) X is measurable if the mapping $X : I \times \Omega \to E$ is $\mathcal{B}(I) \times \mathcal{F}$-measurable,

(ii) X is stochastically continuous at $t_0 \in I$ if for all $\varepsilon > 0$ and $\delta > 0$ there exists $\rho > 0$ such that

$$\mathbb{P}(\|X(t) - X(t_0)\|_E \geq \varepsilon) \leq \delta$$

for all $t \in [t_0 - \rho, t_0 + \rho] \cap I$,

(iii) X is stochastically continuous on I if it is stochastically continuous at every point in I,

(iv) X is uniformly stochastically continuous on I if for all $\varepsilon > 0$ and $\delta > 0$ there exists $\rho > 0$ such that

$$\mathbb{P}(\|X(t) - X(s)\|_E \geq \varepsilon) \leq \delta$$

for all $s, t \in I$ such that $|t - s| \leq \rho$,

(v) X is mean square continuous at $t_0 \in I$ if

$$\lim_{t \to t_0} \mathbb{E}\|X(t) - X(t_0)\|_E^2 = 0,$$

(vi) X is mean square continuous on I if it is mean square continuous at every point in I,

(vii) X is continuous (with probability 1) if its trajectories $X(\cdot,\omega)$ are continuous for almost all $\omega \in \Omega$.

In the following lemma we show the relation between stochastically and mean square continuous processes.

Lemma 4.20. *A mean square continuous process is stochastically continuous.*

Proof. Let X be mean square continuous on I and $t_0 \in I$. Let $\varepsilon > 0$ and $\delta > 0$. Then there exists $\rho > 0$ such that $\mathbb{E}\|X(t) - X(t_0)\|_E^2 < \varepsilon^2 \delta$ for all $t \in [t_0 - \rho, t_0 + \rho] \cap I$. By Tšebyšev's inequality,

$$\mathbb{P}(\|X(t) - X(t_0)\|_E \geq \varepsilon) \leq \frac{\mathbb{E}\|X(t) - X(t_0)\|_E^2}{\varepsilon^2} < \delta$$

for all $t \in [t_0 - \rho, t_0 + \rho] \cap I$. Thus X is stochastically continuous on I. □

The stochastical continuity is uniform if the interval is compact.

Lemma 4.21. *If I is a compact interval, a stochastically continuous process on I is uniformly stochastically continuous.*

Proof. Let X be a stochastically continuous process on I. Let $\varepsilon > 0$ and $\delta > 0$. Then for all $r \in I$ there exists a closed interval $[r - \rho_r, r + \rho_r]$ with $\rho_r > 0$ such that

$$\mathbb{P}\left(\|X(s) - X(r)\|_E \geq \frac{\varepsilon}{2}\right) \leq \frac{\delta}{2}$$

for all $s \in [r - \rho_r, r + \rho_r] \cap I$. Consequently, for all $s, t \in [r - \rho_r, r + \rho_r] \cap I$

$$\begin{aligned}\mathbb{P}(\|X(s) - X(t)\|_E \geq \varepsilon) &\leq \mathbb{P}\left(\|X(s) - X(r)\|_E \geq \frac{\varepsilon}{2} \text{ or } \|X(t) - X(r)\|_E \geq \frac{\varepsilon}{2}\right) \\ &\leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.\end{aligned}$$

Since the interval I is compact, there exists a finite family of intervals $[r_i - \rho_{r_i}/2, r_i + \rho_{r_i}/2]$ which covers I. The common ρ is then $\min_i (\rho_{r_i}/2)$. □

A stochastic process Y is called a *modification* or a *version* of X if

$$\mathbb{P}(\omega \in \Omega : X(t, \omega) \neq Y(t, \omega)) = 0$$

for all $t \in I$. If X is a stochastic process on I, then X needs not to be measurable in the product space $I \times \Omega$. If X is stochastically continuous on a compact interval, X has a measurable modification. We use the notation $\Omega_T := [0, T] \times \Omega$ for all $T > 0$. We mark with $\mathbb{P}_T$ the product measure of the Lebesgue measures on $[0, T]$ with the probability measure $\mathbb{P}$ for all $T > 0$.

Proposition 4.22. *Let $X(t)$, $t \in [0, T]$, be a stochastically continuous process. Then X has a measurable modification.*

Proof. A stochastically continuous process X on $[0, T]$ is uniformly stochastically continuous by Lemma 4.21. Thus for each positive integer m there exists a partition $0 = t_{m,0} < t_{m,1} < \ldots < t_{m,n(m)} = T$ such that for all $t \in (t_{m,k}, t_{m,k+1}]$

$$\mathbb{P}(\|X(t_{m,k}, \omega) - X(t, \omega)\|_E \geq 2^{-m}) \leq 2^{-m}$$

if $k = 0, 1, \ldots, n(m) - 1$. We define

$$X_m(t, \omega) := \begin{cases} X(0, \omega) & \text{if } t = 0, \\ X(t_{m,k}, \omega) & \text{if } t \in (t_{m,k}, t_{m,k+1}] \text{ and } k \leq n(m) - 1, \end{cases}$$

for all $t \in [0, T]$ and $\omega \in \Omega$. Since for all $B \in \mathcal{B}(E)$

$$\{(t, \omega) \in \Omega_T : X_m(t, \omega) \in B\} = \{0\} \times C_0 \cup \bigcup_{k=1}^{n(m)-1} (t_{m,k}, t_{m,k+1}] \times C_k$$

where $C_k \in \mathcal{F}$ for all $k = 0, 1, \ldots, n(m) - 1$, the process X_m is measurable with respect to the σ-algebra $\mathcal{B}([0, T]) \times \mathcal{F}$. We denote by A the set of all those $(t, \omega) \in \Omega_T$ for which the sequence $\{X_m(t, \omega)\}_{m=1}^{\infty}$ is convergent. Then $A \in \mathcal{B}([0, T]) \times \mathcal{F}$ since

$$A = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{m=N}^{\infty} \bigcap_{n=N}^{\infty} \left\{(t, \omega) \in \Omega_T : \|X_m(t, \omega) - X_n(t, \omega)\|_E \leq \frac{1}{k}\right\}.$$

Hence the process Y defined by

$$Y(t,\omega) := \lim_{m\to\infty} \chi_A(t,\omega)X_m(t,\omega)$$

is $\mathcal{B}([0,T]) \times \mathcal{F}$-measurable since by Corollary B.7 in a separable Banach space the weak measurability is equivalent to the measurability and therefore the limit of random variables is a random variable in a separable Banach space.

For a fixed $t \in [0,T]$ we denote

$$B_m := \{\omega \in \Omega : \|X_m(t,\omega) - X(t,\omega)\|_E \geq 2^{-m}\}.$$

Then $\mathbb{P}(B_m) \leq 2^{-m}$. Since

$$\sum_{m=1}^{\infty} \mathbb{P}(B_m) \leq \sum_{m=1}^{\infty} 2^{-m} = 1,$$

according to the Borel–Cantelli lemma $\mathbb{P}(\limsup B_m) = 0$. Hence for almost all $\omega \in \Omega$ there exists $m(\omega) > 0$ such that $\|X_n(t,\omega) - X(t,\omega)\|_E < 2^{-n}$ for all $n \geq m(\omega)$. Therefore $\mathbb{P}_T(A) = 1$ and $X_n(t,\omega)$ converges pointwise to $X(t,\omega)$ for all $t \in [0,T]$ and almost all $\omega \in \Omega$. Hence $X(t) = Y(t)$ almost surely for all $t \in [0,T]$ and the process Y is the required modification. □

4.4.1 Processes with Filtration

Let I be an interval. A family $\{\mathcal{F}_t\}_{t\in I}$ of σ-algebras $\mathcal{F}_t \subseteq \mathcal{F}$ is called a *filtration* if $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s,t \in I$ such that $s < t$. We denote by $\mathcal{F}_{t^+}$ the intersection of all $\mathcal{F}_s$ where $s > t$, i.e.,

$$\mathcal{F}_{t^+} := \bigcap_{s>t} \mathcal{F}_s.$$

Then $\mathcal{F}_{t^+}$ is a σ-algebra for all $t \in I$. The family $\{\mathcal{F}_t\}_{t\in I}$ is said to be *right-continuous* if $\mathcal{F}_t = \mathcal{F}_{t^+}$ for all $t \in I$. The filtration $\{\mathcal{F}_t\}_{t\in I}$ is called *normal* if it is right-continuous and $\mathcal{F}_0$ contains all $A \in \mathcal{F}$ such that $\mathbb{P}(A) = 0$.

If the random variable $X(t)$ is $\mathcal{F}_t$-measurable for all $t \in I$, the process X is said to be *adapted* (to the filtration $\{\mathcal{F}_t\}_{t\in I}$). If $X(t)$, $t \in I$, is a stochastic process, the filtration $\{\mathcal{F}_t^X\}_{t\in I}$ generated by the process X is defined by $\mathcal{F}_t^X := \sigma(X(s),\ s \leq t)$ for all $t \in I$. Every process is adapted to the filtration generated by its own history.

We denote the collection of $\mathbb{P}$-*null sets* by

$$\mathcal{N}^{\mathbb{P}} := \{A \subseteq \Omega : \text{there exists } B \in \mathcal{F} \text{ such that } A \subseteq B \text{ and } \mathbb{P}(B) = 0\}.$$

The *augmentation* $\{\mathcal{F}_t^{\mathbb{P}}\}_{t\in I}$ of the filtration $\{\mathcal{F}_t\}_{t\in I}$ is defined by $\mathcal{F}_t^{\mathbb{P}} := \sigma(\mathcal{F}_t \cup \mathcal{N}^{\mathbb{P}})$ for all $t \in I$. The augmentation is a filtration in $(\Omega, \mathcal{F}, \mathbb{P})$ if and only if the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete. In that case $\{\mathcal{F}_t^{\mathbb{P}}\}_{t\in I}$ is called the *augmented filtration*.

Let $T > 0$ and $\{\mathcal{F}_t\}_{t\in[0,T]}$ be a filtration. We denote by $\mathcal{P}_T$ the σ-algebra on Ω_T generated by sets of the form

$$\begin{cases} (s,t] \times F & \text{where } 0 \leq s < t \leq T \text{ and } F \in \mathcal{F}_s, \\ \{0\} \times F & \text{where } F \in \mathcal{F}_0. \end{cases} \tag{4.5}$$

Then $\mathcal{P}_T \subset \mathcal{B}([0,T]) \times \mathcal{F}$. Hence $(\Omega_T, \mathcal{P}_T, \mathbb{P}_T)$ is a measure space. The σ-algebra $\mathcal{P}_T$ is said to be the *predictable* σ-algebra and its element are called the *predictable sets*.

Lemma 4.23. *Let $T > 0$ and A be a predictable subset of Ω_T. Then for all $\varepsilon > 0$ there exists a finite union Γ of disjoint sets of the form (4.5) such that*

$$\mathbb{P}_T((A \setminus \Gamma) \cup (\Gamma \setminus A)) < \varepsilon.$$

Proof. Let $\mathcal{K}$ denote the family of all finite unions of disjoint sets of the form (4.5). Then $\mathcal{K}$ is closed under finite unions and intersections and the complement since

$$\begin{aligned}
&((s_1, t_1] \times F_1) \cap ((s_2, t_2] \times F_2) \\
&= \begin{cases} (s_1 \vee s_2, t_1 \wedge t_2] \times (F_1 \cap F_2) & \text{if } s_1 \vee s_2 \leq t_1 \wedge t_2, \\ \emptyset & \text{if } s_1 \vee s_2 > t_1 \wedge t_2, \end{cases}
\end{aligned}$$

for $s_1 < s_2$

$$\begin{aligned}
&((s_1, t_1] \times F_1) \cup ((s_2, t_2] \times F_2) \\
&= \begin{cases} ((s_1, t_1] \times F_1) \cup ((s_2, t_2] \times F_2) & \text{if } t_1 \leq s_2, \\ ((s_1, s_2] \times F_1) \cup ((s_2, t_1] \times F_1 \cup F_2) \cup ((t_1, t_2] \times F_2) & \text{if } s_2 < t_1 \leq t_2, \\ ((s_1, s_2] \times F_1) \cup ((s_2, t_2] \times F_1 \cup F_2) \cup ((t_2, t_1] \times F_1) & \text{if } s_2 < t_2 < t_1 \end{cases}
\end{aligned}$$

and

$$((s,t] \times F)^c = (\{0\} \times \Omega) \cup ((0, s] \times \Omega) \cup ((s,t] \times (\Omega \setminus F)) \cup ((t, T] \times \Omega).$$

Thus $\mathcal{K}$ is a π-system. Let $\mathcal{G}$ be the family of such $A \in \mathcal{P}_T$ that for all $\varepsilon > 0$ there exists $\Gamma \in \mathcal{K}$ such that $\mathbb{P}_T((A \setminus \Gamma) \cup (\Gamma \setminus A)) < \varepsilon$. Then $\mathcal{K} \subset \mathcal{G}$. Let $A \in \mathcal{G}$ and $\varepsilon > 0$. Then there exists $\Gamma \in \mathcal{K}$ such that $\mathbb{P}_T((A \setminus \Gamma) \cup (\Gamma \setminus A)) < \varepsilon$. Thus $\mathbb{P}_T((A^c \setminus \Gamma^c) \cup (\Gamma^c \setminus A^c)) < \varepsilon$ since

$$(A \setminus \Gamma) \cup (\Gamma \setminus A) = (A \cap \Gamma^c) \cup (\Gamma \cap A^c) = (\Gamma^c \setminus A^c) \cup (A^c \setminus \Gamma^c).$$

Hence $A^c \in \mathcal{G}$ because $\Gamma^c \in \mathcal{K}$. Let $A_i \in \mathcal{G}$ for all $i \in \mathbb{N}$ be such that $A_i \cap A_j = \emptyset$ for all $i \neq j$ and $\varepsilon > 0$. Then there exist $\Gamma_i \in \mathcal{K}$ for all $i \in \mathbb{N}$ such that

$$\mathbb{P}_T((A_i \setminus \Gamma_i) \cup (\Gamma_i \setminus A_i)) < \frac{\varepsilon}{2^{i+1}}.$$

Let $m \in \mathbb{N}$ be such that

$$\sum_{i=m+1}^{\infty} \mathbb{P}_T(A_i) < \frac{\varepsilon}{2}.$$

Then $\cup_{i=1}^m \Gamma_i \in \mathcal{K}$,

$$\begin{aligned}
&\left(\bigcup_{i=1}^{\infty} A_i\right) \setminus \left(\bigcup_{i=1}^{m} \Gamma_i\right) \bigcup \left(\bigcup_{i=1}^{m} \Gamma_i\right) \setminus \left(\bigcup_{i=1}^{\infty} A_i\right) \\
&\subset \left[\left(\bigcup_{i=1}^{m} A_i\right) \setminus \left(\bigcup_{i=1}^{m} \Gamma_i\right) \bigcup \left(\bigcup_{i=1}^{m} \Gamma_i\right) \setminus \left(\bigcup_{i=1}^{m} A_i\right)\right] \bigcup \left(\bigcup_{i=m+1}^{\infty} A_i\right)
\end{aligned}$$

and

$$\left(\bigcup_{i=1}^{m} A_i\right) \setminus \left(\bigcup_{i=1}^{m} \Gamma_i\right) = \bigcup_{i=1}^{m} \left(A_i \cap \left(\bigcap_{i=1}^{m} \Gamma_i^c\right)\right) \subset \bigcup_{i=1}^{m} A_i \cap \Gamma_i^c = \bigcup_{i=1}^{m} A_i \setminus \Gamma_i.$$

Thus

$$\begin{aligned}
&\mathbb{P}_T\left(\left(\bigcup_{i=1}^{m} A_i\right) \setminus \left(\bigcup_{i=1}^{m} \Gamma_i\right) \bigcup \left(\bigcup_{i=1}^{m} \Gamma_i\right) \setminus \left(\bigcup_{i=1}^{m} A_i\right)\right) \\
&\leq \mathbb{P}_T\left(\bigcup_{i=1}^{m} (A_i \setminus \Gamma_i) \cup (\Gamma_i \setminus A_i)\right) \\
&\leq \sum_{i=1}^{m} \mathbb{P}_T((A_i \setminus \Gamma_i) \cup (\Gamma_i \setminus A_i)) \\
&< \sum_{i=1}^{m} \frac{\varepsilon}{2^{i+1}} < \frac{\varepsilon}{2}.
\end{aligned}$$

Hence

$$\mathbb{P}_T\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \setminus \left(\bigcup_{i=1}^{m} \Gamma_i\right) \bigcup \left(\bigcup_{i=1}^{m} \Gamma_i\right) \setminus \left(\bigcup_{i=1}^{\infty} A_i\right)\right) < \varepsilon.$$

Therefore $\cup_{i=1}^{\infty} A_i \in \mathcal{G}$. Thus $\mathcal{G} = \sigma(\mathcal{K}) = \mathcal{P}_T$ by Proposition 4.1. □

A measurable function from $(\Omega_T, \mathcal{P}_T, \mathbb{P}_T)$ to $(E, \mathcal{B}(E))$ is called a *predictable process.* A predictable process is necessarily an adapted one.

Proposition 4.24. *Let X be an adapted stochastically continuous process on the interval $[0, T]$. Then the process X has a predictable version on $[0, T]$.*

Proof. The proof is exactly the same as the one of Proposition 4.22. Since X is adapted, X_m is predictable. Hence the set A is a predictable set and the process Y is predictable. □

A E-valued stochastic process $X(t)$, $t \in [0, T]$, taking only a finite number of values is said to be *elementary* if there exist a sequence $0 = t_0 < t_1 < \ldots < t_k = T$ and a sequence $\{X_m\}_{m=0}^{k-1}$ of E-valued simple random variables such that X_m is $\mathcal{F}_{t_m}$-measurable and $X(t) = X_m$ if $t \in (t_m, t_{m+1}]$ for all $m = 0, 1, \ldots, k-1$. Elementary processes are a simple example of predictable processes. Actually, predictable processes can be approximated by elementary processes if they are integrable.

Proposition 4.25. *Let $X(t)$, $t \in [0, T]$, be an E-valued predictable process. If*

$$\mathbb{E} \int_0^T \|X(t)\|_E \, dt < \infty, \tag{4.6}$$

there exists a sequence $\{X_n\}_{n=1}^{\infty}$ of elementary processes such that

$$\mathbb{E} \int_0^T \|X(t) - X_n(t)\|_E \, dt \longrightarrow 0$$

as $n \to \infty$.

Proof. Let X be an E-valued predictable process satisfying Condition (4.6). Then by the Fubini theorem $X \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)$. According to Theorem B.12 there exists a sequence $\{X_n\}_{n=1}^{\infty}$ of simple E-valued predictable processes such that

$$\int_{\Omega_T} \|X(t,\omega) - X_n(t,\omega)\|_E \, \mathbb{P}_T(dt, d\omega) = \mathbb{E} \int_0^T \|X(t) - X_n(t)\|_E \, dt \longrightarrow 0$$

as $n \to \infty$. Since X_n is a simple E-valued $\mathcal{P}_T$-measurable function, it is of the form

$$X_n(t,\omega) = \sum_{l=1}^{m_n} X_n^l \chi_{A_n^l}(t,\omega)$$

for all $(t,\omega) \in \Omega_T$ where $m_n \in \mathbb{N}$ and $X_n^l \in E$ and $A_n^l \in \mathcal{P}_T$ for all $l = 1, \dots, m_n$ and $A_n^i \cap A_n^j = \emptyset$ if $i \neq j$. We denote $C := \sum_{l=1}^{m_n} \|X_n^l\|_E$. According to Lemma 4.23 for all $l = 1, \dots, m_n$ there exists a finite union Γ_n^l of disjoint sets of the form (4.5) such that

$$\mathbb{P}_T((A_n^l \setminus \Gamma_n^l) \cup (\Gamma_n^l \setminus A_n^l)) < \frac{1}{nCm_n}.$$

Then

$$Y_n(t,\omega) := \sum_{l=1}^{m_n} X_n^l \chi_{\Gamma_n^l}(t,\omega)$$

for all $(t,\omega) \in \Omega_T$ is an elementary process and

$$\begin{aligned}\mathbb{E} \int_0^T \|X_n(t,\omega) - Y_n(t,\omega)\|_E \, dt &\leq C\mathbb{P}_T \left(\bigcup_{l=1}^{m_n} (A_n^l \setminus \Gamma_n^l) \cup (\Gamma_n^l \setminus A_n^l) \right) \\ &\leq C \sum_{l=1}^{m_n} \mathbb{P}_T((A_n^l \setminus \Gamma_n^l) \cup (\Gamma_n^l \setminus A_n^l)) < \frac{1}{n}.\end{aligned}$$

Thus $\{Y_n\}_{n=1}^{\infty}$ is a sequence of elementary processes such that for every $\varepsilon > 0$

$$\|X - Y_n\|_{L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)} \leq \|X - X_n\|_{L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)} + \|X_n - Y_n\|_{L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)} < \varepsilon$$

if $n \in \mathbb{N}$ is so large that $\|X - X_n\|_{L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)} < \varepsilon/2$ and $n \geq 2/\varepsilon$. □

The integral of an integrable predictable process has a predictable version.

Lemma 4.26. *If $X \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)$, the process*

$$Y(t) := \begin{cases} \int_0^t X(s) \, ds & \text{if } \int_0^t \|X(s)\|_E \, ds < \infty, \\ 0 & \text{if } \int_0^t \|X(s)\|_E \, ds = \infty \end{cases}$$

on $[0,T]$ is continuous and has a predictable version.

Proof. Since $X \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)$, the trajectories of X are Bochner integrable almost surely. Hence $Y(t) = \int_0^t X(s) \, ds$ almost surely and $Y(t)$ is $\mathcal{F}_t$-measurable for all $t \in [0,T]$. Since for all $0 \leq s < t \leq T$

$$\|Y(t) - Y(s)\|_E = \left\| \int_0^t X(r) \, dr - \int_0^s X(r) \, dr \right\|_E \leq \int_0^T \chi_{[s,t]}(r) \|X(r)\|_E \, dr$$

almost surely and $X \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)$, the process Y is continuous by Lebesgue's dominated convergence theorem. Furthermore, by Lebesgue's dominated convergence theorem for all $0 \le s < t \le T$

$$\mathbb{E}\|Y(t) - Y(s)\|_E \le \mathbb{E}\int_0^T \chi_{[s,t]}(r)\|X(r)\|_E \, dr \longrightarrow 0$$

as $|t-s| \to 0$. Thus for all $\varepsilon > 0$ and $\delta > 0$ there exists $\rho > 0$ such that $\mathbb{E}\|Y(t) - Y(s)\|_E \le \varepsilon\delta$ if $|t-s| \le \rho$. Hence

$$\mathbb{P}(\|Y(t) - Y(s)\|_E \ge \varepsilon) \le \frac{\mathbb{E}\|Y(t) - Y(s)\|_E}{\varepsilon} \le \delta$$

if $|t-s| \le \rho$. Therefore Y is stochastically continuous. By Proposition 4.24 the process $Y(t)$ has a predictable version. □

In future if $X \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; E)$, we denote the integral process by $\int_0^t X(s)\, ds$, $0 \le t \le T$, even though in a set of probability zero its value is zero.

4.4.2 Martingales

If $\mathbb{E}\|X(t)\|_E < \infty$ for all $t \in I$, the process $X(t)$ is called *integrable*. Let $\{\mathcal{F}_t\}_{t\in I}$ be a filtration. An integrable adapted E-valued process $X(t)$, $t \in I$, is said to be a *martingale* if

$$\mathbb{E}(X(t)|\mathcal{F}_s) = X(s) \tag{4.7}$$

almost surely for all $s, t \in I$ such that $s \le t$. A real valued integrable adapted process $X(t)$, $t \in I$, is said to be a *submartingale* if

$$\mathbb{E}(X(t)|\mathcal{F}_s) \ge X(s)$$

almost surely for all $s, t \in I$ such that $s \le t$.

Proposition 4.27. *Let $I = [0,T]$ for some $T > 0$.*

(i) If $M(t)$, $t \in I$, is a martingale, $\|M(t)\|_E$, $t \in I$, is a submartingale.

(ii) If $M(t)$, $t \in I$, is a martingale, g is an increasing convex function from $[0,\infty)$ to $[0,\infty)$ and $\mathbb{E}[g(\|M(t)\|_E)] < \infty$ for all $t \in I$, then $g(\|M(t)\|_E)$, $t \in I$, is a submartingale.

Proof. (i) Let M be a martingale and $s, t \in I$ such that $s < t$. Then according to Theorem 4.9,

$$\|M(s)\|_E = \|\mathbb{E}(M(t)|\mathcal{F}_s)\|_E \le \mathbb{E}(\|M(t)\|_E|\mathcal{F}_s)$$

almost surely. Hence $\|M(t)\|_E$, $t \in I$, is a submartingale.

(ii) Since M is a martingale, $\|M(t)\|_E$, $t \in I$, is a submartingale by the statement (i). Since g is increasing,

$$g(\|M(s)\|_E) \le g(\mathbb{E}[\|M(t)\|_E|\mathcal{F}_s])$$

almost surely for all $s < t$. Since g is convex and $\mathbb{E}[g(\|M(t)\|_E)] < \infty$ for all $t \in I$,

$$g(\|M(s)\|_E) \leq \mathbb{E}[g(\|M(t)\|_E)|\mathcal{F}_s]$$

almost surely for all $s < t$ [53, Theorem 9.7]. Hence $g(\|M(t)\|_E)$, $t \in I$, is a submartingale. $\square$

We need the maximal inequality for real valued submartingales.

Theorem 4.28. [20, Theorem 1.3.8] *Let $X(t)$, $t \in I$, be a real valued continuous submartingale. If $X(t)$ is non-negative for all $t \in I$ and $p > 1$,*

$$\mathbb{E}\left(\sup_{t\in I}(X(t))^p\right) \leq \left(\frac{p}{p-1}\right)^p \sup_{t\in I}\mathbb{E}(X(t))^p.$$

As an immediate consequence of Proposition 4.27 and Theorem 4.28 we have the following corollary.

Corollary 4.29. *Let $M(t)$, $t \in I$, be an E-valued continuous martingale. If $p > 1$,*

$$\mathbb{E}\left(\sup_{t\in I}\|M(t)\|_E^p\right) \leq \left(\frac{p}{p-1}\right)^p \sup_{t\in I}\mathbb{E}\|M(t)\|_E^p.$$

If $M(t)$, $t \in [0,T]$, is an E-valued continuous martingale and $\mathbb{E}\|M(t)\|_E^p < \infty$ for all $t \in [0,T]$,

$$\sup_{t\in[0,T]} \mathbb{E}\|M(t)\|_E^p = \mathbb{E}\|M(T)\|_E^p$$

for all $p > 1$ by Theorem 4.9 and Proposition 4.27.

Theorem 4.30. *Let us denote by $\mathcal{M}_T^2(E)$ the vector space of E-valued continuous square integrable martingales on $[0,T]$. Then $\mathcal{M}_T^2(E)$ equipped with the norm*

$$\|M\|_{\mathcal{M}_T^2(E)} := \left(\mathbb{E}\sup_{t\in[0,T]}\|M(t)\|_E^2\right)^{\frac{1}{2}}$$

for all $M \in \mathcal{M}_T^2(E)$ is a Banach space.

Proof. If $M \in \mathcal{M}_T^2(E)$,

$$\|M\|^2_{\mathcal{M}_T^2(E)} = \mathbb{E}\sup_{t\in[0,T]}\|M(t)\|_E^2 \leq 4\sup_{t\in[0,T]}\mathbb{E}\|M(t)\|_E^2 = 4\mathbb{E}\|M(T)\|_E^2 < \infty$$

by Theorem 4.9, Proposition 4.27 and Corollary 4.29. Since for all $M \in \mathcal{M}_T^2(E)$

$$\|M\|_{\mathcal{M}_T^2(E)} = \left(\mathbb{E}\left(\sup_{t\in[0,T]}\|M(t)\|_E\right)^2\right)^{\frac{1}{2}} = \big\|\|M\|_{L^\infty(0,T;E)}\big\|_{L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R})},$$

$\|\cdot\|_{\mathcal{M}_T^2(E)}$ is a norm. Hence $\mathcal{M}_T^2(E)$ is a norm space.

To prove the completeness we assume that $\{M_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathcal{M}_T^2(E)$, i.e.,

$$\mathbb{E}\left(\sup_{t\in[0,T]}\|M_n(t)-M_m(t)\|_E^2\right)\longrightarrow 0$$

as $m,n\to\infty$. Since

$$c^2\mathbb{P}\left(\sup_{t\in[0,T]}\|M_n(t)-M_m(t)\|_E\geq c\right)\leq\mathbb{E}\left(\sup_{t\in[0,T]}\|M_n(t)-M_m(t)\|_E^2\right),$$

one can find a subsequence $\{M_{n_k}\}_{k=1}^{\infty}$ such that

$$\mathbb{P}\left(\sup_{t\in[0,T]}\|M_{n_{k+1}}(t)-M_{n_k}(t)\|_E\geq 2^{-k}\right)\leq 2^{-k}.$$

We denote

$$A_k:=\left\{\omega\in\Omega:\sup_{t\in[0,T]}\|M_{n_{k+1}}(t)-M_{n_k}(t)\|_E\geq 2^{-k}\right\}.$$

Then $\mathbb{P}(A_k)\leq 2^{-k}$. Since

$$\sum_{k=1}^{\infty}\mathbb{P}(A_k)\leq\sum_{k=1}^{\infty}2^{-k}=1,$$

according to the Borel–Cantelli lemma $\mathbb{P}(\limsup A_k)=0$. Thus for almost all $\omega\in\Omega$ there exists $l(\omega)\in\mathbb{N}$ such that $\omega\notin A_k$, i.e.,

$$\sup_{t\in[0,T]}\|M_{n_{k+1}}(t)-M_{n_k}(t)\|_E<2^{-k}$$

for all $k\geq l(\omega)$. Hence

$$\begin{aligned}\sup_{t\in[0,T]}\|M_{n_j}(t)-M_{n_k}(t)\|_E&\leq\sum_{i=k}^{j-1}\sup_{t\in[0,T]}\|M_{n_{i+1}}(t)-M_{n_i}(t)\|_E\\&<\sum_{i=k}^{j-1}2^{-i}\leq\sum_{i=k}^{\infty}2^{-i}=2^{-k+1}\end{aligned}$$

for all $j\geq k\geq l(\omega)$ for almost all $\omega\in\Omega$. Thus $\{M_{n_k}(\cdot,\omega)\}_{k=1}^{\infty}$ is a Cauchy sequence in $L^{\infty}(0,T;E)$ for almost all $\omega\in\Omega$. Therefore for almost all $\omega\in\Omega$ there exists $M(\cdot,\omega)\in L^{\infty}(0,T;E)$ such that $M_{n_k}(\cdot,\omega)\to M(\cdot,\omega)$ in $L^{\infty}(0,T;E)$ as $k\to\infty$. Since M_{n_k} is continuous for all $k\in\mathbb{N}$ and the convergence is uniform, M is continuous.

Let $t\in[0,T]$ be fixed. Then

$$\mathbb{E}\|M_{n_k}(t)-M_{n_l}(t)\|_E^2\leq\mathbb{E}\left(\sup_{t\in[0,T]}\|M_{n_k}(t)-M_{n_l}(t)\|_E^2\right)\longrightarrow 0$$

as $k,l\to\infty$. Thus $\{M_{n_k}(t)\}_{k=1}^{\infty}$ is a Cauchy sequence in $L^2(\Omega,\mathcal{F},\mathbb{P};E)$ for all $t\in[0,T]$. Therefore for all $t\in[0,T]$ there exists $N(t)\in L^2(\Omega,\mathcal{F},\mathbb{P};E)$ such that

$M_{n_k}(t) \to N(t)$ in $L^2(\Omega, \mathcal{F}, \mathbb{P}; E)$ as $k \to \infty$. By Theorem B.16 for all $t \in [0,T]$ there exists a subsequence $\{M_{n_{k_l}}(t,\omega)\}_{l=1}^\infty$ which converges pointwise to $N(t,\omega)$ for almost all $\omega \in \Omega$. Since $M_{n_k}(\cdot,\omega) \to M(\cdot,\omega)$ in $L^\infty(0,T;E)$ as $k \to \infty$ for almost all $\omega \in \Omega$ and M_{n_k} and M are continuous, $M_{n_k}(t,\omega)$ converges pointwise to $M(t,\omega)$ for all $t \in [0,T]$ and almost all $\omega \in \Omega$. Hence $N(t,\omega) = M(t,\omega)$ for all $t \in [0,T]$ and almost all $\omega \in \Omega$. Thus M is square integrable.

If $0 \le s \le t \le T$, then $\mathbb{E}(M_{n_k}(t)|\mathcal{F}_s) = M_{n_k}(s)$ almost surely for all $k \in \mathbb{N}$. By Theorem 4.9 for all $0 \le s \le t \le T$

$$\|\mathbb{E}(M_{n_k}(t) - M(t)|\mathcal{F}_s)\|_E \le \mathbb{E}(\|M_{n_k}(t) - M(t)\|_E|\mathcal{F}_s)$$

almost surely. Thus for all $t \in [0,T]$ and $s \le t$

$$\begin{aligned}\int_\Omega \|\mathbb{E}(M_{n_k}(t) - M(t)|\mathcal{F}_s)\|_E \, d\mathbb{P} &\le \int_\Omega \mathbb{E}(\|M_{n_k}(t) - M(t)\|_E|\mathcal{F}_s) \, d\mathbb{P} \\ &= \int_\Omega \|M_{n_k}(t) - M(t)\|_E \, d\mathbb{P} \\ &\le \left(\mathbb{E}\|M_{n_k}(t) - M(t)\|_E^2\right)^{\frac{1}{2}} \longrightarrow 0\end{aligned}$$

as $k \to \infty$. Hence $\mathbb{E}(M_{n_k}(t)|\mathcal{F}_s) \to \mathbb{E}(M(t)|\mathcal{F}_s)$ in $L^1(\Omega, \mathcal{F}, \mathbb{P}; E)$ as $k \to \infty$ for all $0 \le s \le t \le T$. Thus by Theorem B.16 for all $0 \le s \le t \le T$ there exists a subsequence $\{M_{n_{k_l}}(t)\}_{l=1}^\infty$ such that $\mathbb{E}(M_{n_{k_l}}(t)|\mathcal{F}_s)$ converges pointwise to $\mathbb{E}(M(t)|\mathcal{F}_s)$ almost surely. Then

$$\mathbb{E}(M(t)|\mathcal{F}_s) = \lim_{l\to\infty} \mathbb{E}(M_{n_{k_l}}(t)|\mathcal{F}_s) = \lim_{l\to\infty} M_{n_{k_l}}(s) = M(s)$$

almost surely for all $0 \le s \le t \le T$. Hence M is a martingale. Therefore $M \in \mathcal{M}_T^2(E)$.

Since M_{n_k} converges pointwise to M for all $0 \le t \le T$ almost surely, by Fatou's lemma,

$$\begin{aligned}\|M_{n_k} - M\|_{\mathcal{M}_T^2(E)}^2 &= \mathbb{E}\left(\sup_{t\in[0,T]} \|M_{n_k}(t) - M(t)\|_E^2\right) \\ &= \mathbb{E}\left(\lim_{l\to\infty} \sup_{t\in[0,T]} \|M_{n_k}(t) - M_{n_l}(t)\|_E^2\right) \\ &\le \liminf_{l\to\infty} \|M_{n_k} - M_{n_l}\|_{\mathcal{M}_T^2(E)}^2 < \varepsilon\end{aligned}$$

for $k \in \mathbb{N}$ large enough. Thus $M_{n_k} \to M$ in $\mathcal{M}_T^2(E)$ as $k \to \infty$ and hence $\mathcal{M}_T^2(E)$ is complete. □

4.4.3 Hilbert Space Valued Wiener Processes

Let $(H, (\cdot,\cdot)_H)$ be a real separable Hilbert space and $Q \in B(H)$ a positive self-adjoint trace class operator with $\operatorname{Ker} Q = \{0\}$. Then there exist an orthonormal basis $\{e_k\}_{k=1}^\infty$ in H and a bounded sequence $\{\lambda_k\}_{k=1}^\infty$ of positive numbers such that $Qe_k = \lambda_k e_k$ for all $k \in \mathbb{N}$ since Q is compact by Proposition D.9 [14, Theorem 5.1, Observation 6.1.b, pp. 113–116].

Definition 4.31. *An H-valued stochastic process $W(t)$, $t \geq 0$, is called a Q-Wiener process if*

(i) $W(0) = 0$,

(ii) W *is continuous,*

(iii) W *has independent increments, i.e.,* $W(u) - W(t)$ *and* $W(s) - W(r)$ *are independent for all* $0 \leq r < s \leq t < u < \infty$ *and*

(iv) $\mathcal{L}(W(t) - W(s)) = \mathcal{N}(0, (t-s)Q)$ *for all* $0 \leq s < t < \infty$.

If a process $W(t)$, $t \in [0,T]$, satisfies (i)-(ii) and (iii)-(iv) for $r, s, t, u \in [0,T]$, then W is a Q-Wiener process on $[0,T]$.

Let $(E, \|\cdot\|_E)$ be a real Banach space. An E-valued stochastic process X on I is said to be *Gaussian* if for any $n \in \mathbb{N}$ and for all $t_1, t_2, \ldots, t_n \in I$ the E^n-valued random variable $(X(t_1), X(t_2), \ldots, X(t_n))$ is Gaussian.

Proposition 4.32. *Let W be a Q-Wiener process. Then W is a Gaussian process on H such that $\mathbb{E}W(t) = 0$ and $\operatorname{Cov}(W(t)) = tQ$ for all $t \geq 0$. Furthermore, W has the expansion*

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \beta_k(t) e_k \tag{4.8}$$

for each $t \geq 0$ where

$$\beta_k(t) := \frac{1}{\sqrt{\lambda_k}} (W(t), e_k)_H$$

for all $k \in \mathbb{N}$ and $t \geq 0$ are mutually independent real valued Wiener processes on $(\Omega, \mathcal{F}, \mathbb{P})$ and the series on the right hand side of (4.8) converges in $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$.

Proof. Let W be a Q-Wiener process. We want to show that for all $t_1, \ldots, t_n \in [0, \infty)$ the H^n-valued random variable $(W(t_1), \ldots, W(t_n))$ is Gaussian. Thus we need to prove that

$$Z := ((W(t_1), \ldots, W(t_n)), (h_1, \ldots, h_n))_{H^n} := \sum_{i=1}^{n} (W(t_i), h_i)_H$$

is a real valued Gaussian random variable for all $h_1, \ldots, h_n \in H$. We may assume that $0 < t_1 < \ldots < t_n < \infty$. Then

$$Z = \left(W(t_1), \sum_{i=1}^{n} h_i \right)_H + \sum_{i=2}^{n} \left(W(t_i) - W(t_{i-1}), \sum_{j=i}^{n} h_j \right)_H .$$

Since W has independent increments, $W(t_1)$ and $W(t_i) - W(t_{i-1})$ for $i = 2, \ldots, n$ are mutually independent Gaussian random variables. Hence $(W(t_1), \sum_{i=1}^{n} h_i)_H$ and $\left(W(t_i) - W(t_{i-1}), \sum_{j=i}^{n} h_j\right)_H$ for $i = 2, \ldots, n$ are mutually independent real Gaussian random variables. Thus Z is Gaussian and therefore W is a Gaussian process. Additionally, $\mathbb{E}W(t) = 0$ and $\operatorname{Cov}(W(t)) = tQ$ for all $t \geq 0$ by Remark 4.18 and the conditions (i) and (iv) in Definition 4.31.

By the definition $\beta_k(t)$, $t \geq 0$, is a real valued Gaussian process for every $k \in \mathbb{N}$ [21, Theorem A.5]. In addition, $\beta_k(t)$, $t \geq 0$, satisfies the conditions (i)-(iii) in Definition 4.31. Let $0 < s < t$. Since the Wiener process has independent increments,

$$\begin{aligned}
&\mathbb{E}(\beta_k(t)\beta_l(s)) \\
&= \frac{1}{\sqrt{\lambda_k\lambda_l}}\mathbb{E}(W(t), e_k)_H(W(s), e_l)_H \\
&= \frac{1}{\sqrt{\lambda_k\lambda_l}}[\mathbb{E}(W(t) - W(s), e_k)_H(W(s), e_l)_H + \mathbb{E}(W(s), e_k)_H(W(s), e_l)_H] \\
&= \frac{1}{\sqrt{\lambda_k\lambda_l}}[\mathbb{E}(W(t) - W(s), e_k)_H\mathbb{E}(W(s), e_l)_H + \mathbb{E}(W(s), e_k)_H(W(s), e_l)_H] \\
&= \frac{s}{\sqrt{\lambda_k\lambda_l}}(Qe_k, e_l)_H = s\sqrt{\frac{\lambda_k}{\lambda_l}}\delta_{kl} = s\delta_{kl}
\end{aligned}$$

for all $k, l \in \mathbb{N}$. Since $\beta_k(t)$, $t \geq 0$, is Gaussian for all $k \in \mathbb{N}$, the calculation above implies that they are mutually independent. Furthermore, the covariance of $\beta_k(t) - \beta_k(s)$ is $t - s$ for all $k \in \mathbb{N}$ and $0 \leq s < t$. Hence $\beta_k(t)$, $t \geq 0$, is a real valued Wiener process. For $m, n \in \mathbb{N}$ such that $m > n$

$$\begin{aligned}
\mathbb{E}\left\|\sum_{k=n+1}^{m}\sqrt{\lambda_k}\beta_k(t)e_k\right\|_H^2 &= \mathbb{E}\left(\sum_{k=n+1}^{m}\lambda_k\beta_k^2(t)\right) = \sum_{k=n+1}^{m}\lambda_k\mathbb{E}(\beta_k^2(t)) \\
&= t\sum_{k=n+1}^{m}\lambda_k \longrightarrow 0
\end{aligned}$$

as $m, n \to \infty$ since $\operatorname{Tr} Q = \sum_{k=1}^{\infty}\lambda_k < \infty$ by Proposition D.14. Therefore the series on the right hand side of (4.8) converges in $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$. The set $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis in H. Thus

$$W(t) = \sum_{j=1}^{\infty}\sqrt{\lambda_j}\beta_j(t)e_j$$

for all $t \geq 0$. □

Some basic properties of Q-Wiener processes has been gathered in the following lemma.

Lemma 4.33. *Let W be a Q-Wiener process. Then*

(i) $\mathbb{E}\|W(t) - W(s)\|_H^2 = (t - s)\operatorname{Tr} Q$,

(ii) $\mathbb{E}\|W(t) - W(s)\|_H^4 \leq 3(\operatorname{Tr} Q)^2(t - s)^2$

for all $0 \leq s < t < \infty$.

Proof. (i) Since $\operatorname{Cov}(W(t) - W(s)) = (t - s)Q$ for all $0 \leq s < t < \infty$, by Lebesgue's monotone convergence theorem,

$$\begin{aligned}
\mathbb{E}\|W(t) - W(s)\|_H^2 &= \mathbb{E}\sum_{k=1}^{\infty}(W(t) - W(s), e_k)_H^2 = \sum_{k=1}^{\infty}\mathbb{E}(W(t) - W(s), e_k)_H^2 \\
&= \sum_{k=1}^{\infty}((t - s)Qe_k, e_k)_H = (t - s)\operatorname{Tr} Q
\end{aligned}$$

for all $0 \leq s < t < \infty$.

(ii) By Proposition 4.32 for all $k \in \mathbb{N}$

$$\beta_k(t) := \frac{1}{\sqrt{\lambda_k}}(W(t), e_k)_H$$

are mutually independent real valued Wiener processes on $[0, \infty)$. Then for all $0 \leq s < t < \infty$

$$\begin{aligned}
&\mathbb{E}\|W(t) - W(s)\|_H^4 \\
&= \mathbb{E}\left(\sum_{k=1}^{\infty}(W(t) - W(s), e_k)_H^2\right)^2 = \mathbb{E}\left(\sum_{k=1}^{\infty}\lambda_k(\beta_k(t) - \beta_k(s))^2\right)^2 \\
&= \mathbb{E}\sum_{k=1}^{\infty}\lambda_k^2(\beta_k(t) - \beta_k(s))^4 + 2\mathbb{E}\sum_{k=1}^{\infty}\sum_{l<k}\lambda_k\lambda_l(\beta_k(t) - \beta_k(s))^2(\beta_l(t) - \beta_l(s))^2.
\end{aligned}$$

Since $\beta_k(t) - \beta_k(s) \sim \mathcal{N}(0, t-s)$ and $\beta_k(t) - \beta_k(s)$ and $\beta_l(t) - \beta_l(s)$ are independent, $\mathbb{E}(\beta_k(t) - \beta_k(s))^4 = 3(t-s)^2$ and

$$\mathbb{E}(\beta_k(t) - \beta_k(s))^2(\beta_l(t) - \beta_l(s))^2 = \mathbb{E}(\beta_k(t) - \beta_k(s))^2\mathbb{E}(\beta_l(t) - \beta_l(s))^2 = (t-s)^2$$

for all $k \neq l$ and $0 \leq s < t < \infty$. Thus by Lebesgue's monotone convergence theorem,

$$\begin{aligned}
\mathbb{E}\|W(t) - W(s)\|_H^4 &= 3(t-s)^2\sum_{k=1}^{\infty}\lambda_k^2 + 2(t-s)^2\sum_{k=1}^{\infty}\sum_{l<k}\lambda_k\lambda_l \\
&= 2(t-s)^2\sum_{k=1}^{\infty}\lambda_k^2 + (t-s)^2\left(\sum_{k=1}^{\infty}\lambda_k\right)^2 \\
&\leq 3(t-s)^2\left(\sum_{k=1}^{\infty}\lambda_k\right)^2 = 3(t-s)^2(\operatorname{Tr} Q)^2
\end{aligned}$$

for all $0 \leq s < t < \infty$. □

Let $\{\mathcal{F}_t\}_{t\geq 0}$ be a filtration and W a Q-Wiener process. We say that W is a Q-Wiener process with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if $W(t)$ is $\mathcal{F}_t$-measurable and $W(t+h) - W(t)$ is independent of $\mathcal{F}_t$ for all $t \geq 0$ and $h > 0$. In that case W is a martingale because by Theorem 4.9,

$$\begin{aligned}
\mathbb{E}(W(t)|\mathcal{F}_s) &= \mathbb{E}(W(t) - W(s)|\mathcal{F}_s) + \mathbb{E}(W(s)|\mathcal{F}_s) \\
&= \mathbb{E}(W(t) - W(s)) + W(s) = W(s)
\end{aligned}$$

almost surely for all $0 \leq s < t < \infty$. Let $\{\mathcal{F}_t^W\}_{t\geq 0}$ be the filtration generated by the Wiener process W, i.e., $\mathcal{F}_t^W = \sigma(W(s),\ s \leq t)$ for all $t \geq 0$. Since W is a Q-Wiener process with respect to the filtration $\{\mathcal{F}_t^W\}_{t\geq 0}$, then W is a martingale with respect to its own history.

Let I be an interval. We denote all functions from I to H by $F(I, H)$. Let $\mathcal{F}(I, H)$ be the σ-algebra generated by set of the form

$$\{f \in F(I, H) : f(t_1) \in A_1, \ldots, f(t_n) \in A_n\} \tag{4.9}$$

where $n \in \mathbb{N}$ and $t_i \in I$ and $A_i \in \mathcal{B}(H)$ for $i = 1, \ldots, n$. If $\{X(t)\}_{t \in I}$ is an H-valued stochastic process, $X(\cdot, \omega)$ belongs to $F(I, H)$ for all $\omega \in \Omega$ and $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (F(I, H), \mathcal{F}(I, H))$ is measurable. We want to prove that the augmentation of the filtration generated by a Wiener process is right-continuous. We need the following lemma.

Lemma 4.34. *Let $t > 0$ and $\{Y(s)\}_{s \in [0,t]}$ be an H-valued stochastic process. A random variable $X : \Omega \to \mathbb{R}$ is $\sigma(Y(s),\ s \leq t)$-measurable if and only if $X = g(Y)$ where g is $\mathcal{F}([0,t], H)$-measurable function from $F([0,t], H)$ to $\mathbb{R}$.*

Proof. "$\Leftarrow$" The statement is obvious.

"$\Rightarrow$" It is enough to prove that if $X : \Omega \to \mathbb{R}$ is a bounded $\sigma(Y(s),\ s \leq t)$-measurable function, there exists a bounded $\mathcal{F}([0,t], H)$-measurable function g from $F([0,t], H)$ to $\mathbb{R}$ such that $X = g(Y)$. We define $\mathcal{H}$ to be the set of all bounded random variables $X : \Omega \to \mathbb{R}$ such that $X = g(Y)$ for some bounded $\mathcal{F}([0,t], H)$-measurable function. Then $\mathcal{H}$ is a vector space and the constant function 1 belongs to $\mathcal{H}$. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of non-negative random variables in $\mathcal{H}$ such that X_n increases monotonically pointwise to X and X is bounded, i.e., $0 \leq X \leq M$ for some $M > 0$. Then for all $n \in \mathbb{N}$ there exists a bounded $\mathcal{F}([0,t], H)$-measurable function g_n such that $X_n = g_n(Y)$. We denote $g := \chi_A \limsup_{n \to \infty} g_n$ where $A := \{\limsup_{n \to \infty} g_n \in [0, M]\}$. Then g is a bounded $\mathcal{F}([0,t], H)$-measurable function and $X = g(Y)$. Hence X belongs to $\mathcal{H}$. We define $\mathcal{I}$ to be the set of all sets of form $\{\omega : Y(t_1) \in A_1, \ldots, Y(t_n) \in A_n\}$ where $n \in \mathbb{N}$ and $0 \leq t_i \leq t$ and $A_i \in \mathcal{B}(H)$ for all $i = 1, \ldots, n$. Then $\mathcal{I}$ is a π-system and $\sigma(\mathcal{I}) = \sigma(Y(s),\ s \leq t)$. Furthermore,

$$\chi_{\{\omega \in \Omega : Y(t_1) \in A_1, \ldots, Y(t_n) \in A_n\}} = \chi_{\{f \in F([0,t], H) : f(t_1) \in A_1, \ldots, f(t_n) \in A_n\}}(Y)$$

for all $n \in \mathbb{N}$ and $0 \leq t_i \leq t$ and $A_i \in \mathcal{B}(H)$ for $i = 1, \ldots, n$. Therefore $\chi_B \in \mathcal{H}$ for all $B \in \mathcal{I}$. By the monotone class theorem $\mathcal{H}$ contains every bounded $\sigma(Y(s),\ s \leq t)$-measurable random variable. $\square$

Proposition 4.35. *If W is a Q-Wiener process, the augmentation $\{\mathcal{F}_t^{W,\mathbb{P}}\}_{t \geq 0}$ of the filtration $\{\mathcal{F}_t^W\}_{t \geq 0}$ is right-continuous. If, in addition, $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, the augmented filtration $\{\mathcal{F}_t^{W,\mathbb{P}}\}_{t \geq 0}$ is normal.*

Proof. Since W is a Wiener process with respect to its own history $\{\mathcal{F}_t^W\}_{t \geq 0}$, it is a Wiener process with respect to the augmented filtration $\{\mathcal{F}_t^{W,\mathbb{P}}\}_{t \geq 0}$. Let $s < t$. Then for all $h \in H$

$$\begin{aligned}
\mathbb{E}\left(e^{i(W(t),h)_H} | \mathcal{F}_s^{W,\mathbb{P}}\right) &= e^{i(W(s),h)_H} \mathbb{E}\left(e^{i(W(t)-W(s),h)_H} | \mathcal{F}_s^{W,\mathbb{P}}\right) \\
&= e^{i(W(s),h)_H} \mathbb{E}\left(e^{i(W(t)-W(s),h)_H}\right) \\
&= e^{i(W(s),h)_H - \frac{1}{2}(t-s)(Qh,h)_H}
\end{aligned}$$

almost surely. Let ε be such that $0 < \varepsilon < t - s$. Then

$$\begin{aligned}
\mathbb{E}\left(e^{i(W(t),h)_H} | \mathcal{F}_{s^+}^{W,\mathbb{P}}\right) &= \mathbb{E}\left(\mathbb{E}\left(e^{i(W(t),h)_H} | \mathcal{F}_{s+\varepsilon}^{W,\mathbb{P}}\right) | \mathcal{F}_{s^+}^{W,\mathbb{P}}\right) \\
&= \mathbb{E}\left(e^{i(W(s+\varepsilon),h)_H - \frac{1}{2}(t-s-\varepsilon)(Qh,h)_H} | \mathcal{F}_{s^+}^{W,\mathbb{P}}\right)
\end{aligned}$$

for all $h \in H$ almost surely. By passing to the limit $\varepsilon \to 0$

$$\mathbb{E}\left(e^{i(W(t),h)_H}|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right) = e^{i(W(s),h)_H - \frac{1}{2}(t-s)(Qh,h)_H}$$

for all $h \in H$ almost surely since W is continuous and Q is positive [53, Theorem 9.7]. Furthermore,

$$\begin{aligned}
\mathbb{E}\left(e^{i(W(t),h)_H}|W(s)\right) &= e^{i(W(s),h)_H}\mathbb{E}\left(e^{i(W(t)-W(s),h)_H}|W(s)\right) \\
&= e^{i(W(s),h)_H}\mathbb{E}\left(e^{i(W(t)-W(s),h)_H}\right) \\
&= e^{i(W(s),h)_H - \frac{1}{2}(t-s)(Qh,h)_H}
\end{aligned}$$

for all $h \in H$ almost surely. Thus

$$\mathbb{E}\left(e^{i(W(t),h)_H}|W(s)\right) = \mathbb{E}\left(e^{i(W(t),h)_H}|\mathcal{F}_s^{W,\mathbb{P}}\right) = \mathbb{E}\left(e^{i(W(t),h)_H}|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right)$$

for all $h \in H$ almost surely. Therefore according to Corollary 4.14,

$$E(f(W(t))|W(s)) = \mathbb{E}\left(f(W(t))|\mathcal{F}_s^{W,\mathbb{P}}\right) = \mathbb{E}\left(f(W(t))|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right)$$

for all bounded measurable functions f from H to $\mathbb{R}$ almost surely since

$$e^{i(W(s),h)_H - \frac{1}{2}(t-s)(Qh,h)_H} = \int_H e^{i(x,h)_H}\, \mathcal{N}(W(s),(t-s)Q)(dx)$$

for all $h \in H$. Let $s < t_1 < t_2$ and $f_1, f_2 : H \to \mathbb{R}$ be bounded measurable functions. Then

$$\begin{aligned}
\mathbb{E}\left(f_1(W(t_1))f_2(W(t_2))|\mathcal{F}_s^{W,\mathbb{P}}\right) &= \mathbb{E}\left(f_1(W(t_1))\mathbb{E}\left[f_2(W(t_2))|\mathcal{F}_{t_1}^{W,\mathbb{P}}\right]\Big|\mathcal{F}_s^{W,\mathbb{P}}\right) \\
&= \mathbb{E}\left(f_1(W(t_1))\mathbb{E}[f_2(W(t_2))|W(t_1)]|\mathcal{F}_s^{W,\mathbb{P}}\right) \\
&= \mathbb{E}\left(f_1(W(t_1))\mathbb{E}[f_2(W(t_2))|W(t_1)]|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right) \\
&= \mathbb{E}\left(f_1(W(t_1))\mathbb{E}\left[f_2(W(t_2))|\mathcal{F}_{t_1}^{W,\mathbb{P}}\right]\Big|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right) \\
&= \mathbb{E}\left(f_1(W(t_1))f_2(W(t_2))|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right)
\end{aligned}$$

almost surely because there exists a bounded measurable function f from H to $\mathbb{R}$ such that $\mathbb{E}[f_2(W(t_2))|W(t_1)] = f(W(t_1))$ [53, Lemma A3.2]. By induction,

$$\mathbb{E}\left(\prod_{i=1}^{n} f_i(W(t_i))\Big|\mathcal{F}_s^{W,\mathbb{P}}\right) = \mathbb{E}\left(\prod_{i=1}^{n} f_i(W(t_i))\Big|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right) \tag{4.10}$$

almost surely where $n \in \mathbb{N}$ and $0 \le t_i < \infty$ and f_i are bounded measurable functions from H to $\mathbb{R}$ for $i = 1, \ldots, n$.

We define $\mathcal{H}$ to be the set of all bounded functions g from $F([0,\infty), H)$ to $\mathbb{R}$ such that

$$\mathbb{E}\left(g(W)|\mathcal{F}_s^{W,\mathbb{P}}\right) = \mathbb{E}\left(g(W)|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right)$$

almost surely. Then $\mathcal{H}$ is a vector space and the constant function 1 is an element of $\mathcal{H}$. Furthermore if $\{g_n\}_{n=1}^{\infty}$ is a sequence of non-negative functions in $\mathcal{H}$ such that g_n increases monotonically pointwise to g and g is bounded, g belongs to $\mathcal{H}$ [53, Theorem 9.7]. Let $\mathcal{I}$ be the set of all sets of the form (4.9) where $I=[0,\infty)$. Then $\mathcal{I}$ is a π-system. Since

$$\chi_{\{f\in F([0,\infty),H):f(t_1)\in A_1,\ldots,f(t_n)\in A_n\}}(W)=\prod_{i=1}^{n}\chi_{A_i}(W(t_i))$$

when $n\in\mathbb{N}$ and $0\leq t_i<\infty$ and $A_i\in\mathcal{B}(H)$ for $i=1,\ldots,n$, by Formula (4.10) the characteristic functions of all sets in $\mathcal{I}$ belong to $\mathcal{H}$. According to the monotone class theorem $\mathcal{H}$ contains every bounded $\mathcal{F}([0,\infty),H)$-measurable functions from $F([0,\infty),H)$ to $\mathbb{R}$. Hence

$$\mathbb{E}\left(g(W)|\mathcal{F}_s^{W,\mathbb{P}}\right)=\mathbb{E}\left(g(W)|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right)$$

almost surely for all bounded $\mathcal{F}([0,\infty),H)$-measurable functions g from $F([0,\infty),H)$ to $\mathbb{R}$.

Let $s<t$. According to Lemma 4.34,

$$\mathbb{E}\left(X|\mathcal{F}_s^{W,\mathbb{P}}\right)=\mathbb{E}\left(X|\mathcal{F}_{s^+}^{W,\mathbb{P}}\right)$$

almost surely for every bounded $\mathcal{F}_t^{W,\mathbb{P}}$-measurable function $X:\Omega\to\mathbb{R}$ and hence for every $\mathcal{F}_t^{W,\mathbb{P}}$-measurable function $X:\Omega\to\mathbb{R}$ [53, Theorem 9.7]. Let $X:\Omega\to\mathbb{R}$ be an $\mathcal{F}_{s^+}^{W,\mathbb{P}}$-measurable function. Then $\mathbb{E}(X|\mathcal{F}_s^{W,\mathbb{P}})=X$ almost surely. Since $\mathcal{F}_s^{W,\mathbb{P}}$ is a complete σ-algebra, X is $\mathcal{F}_s^{W,\mathbb{P}}$-measurable. Consequently, $\mathcal{F}_{s^+}^{W,\mathbb{P}}\subseteq\mathcal{F}_s^{W,\mathbb{P}}$. The reverse inclusion $\mathcal{F}_s^{W,\mathbb{P}}\subseteq\mathcal{F}_{s^+}^{W,\mathbb{P}}$ is obvious. Hence $\mathcal{F}_s^{W,\mathbb{P}}=\mathcal{F}_{s^+}^{W,\mathbb{P}}$ for all $s\geq 0$.

If $(\Omega,\mathcal{F},\mathbb{P})$ is complete, $\{\mathcal{F}_t^{W,\mathbb{P}}\}_{t\geq 0}$ is a filtration. Since $\mathcal{F}_0^{W,\mathbb{P}}$ contains all $A\in\mathcal{F}$ such that $\mathbb{P}(A)=0$, the right-continuity assures the normality. □

The version in $\mathbb{R}$ of the previous proposition can be found from the book of Liptser and Shiryayev [27].

4.5 The Stochastic Integral

Let $(H,(\cdot,\cdot)_H)$ and $(U,(\cdot,\cdot)_U)$ be real separable Hilbert spaces and $Q\in B(U)$ a positive self-adjoint trace class operator with $\operatorname{Ker}Q=\{0\}$. Then there exist a complete orthonormal system $\{e_k\}_{k=1}^\infty$ in U and a bounded sequence $\{\lambda_k\}_{k=1}^\infty$ of positive numbers such that $Qe_k=\lambda_k e_k$ for all $k\in\mathbb{N}$ since Q is compact by Proposition D.9 [14, Theorem 5.1, Observation 6.1.b, pp. 113–116]. We introduce the subspace $U_0:=Q^{1/2}(U)$ of U, which endowed with the inner product

$$(u,v)_{U_0}:=\sum_{k=1}^{\infty}\frac{1}{\lambda_k}(u,e_k)_U(v,e_k)_U=(Q^{-1/2}u,Q^{-1/2}v)_U$$

is a Hilbert space. Then $\{g_k\}_{k=1}^\infty$ where $g_k:=\sqrt{\lambda_k}e_k$ for all $k\in\mathbb{N}$ is an orthonormal basis in U_0. In the construction of the stochastic integral an important rôle is played by the space $B_2(U_0,H)$ of Hilbert-Schmidt operators from U_0 to H. Let $\{f_k\}_{k=1}^\infty$

be an orthonormal basis in H. The space $B_2(U_0,H)$ is a separable Hilbert space equipped with the norm

$$\begin{aligned}\|\Psi\|^2_{B_2(U_0,H)} &= \sum_{k=1}^{\infty}\|\Psi g_k\|_H^2 = \sum_{k=1}^{\infty}\|\Psi Q^{1/2}e_k\|_H^2 = \|\Psi Q^{1/2}\|^2_{B_2(U,H)} \\ &= \|Q^{1/2}\Psi^*\|^2_{B_2(H,U)} = \sum_{k=1}^{\infty}\|Q^{1/2}\Psi^* f_k\|_U^2 \\ &= \sum_{k=1}^{\infty}(\Psi Q\Psi^* f_k, f_k)_H = \mathrm{Tr}[\Psi Q \Psi^*]\end{aligned}$$

for all $\Psi \in B_2(U_0,H)$. Clearly, $B(U,H) \subset B_2(U_0,H)$ but not all operators in $B_2(U_0,H)$ can be regarded as restrictions of operators in $B(U,H)$. The space $B_2(U_0,H)$ contains genuinely unbounded operators on U.

Let $W(t)$, $t \in [0,T]$, be a Q-Wiener process in a probability space $(\Omega,\mathcal{F},\mathbb{P})$ with values in U with respect to a normal filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ for a fixed $T > 0$. Let $\Phi(t)$, $t \in [0,T]$, be a $B(U,H)$-valued elementary process, i.e., there exist a sequence $0 = t_0 < t_1 < \ldots < t_k = T$ and a sequence $\{\Phi_m\}_{m=0}^{k-1}$ of $B(U,H)$-valued simple random variables such that Φ_m is $\mathcal{F}_{t_m}$-measurable and $\Phi(t) = \Phi_m$ if $t \in (t_m, t_{m+1}]$ for all $m = 0,1,\ldots,k-1$. We define the stochastic integral for elementary processes Φ by the formula

$$\int_0^t \Phi(s)\,dW(s) := \sum_{m=0}^{k-1}\Phi_m(W(t_{m+1}\wedge t) - W(t_m \wedge t))$$

and denote the stochastic integral by $\Phi\cdot W(t)$, $t \in [0,T]$. Let $\Phi(t)$, $t \in [0,T]$, be a $B_2(U_0,H)$-valued stochastic process. We define the norms

$$|||\Phi|||_t := \left(\mathbb{E}\int_0^t \|\Phi(s)\|^2_{B_2(U_0,H)}\,ds\right)^{\frac{1}{2}} = \left(\mathbb{E}\int_0^t \mathrm{Tr}\left(\Phi(s)Q\Phi^*(s)\right)\,ds\right)^{\frac{1}{2}}$$

for all $t \in [0,T]$. If Φ is an elementary process, $|||\Phi|||_t < \infty$ for all $t \in [0,T]$.

Proposition 4.36. *If Φ is an elementary process, the stochastic process $\Phi \cdot W$ is a continuous square integrable H-valued martingale on $[0,T]$ and*

$$\mathbb{E}\|\Phi\cdot W(t)\|^2 = |||\Phi|||_t^2 \tag{4.11}$$

for all $0 \le t \le T$. Furthermore, $\mathbb{E}(\Phi\cdot W(t)) = 0$ for all $0 \le t \le T$.

Proof. Since $W(t)$ is a continuous square integrable U-valued martingale and Φ is a $B(U,H)$-valued elementary process,

$$\Phi\cdot W(t) = \sum_{m=0}^{k-1}\Phi_m(W(t_{m+1}\wedge t) - W(t_m\wedge t))$$

is a continuous H-valued adapted process on $[0,T]$. Since Φ_m is a simple $\mathcal{F}_{t_m}$-measurable random variable, also Φ_m^* is for all $m = 0,\ldots,k-1$. Hence $\Phi_m^* h$ is a simple $\mathcal{F}_{t_m}$-measurable random variable for all $h \in H$ and $m = 0,\ldots,k-1$. Since

$W(t) - W(t_m)$ is independent of $\mathcal{F}_{t_m}$ for all $t \geq t_m$, by Theorem 4.9 and Proposition 4.10,

$$\begin{aligned}(\mathbb{E}\left(\Phi_m(W(t) - W(t_m))|\mathcal{F}_{t_m}\right), h)_H &= \mathbb{E}\left((\Phi_m(W(t) - W(t_m)), h)_H|\mathcal{F}_{t_m}\right)\\ &= \mathbb{E}\left((W(t) - W(t_m), \Phi_m^* h)_H|\mathcal{F}_{t_m}\right) = 0\end{aligned}$$

almost surely for all $h \in H$ and $m = 0, \ldots, k-1$. Hence

$$\mathbb{E}\left[\Phi_m(W(t) - W(t_m))\right] = \mathbb{E}\left[\mathbb{E}\left(\Phi_m(W(t) - W(t_m))|\mathcal{F}_m\right)\right] = 0$$

for all $t \geq t_m$ and $m = 0, \ldots, k-1$ and therefore $\mathbb{E}(\Phi \cdot W(t)) = 0$ for all $0 \leq t \leq T$.

Let $s \leq t_m < t$. Then

$$\mathbb{E}\left(\Phi_m(W(t) - W(t_m))|\mathcal{F}_s\right) = \mathbb{E}\left(\mathbb{E}\left(\Phi_m(W(t) - W(t_m))|\mathcal{F}_{t_m}\right)|\mathcal{F}_s\right) = 0$$

almost surely. If $t_m < t \leq s$, according to the measurability

$$\mathbb{E}\left(\Phi_m(W(t) - W(t_m))|\mathcal{F}_s\right) = \Phi_m(W(t) - W(t_m))$$

almost surely. If $t_m < s < t$,

$$\begin{aligned}&\mathbb{E}\left(\Phi_m(W(t) - W(t_m))|\mathcal{F}_s\right)\\ &= \mathbb{E}\left(\Phi_m(W(t) - W(s))|\mathcal{F}_s\right) + \mathbb{E}\left(\Phi_m(W(s) - W(t_m))|\mathcal{F}_s\right)\\ &= \Phi_m(W(s) - W(t_m))\end{aligned}$$

almost surely. Hence $\mathbb{E}(\Phi \cdot W(t)|\mathcal{F}_s) = \Phi \cdot W(s)$ almost surely for all $0 \leq s < t \leq T$, i.e., $\Phi \cdot W(t)$ is a martingale on $[0, T]$.

We still have to prove that $\Phi \cdot W(t)$ is square integrable. Let $t_m < t \leq t_{m+1}$. We denote

$$\begin{cases}\zeta_j := W(t_{j+1}) - W(t_j), & j = 0, \ldots, m-1,\\ \zeta_m := W(t) - W(t_m).\end{cases}$$

Then

$$\mathbb{E}\|\Phi \cdot W(t)\|_H^2 = \mathbb{E}\left\|\sum_{j=1}^m \Phi_j \zeta_j\right\|_H^2 = \mathbb{E}\sum_{j=1}^m \|\Phi_j \zeta_j\|_H^2 + 2\mathbb{E}\sum_{i<j=1}^m (\Phi_i \zeta_i, \Phi_j \zeta_j)_H.$$

Since ζ_j is independent of $\mathcal{F}_{t_j}$ and $\Phi_j^* h$ is $\mathcal{F}_{t_j}$-measurable for all $h \in H$ and $j = 0, \ldots, m$, by Lebesgue's monotone convergence theorem, Theorem 4.9 and Proposition 4.10,

$$\begin{aligned}\mathbb{E}\|\Phi_j \zeta_j\|_H^2 &= \mathbb{E}\sum_{l=1}^\infty (\Phi_j \zeta_j, f_l)_H^2 = \sum_{l=1}^\infty \mathbb{E}(\zeta_j, \Phi_j^* f_l)_U^2\\ &= \sum_{l=1}^\infty \mathbb{E}\left(\mathbb{E}\left[\left(\zeta_j, \Phi_j^* f_l\right)_U^2|\mathcal{F}_{t_j}\right]\right)\\ &= (t_{j+1} \wedge t - t_j)\sum_{l=1}^\infty \mathbb{E}(Q\Phi_j^* f_l, \Phi_j^* f_l)_U\\ &= (t_{j+1} \wedge t - t_j)\mathbb{E}\sum_{l=1}^\infty (\Phi_j Q\Phi_j^* f_l, f_l)_U\\ &= (t_{j+1} \wedge t - t_j)\mathbb{E}\operatorname{Tr}(\Phi_j Q\Phi_j^*)\end{aligned}$$

for all $j = 0, \dots, m$. Thus

$$\mathbb{E}\sum_{j=1}^{m} \|\Phi_j \zeta_j\|_H^2 = \mathbb{E}\sum_{j=1}^{m}(t_{j+1}\wedge t - t_j)\|\Phi_j\|^2_{B_2(U_0,H)}.$$

Let $i < j$. Then by Lebesgue's dominated convergence theorem,

$$E(\Phi_i\zeta_i, \Phi_j\zeta_j)_H = \sum_{k=1}^{\infty} \mathbb{E}\left[(\Phi_i\zeta_i, f_k)_H(\Phi_j\zeta_j, f_k)_H\right].$$

Since Φ_i is $\mathcal{F}_{t_i}$-measurable and ζ_i is $\mathcal{F}_{t_{i+1}}$-measurable, $(\Phi_i\zeta_i, f_k)_H$ is $\mathcal{F}_{t_{i+1}}$-measurable. Since ζ_j is independent of $\mathcal{F}_{t_j}$ and Φ_j is simple, $(\Phi_j\zeta_j, f_k)_H$ is independent of $\mathcal{F}_{t_j}$. Hence $(\Phi_i\zeta_i, f_k)_H$ and $(\Phi_j\zeta_j, f_k)_H$ are independent random variables. Thus

$$E(\Phi_i\zeta_i, \Phi_j\zeta_j)_H = \sum_{k=1}^{\infty} \mathbb{E}(\Phi_i\zeta_i, f_k)_H \mathbb{E}(\Phi_j\zeta_j, f_k)_H.$$

Since $\mathbb{E}(\Phi_i\zeta_i, f_k)_H = (\mathbb{E}\Phi_i\zeta_i, f_k)_H = 0$ for all $k \in \mathbb{N}$, then $E(\Phi_i\zeta_i, \Phi_j\zeta_j)_H = 0$. Thus

$$\begin{aligned}\mathbb{E}\|\Phi \cdot W(t)\|_H^2 &= \mathbb{E}\sum_{j=1}^{m}(t_{j+1}\wedge t - t_j)\|\Phi_j\|^2_{B_2(U_0,H)}\\ &= \mathbb{E}\int_0^t \|\Phi(s)\|^2_{B_2(U_0,H)}\, ds = |||\Phi|||_t^2.\end{aligned}$$

Hence $\Phi \cdot W(t)$ is square integrable and Equality (4.11) holds. □

By Corollary 4.29 for all $M \in \mathcal{M}_T^2(H)$

$$\mathbb{E}\|M(T)\|_H^2 \le \|M\|^2_{\mathcal{M}_T^2(H)} \le 4\mathbb{E}\|M(T)\|_H^2.$$

Hence by Equality (4.11),

$$|||\Phi|||_T \le \|\Phi \cdot W\|_{\mathcal{M}_T^2(H)} \le 2|||\Phi|||_T$$

for all elementary processes Φ. Therefore the stochastic integral is a bounded linear operator from the space of elementary processes with the norm $|||\cdot|||_T$ to the space $\mathcal{M}_T^2(H)$ of H-valued continuous square integrable martingales.

The definition of the stochastic integral can be extended to more general processes. The proper class of integrands is predictable processes with values in $B_2(U_0, H)$, more precisely, measurable mappings from $(\Omega_T, \mathcal{P}_T)$ to $(B_2(U_0, H), \mathcal{B}(B_2(U_0, H)))$.

Proposition 4.37. *(i) If Φ is a $B(U, H)$-valued predictable process, Φ is also a $B_2(U_0, H)$-valued predictable process. In particular, elementary processes are $B_2(U_0, H)$-valued predictable processes.*

(ii) If Φ is a $B_2(U_0, H)$-valued predictable process such that $|||\Phi|||_T < \infty$, there exists a sequence $\{\Phi_n\}_{n=1}^{\infty}$ of elementary processes such that

$$|||\Phi - \Phi_n|||_T \longrightarrow 0$$

as $n \to \infty$.

Proof. (i) Operators $(f_k \otimes g_j)u := f_k(u, g_j)_{U_0}$ for all $j, k \in \mathbb{N}$ and $u \in U_0$ are linearly dense in $B_2(U_0, H)$ by the proof of Proposition D.6. For all $T \in B_2(U_0, H)$

$$(f_k \otimes g_j, T)_{B_2(U_0,H)} = \sum_{l=1}^{\infty} ((f_k \otimes g_j)g_l, Tg_l)_H = (f_k, Tg_j)_H = \sqrt{\lambda_j}(Te_j, f_k)_H.$$

Let $\Phi : (\Omega_T, \mathcal{P}_T) \to (B(U, H), \mathcal{B}(U, H))$ be a random variable. Then

$$\sqrt{\lambda_j}(\Phi e_j, f_k)_H : (\Omega_T, \mathcal{P}_T) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

is measurable for all $j, k \in \mathbb{N}$. Thus

$$(f_k \otimes g_j, \Phi)_{B_2(U_0,H)} : \Omega_T \longrightarrow \mathbb{R}$$

is $\mathcal{P}_T$-measurable for all $j, k \in \mathbb{N}$. Hence Φ is a random variable from $(\Omega_T, \mathcal{P}_T)$ to $(B_2(U_0, H), \mathcal{B}(B_2(U_0, H)))$ according to Lemma 4.4.

Elementary processes are $B(U, H)$-valued predictable processes by definition. Hence they are $B_2(U_0, H)$-valued predictable processes.

(ii) The proof is similar to the one of Proposition 4.25. Let Φ be a $B_2(U_0, H)$-valued predictable process. Since $(B_2(U_0, H), \|\cdot\|_{B_2(U_0,H)})$ is a separable Hilbert space, by Lemma 4.3 there exists a sequence $\{\Phi_n\}_{n=1}^{\infty}$ of simple $B_2(U_0, H)$-valued predictable processes such that $\|\Phi_n(t, \omega) - \Phi(t, \omega)\|_{B_2(U_0,H)} \downarrow 0$ as $n \to \infty$ for all $(t, \omega) \in \Omega_T$. Since operators $f_k \otimes g_j = \sqrt{\lambda_j} f_k \otimes e_j$ are linearly dense in $B_2(U_0, H)$ and belong to $B(U, H)$, the space $B(U, H)$ is densely embedded to $B_2(U_0, H)$. Hence, because Φ_n is simple, there exists a $B(U, H)$-valued simple predictable process Ψ_n such that

$$\|\Psi_n(t, \omega) - \Phi_n(t, \omega)\|_{B_2(U_0,H)} < \frac{1}{n}$$

for all $n \in \mathbb{N}$. Thus $\{\Psi_n\}_{n=1}^{\infty}$ is a sequence of simple $B(U, H)$-valued predictable processes such that

$$\begin{aligned}
&\|\Phi(t, \omega) - \Psi_n(t, \omega)\|_{B_2(U_0,H)} \\
&\le \|\Phi(t, \omega) - \Phi_n(t, \omega)\|_{B_2(U_0,H)} + \|\Phi_n(t, \omega) - \Psi_n(t, \omega)\|_{B_2(U_0,H)} \downarrow 0
\end{aligned}$$

as $n \to \infty$ for all $(t, \omega) \in \Omega_T$. According to Lebesgue's monotone convergence theorem,

$$\int_{\Omega_T} \|\Phi(t, \omega) - \Psi_n(t, \omega)\|^2_{B_2(U_0,H)} \, \mathbb{P}_T(dt, d\omega) \longrightarrow 0$$

as $n \to \infty$. Hence by the Fubini theorem,

$$\mathbb{E} \int_0^T \|\Phi(t, \omega) - \Psi_n(t, \omega)\|^2_{B_2(U_0,H)} \, dt = |\!|\!| \Phi - \Psi_n |\!|\!|_T^2 \longrightarrow 0$$

as $n \to \infty$.

Since Ψ_n is a simple $B(U, H)$-valued $\mathcal{P}_T$-measurable random variable, it is of the form

$$\Psi_n(t, \omega) = \sum_{l=1}^{m_n} \Psi_n^l \chi_{A_n^l}(t, \omega)$$

for all $(t, \omega) \in \Omega_T$ where $m_n \in \mathbb{N}$ and $\Psi_n^l \in B(U, H)$ and $A_n^l \in \mathcal{P}_T$ for all $l = 1, \ldots, m_n$ and $A_n^i \cap A_n^j = \emptyset$ if $i \neq j$. We denote $C := \sum_{l=1}^{m_n} \|\Psi_n^l\|^2_{B_2(U_0,H)}$. According to Lemma 4.23 for all $l = 1, \dots, m_n$ there exists a finite union Γ_n^l of disjoint sets of the form (4.5) such that

$$\mathbb{P}_T((A_n^l \setminus \Gamma_n^l) \cup (\Gamma_n^l \setminus A_n^l)) < \frac{1}{n^2 C m_n}.$$

Then

$$\hat{\Psi}_n(t,\omega) := \sum_{l=1}^{m_n} \Psi_n^l \chi_{\Gamma_n^l}(t,\omega)$$

for all $(t,\omega) \in \Omega_T$ is an elementary process and

$$\begin{aligned} \|\!|\Psi_n - \hat{\Psi}_n\|\!|_T^2 &= \mathbb{E}\int_0^T \|\Psi_n(t,\omega) - \hat{\Psi}_n(t,\omega)\|_{B_2(U_0,H)}^2\, dt \\ &\le C\mathbb{P}_T\left(\bigcup_{l=1}^{m_n}(A_n^l \setminus \Gamma_n^l) \cup (\Gamma_n^l \setminus A_n^l)\right) \\ &\le C\sum_{l=1}^{m_n} \mathbb{P}_T((A_n^l \setminus \Gamma_n^l) \cup (\Gamma_n^l \setminus A_n^l)) < \frac{1}{n^2}. \end{aligned}$$

Thus $\{\hat{\Psi}_n\}_{n=1}^\infty$ is a sequence of elementary processes such that for every $\varepsilon > 0$

$$\|\!|\Phi - \hat{\Psi}_n\|\!|_T \le \|\!|\Phi - \Psi_n\|\!|_T + \|\!|\Psi_n - \hat{\Psi}_n\|\!|_T < \varepsilon$$

if $n \in \mathbb{N}$ is so large that $\|\!|\Phi - \Psi_n\|\!|_T < \varepsilon/2$ and $n \ge 2/\varepsilon$. □

We are able to extend the stochastic integral for $B_2(U_0, H)$-valued predictable processes Φ such that $\|\!|\Phi\|\!|_T < \infty$. They form the space $\mathcal{N}_W^2(0,T)$. Since $\mathcal{N}_W^2(0,T) = L^2(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; B_2(U_0,H))$, it is a Hilbert space. By Proposition 4.37 elementary processes form a dense set in $\mathcal{N}_W^2(0,T)$. Let $\Phi \in \mathcal{N}_W^2(0,T)$. Then there exists a sequence $\{\Phi_n\}_{n=1}^\infty \subset \mathcal{N}_W^2(0,T)$ of elementary processes such that $\|\!|\Phi - \Phi_n\|\!|_T \to 0$ as $n \to \infty$. Then the sequence $\{\Phi_n \cdot W\}_{n=1}^\infty$ is a Cauchy sequence in $\mathcal{M}_T^2(H)$ because the stochastic integral is a bounded linear operator for elementary processes. Since $\mathcal{M}_T^2(H)$ is complete, there exists $M \in \mathcal{M}_T^2(H)$ such that $\|\Phi_n \cdot W - M\|_{\mathcal{M}_T^2(H)} \to 0$ as $n \to \infty$. We define $\Phi \cdot W(t) := M(t)$ for all $t \in [0,T]$. Thus $\Phi \cdot W$ is an H-valued continuous square integrable martingale for every $\Phi \in \mathcal{N}_W^2(0,T)$. There are two equivalent norms in $\mathcal{M}_T^2(H)$. Hence

$$E\|\Phi \cdot W(t)\|_H^2 = \lim_{n\to\infty} \mathbb{E}\|\Phi_n \cdot W(t)\|_H^2 = \lim_{n\to\infty} \|\!|\Phi_n\|\!|_t^2 = \|\!|\Phi\|\!|_t^2$$

for all $t \in [0,T]$. Thus Equality (4.11) is valid for all $\Phi \in \mathcal{N}_W^2(0,T)$. The following theorem summarizes the main results of this section.

Theorem 4.38. *Let $\Phi \in \mathcal{N}_W^2(0,T)$. Then the stochastic integral $\Phi \cdot W$ is an H-valued continuous square integrable martingale and*

$$\mathbb{E}\|\Phi \cdot W(t)\|_H^2 = \|\!|\Phi\|\!|_t^2$$

for all $t \in [0,T]$.

4.5.1 Properties of the Stochastic Integral

In this subsection we introduce some further properties of the stochastic integral. In the following theorem we show that the stochastic integral is similar to the deterministic integral.

Theorem 4.39. *The stochastic integral is a bounded linear operator from $\mathcal{N}_W^2(0,T)$ to $\mathcal{M}_T^2(H)$. Furthermore if $\Phi \in \mathcal{N}_W^2(0,T)$,*

(i) for all $t \in [0,T]$

$$\int_0^T \Phi(s)\, dW(s) = \int_0^t \Phi(s)\, dW(s) + \int_t^T \Phi(s)\, dW(s),$$

(ii) if $(E, (\cdot,\cdot)_E)$ is a real separable Hilbert space and $A \in B(H,E)$,

$$A \int_0^t \Phi(s)\, dW(s) = \int_0^t A\Phi(s)\, dW(s)$$

for all $t \in [0,T]$.

Proof. Since $\Phi \cdot W \in \mathcal{M}_T^2(H)$ and $\mathbb{E}\|\Phi \cdot W(t)\|_H^2 = |||\Phi|||_t^2$ for all $\Phi \in \mathcal{N}_W^2(0,T)$ and $t \in [0,T]$, the stochastic integral is bounded by Corollary 4.29. Since the stochastic integral is linear on elementary processes and elementary processes are dense in $\mathcal{N}_W^2(0,T)$, it is linear also on $\mathcal{N}_W^2(0,T)$. The statements (i) and (ii) are valid for elementary processes. By the density they are valid also in $\mathcal{N}_W^2(0,T)$. □

The stochastic integral has a predictable version.

Proposition 4.40. *Let $\Phi \in \mathcal{N}_W^2(0,T)$. Then the stochastic integral $\Phi \cdot W$ has a predictable modification and $\mathbb{E}(\Phi \cdot W(t)) = 0$ for all $t \in [0,T]$.*

Proof. Let $\Phi \in \mathcal{N}_W^2(0,T)$ and $s < t$. Then the stochastic integral $\Phi \cdot W$ is mean square continuous since

$$\begin{aligned}\mathbb{E}\|\Phi \cdot W(t) - \Phi \cdot W(s)\|_H^2 &= \mathbb{E}\|((1-\chi_{[0,s]})\Phi) \cdot W(t)\|_H^2 = |||\chi_{[s,t]}\Phi|||_t^2 \\ &= \mathbb{E} \int_0^T \chi_{[s,t]}(r) \|\Phi(r)\|_{B_2(U_0,H)}^2\, dr\end{aligned}$$

and by Lebesgue's dominated convergence theorem $\mathbb{E}\|\Phi \cdot W(t) - \Phi \cdot W(s)\|_H^2 \to 0$ as $|t-s| \to 0$. By Lemma 4.20 the stochastic integral $\Phi \cdot W$ is stochastically continuous. Since $\Phi \cdot W$ is adapted, it has a predictable version by Proposition 4.24.

We still need to prove that $\mathbb{E}(\Phi \cdot W(t)) = 0$ for all $t \in [0,T]$. Let $\Phi \in \mathcal{N}_W^2(0,T)$ and $\{\Phi_n\}_{n=1}^\infty$ be a sequence of elementary processes defined by Proposition 4.37. Then for each $t \in [0,T]$

$$\begin{aligned}\|\mathbb{E}(\Phi \cdot W(t)) - \mathbb{E}(\Phi_n \cdot W(t))\|_H &= \|\mathbb{E}[(\Phi - \Phi_n) \cdot W(t)]\|_H \le \mathbb{E}\|(\Phi - \Phi_n) \cdot W(t)\|_H \\ &\le \left(\mathbb{E}\|(\Phi - \Phi_n) \cdot W(t)\|_H^2\right)^{\frac{1}{2}} = |||\Phi - \Phi_n|||_t \longrightarrow 0\end{aligned}$$

as $n \to \infty$. Since $\mathbb{E}(\Phi_n \cdot W(t)) = 0$ for all $n \in \mathbb{N}$, then $\mathbb{E}(\Phi \cdot W(t)) = 0$ for all $t \in [0,T]$. □

The correlation operator of two stochastic integrals is presented in the following proposition.

Proposition 4.41. *Let $\Phi_1, \Phi_2 \in \mathcal{N}_W^2(0,T)$. Then the correlation operator of the stochastic integrals $\Phi_1 \cdot W(t)$ and $\Phi_2 \cdot W(s)$ is given by the formula*

$$\operatorname{Cor}(\Phi_1 \cdot W(t), \Phi_2 \cdot W(s)) = \mathbb{E} \int_0^{s \wedge t} (\Phi_1(r) Q^{1/2})(\Phi_2(r) Q^{1/2})^* \, dr \tag{4.12}$$

for all $s, t \in [0,T]$.

Proof. Let $\Phi_1, \Phi_2 \in \mathcal{N}_W^2(0,T)$. Then $\Phi_i(t) Q^{1/2}$, $t \in [0,T]$, is a $B_2(U,H)$-valued predictable process for both $i = 1, 2$. Therefore $(\Phi_1(t) Q^{1/2})(\Phi_2(t) Q^{1/2})^*$, $t \in [0,T]$, is a $B_1(H)$-valued strongly measurable function on $(\Omega_T, \mathcal{P}_T, \mathbb{P}_T)$ and

$$\left\| (\Phi_1(t) Q^{1/2})(\Phi_2(t) Q^{1/2})^* \right\|_{B_1(H)} \leq \left\| \Phi_1(t) Q^{1/2} \right\|_{B_2(U,H)} \left\| \Phi_2(t) Q^{1/2} \right\|_{B_2(U,H)}$$

for all $t \in [0,T]$ by Proposition D.12. Consequently, for all $h \in H$

$$\begin{aligned}
&\mathbb{E} \int_0^T \left\| (\Phi_1(t) Q^{1/2})(\Phi_2(t) Q^{1/2})^* h \right\|_H \, dt \\
&\leq \|h\|_H \mathbb{E} \int_0^T \left\| (\Phi_1(t) Q^{1/2})(\Phi_2(t) Q^{1/2})^* \right\|_{B(H)} \, dt \\
&\leq \|h\|_H \mathbb{E} \int_0^T \left\| (\Phi_1(t) Q^{1/2})(\Phi_2(t) Q^{1/2})^* \right\|_{B_1(H)} \, dt \\
&\leq \|h\|_H \mathbb{E} \int_0^T \|\Phi_1(t)\|_{B_2(U_0,H)} \|\Phi_2(t)\|_{B_2(U_0,H)} \, dt \\
&\leq \|h\|_H \|\!|\Phi_1|\!\|_T \|\!|\Phi_2|\!\|_T.
\end{aligned} \tag{4.13}$$

Therefore the right hand side of (4.12) exists for all $s, t \in [0,T]$ as a strong Bochner integral.

The correlation operator $\operatorname{Cor}(\Phi_1 \cdot W(t), \Phi_2 \cdot W(s))$ is defined by

$$(\operatorname{Cor}(\Phi_1 \cdot W(t), \Phi_2 \cdot W(s)) a, b)_H = \mathbb{E}[(\Phi_1 \cdot W(t), b)_H (\Phi_2 \cdot W(s), a)_H]$$

for all $a, b \in H$ and $s, t \in [0,T]$. If Φ_1 and Φ_2 are elementary processes, there exists a partition $0 = t_0 < t_1 < \ldots < t_k = T$ such that $\Phi_1(t) = \Phi_m^1$ and $\Phi_2(t) = \Phi_m^2$ if $t \in (t_m, t_{m+1}]$ and Φ_m^1 and Φ_m^2 are simple $B(U,H)$-valued $\mathcal{F}_{t_m}$-measurable random variables for all $m = 0, 1, \ldots, k-1$. Let $s, t \in [0,T]$. Then there exist l and m such that $t_l < s \leq t_{l+1}$ and $t_m < t \leq t_{m+1}$. We denote

$$\begin{cases} \eta_j := W(t_{j+1}) - W(t_j), & j = 0, \ldots, l-1, \\ \eta_l := W(s) - W(t_l) \end{cases}$$

and

$$\begin{cases} \zeta_j := W(t_{j+1}) - W(t_j), & j = 0, \ldots, m-1, \\ \zeta_m := W(t) - W(t_m). \end{cases}$$

Then $\eta_j = \zeta_j$ for all $j < l \wedge m$ and

$$\mathbb{E}\left[(\Phi_1 \cdot W(t), b)_H (\Phi_2 \cdot W(s), a)_H\right] = \mathbb{E}\left[\left(\sum_{i=0}^{m} \left(\Phi_i^1 \zeta_i, b\right)_H\right)\left(\sum_{j=0}^{l} \left(\Phi_j^2 \eta_j, a\right)_H\right)\right]$$
$$= \sum_{i=0}^{m}\sum_{j=0}^{l} \mathbb{E}\left[\left(\Phi_i^1 \zeta_i, b\right)_H \left(\Phi_j^2 \eta_j, a\right)_H\right]$$

for all $a, b \in H$. We notice that $\Phi_i^1 \zeta_i$ is $\mathcal{F}_{t_{i+1}}$-measurable and independent of $\mathcal{F}_{t_i}$ for all $i = 0, \ldots, m$ and $\Phi_j^2 \eta_j$ is $\mathcal{F}_{t_{j+1}}$-measurable and independent of $\mathcal{F}_{t_j}$ for all $j = 0, \ldots, l$ because Φ_j^i is a simple random variable for all $i = 1, 2$ and $j = 0, \ldots, k-1$. Hence $(\Phi_i^1 \zeta_i, b)_H$ and $(\Phi_j^2 \eta_j, a)_H$ are independent for all $a, b \in H$ if $i \neq j$. Thus if $i \neq j$,

$$\mathbb{E}\left[\left(\Phi_i^1 \zeta_i, b\right)_H \left(\Phi_j^2 \eta_j, a\right)_H\right] = \mathbb{E}\left(\Phi_i^1 \zeta_i, b\right)_H \mathbb{E}\left(\Phi_j^2 \eta_j, a\right)_H.$$

On the other hand since Φ_j^i is a simple $B(U, H)$-valued $\mathcal{F}_{t_j}$-measurable random variable, also $(\Phi_j^i)^*$ is for all $i = 1, 2$ and $j = 0, \ldots, k-1$. Thus by Proposition 4.10,

$$\mathbb{E}\left(\Phi_i^1 \zeta_i, b\right)_H = \mathbb{E}\left(\zeta_i, (\Phi_i^1)^* b\right)_H = \mathbb{E}\left(\mathbb{E}\left[\left(\zeta_i, (\Phi_i^1)^* b\right)_H | \mathcal{F}_{t_i}\right]\right) = 0$$

for all $i = 0, \ldots, m$ and $b \in H$. Similarly $\mathbb{E}(\Phi_j^2 \eta_j, a)_H = 0$ for all $j = 0, \ldots, l$ and $a \in H$. Hence

$$\mathbb{E}\left[(\Phi_1 \cdot W(t), b)_H (\Phi_2 \cdot W(s), a)_H\right] = \sum_{j=0}^{l \wedge m} \mathbb{E}\left[\left(\Phi_j^1 \zeta_j, b\right)_H \left(\Phi_j^2 \eta_j, a\right)_H\right]$$

for all $a, b \in H$. Let $j < l \wedge m$. Then for all $a, b \in H$

$$\begin{aligned}
\mathbb{E}\left[\left(\Phi_j^1 \zeta_j, b\right)_H \left(\Phi_j^2 \eta_j, a\right)_H\right] &= \mathbb{E}\left[\left(\zeta_j, (\Phi_j^1)^* b\right)_H \left(\zeta_j, (\Phi_j^2)^* a\right)_H\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(\left(\zeta_j, (\Phi_j^1)^* b\right)_H \left(\zeta_j, (\Phi_j^2)^* a\right)_H | \mathcal{F}_{t_j}\right)\right] \\
&= \mathbb{E}\left[\left((t_{j+1} - t_j) Q (\Phi_j^2)^* a, (\Phi_j^1)^* b\right)_H\right] \\
&= (t_{j+1} - t_j) \mathbb{E}\left(\Phi_j^1 Q (\Phi_j^2)^* a, b\right)_H
\end{aligned}$$

by Proposition 4.10. If $l \wedge m = m$ and $j = m$, for all $a, b \in H$

$$\begin{aligned}
&\mathbb{E}\left[\left(\Phi_m^1 \zeta_m, b\right)_H \left(\Phi_m^2 \eta_m, a\right)_H\right] \\
&= \mathbb{E}\left[\left(W(t) - W(t_m), (\Phi_m^1)^* b\right)_H \left(W(t_{m+1} \wedge s) - W(t_m), (\Phi_m^2)^* a\right)_H\right] \\
&= \mathbb{E}\left[\left(W(t) - W(s \wedge t), (\Phi_m^1)^* b\right)_H \left(W(t_{m+1} \wedge s) - W(s \wedge t), (\Phi_m^2)^* a\right)_H\right] + \\
&\quad + \mathbb{E}\left[\left(W(t) - W(s \wedge t), (\Phi_m^1)^* b\right)_H \left(W(s \wedge t) - W(t_m), (\Phi_m^2)^* a\right)_H\right] + \\
&\quad + \mathbb{E}\left[\left(W(s \wedge t) - W(t_m), (\Phi_m^1)^* b\right)_H \left(W(t_{m+1} \wedge s) - W(s \wedge t), (\Phi_m^2)^* a\right)_H\right] + \\
&\quad + \mathbb{E}\left[\left(W(s \wedge t) - W(t_m), (\Phi_m^1)^* b\right)_H \left(W(s \wedge t) - W(t_m), (\Phi_m^2)^* a\right)_H\right] \\
&= \mathbb{E}\left[\left(W(s \wedge t) - W(t_m), (\Phi_m^1)^* b\right)_H \left(W(s \wedge t) - W(t_m), (\Phi_m^2)^* a\right)_H\right] \\
&= (s \wedge t - t_m) \mathbb{E}\left(\Phi_m^1 Q (\Phi_m^2)^* a, b\right)_H.
\end{aligned}$$

Similarly if $l \wedge m = l$ and $j = l$,

$$\mathbb{E}\left[\left(\Phi_l^1 \zeta_l, b\right)_H \left(\Phi_l^2 \eta_l, a\right)_H\right] = (s \wedge t - t_l) \mathbb{E}\left(\Phi_l^1 Q (\Phi_l^2)^* a, b\right)_H$$

for all $a, b \in H$. Hence

$$\begin{aligned}
&\mathbb{E}\left[(\Phi_1 \cdot W(t), b)_H (\Phi_2 \cdot W(s), a)_H\right] \\
&= \sum_{j=0}^{l\wedge m} (t_{j+1} \wedge s \wedge t - t_j)\mathbb{E}\big(\Phi_j^1 Q(\Phi_j^2)^* a, b\big)_H \\
&= \mathbb{E}\sum_{j=0}^{k-1} (t_{j+1} \wedge s \wedge t - t_j \wedge s \wedge t)\big(\Phi_j^1 Q(\Phi_j^2)^* a, b\big)_H \\
&= \mathbb{E}\int_0^{s\wedge t} (\Phi_1(r) Q \Phi_2^*(r) a, b)_H \, dr \\
&= \left(\mathbb{E}\int_0^{s\wedge t} \Phi_1(r) Q \Phi_2^*(r) a \, dr, b\right)_H
\end{aligned}$$

for all $a, b \in H$. Thus

$$\operatorname{Cor}(\Phi_1 \cdot W(t), \Phi_2 \cdot W(s)) = \mathbb{E}\int_0^{s\wedge t} \Phi_1(r) Q \Phi_2^*(r) \, dr$$

for elementary processes Φ_1 and Φ_2.

Let $\Phi_1, \Phi_2 \in \mathcal{N}_W^2(0,T)$ and $s, t \in [0,T]$. By Proposition 4.37 for both $i = 1, 2$ there exists a sequence $\{\Phi_i^n\}_{n=1}^\infty$ of elementary processes such that $|||\Phi_i - \Phi_i^n|||_T \to 0$ as $n \to \infty$. Then for all $a, b \in H$

$$\begin{aligned}
&|\mathbb{E}\left[(\Phi_1 \cdot W(t), b)_H (\Phi_2 \cdot W(s), a)_H\right] - \mathbb{E}\left[(\Phi_1^n \cdot W(t), b)_H (\Phi_2^m \cdot W(s), a)_H\right]| \\
&\le |\mathbb{E}\left[((\Phi_1 - \Phi_1^n) \cdot W(t), b)_H (\Phi_2 \cdot W(s), a)_H\right]| + \\
&\quad + |\mathbb{E}\left[(\Phi_1^n \cdot W(t), b)_H ((\Phi_2 - \Phi_2^m) \cdot W(s), a)_H\right]| \\
&\le \|a\|_H \|b\|_H \mathbb{E}\|(\Phi_1 - \Phi_1^n) \cdot W(t)\|_H \|\Phi_2 \cdot W(s)\|_H + \\
&\quad + \|a\|_H \|b\|_H \mathbb{E}\|\Phi_1^n \cdot W(t)\|_H \|(\Phi_2 - \Phi_2^m) \cdot W(s)\|_H \\
&\le \|a\|_H \|b\|_H \left(\mathbb{E}\|(\Phi_1 - \Phi_1^n) \cdot W(t)\|_H^2\right)^{\frac{1}{2}} \left(\mathbb{E}\|\Phi_2 \cdot W(s)\|_H^2\right)^{\frac{1}{2}} + \\
&\quad + \|a\|_H \|b\|_H \left(\mathbb{E}\|\Phi_1^n \cdot W(t)\|_H^2\right)^{\frac{1}{2}} \left(\mathbb{E}\|(\Phi_2 - \Phi_2^m) \cdot W(s)\|_H^2\right)^{\frac{1}{2}} \\
&= \|a\|_H \|b\|_H \left[|||\Phi_1 - \Phi_1^n|||_t |||\Phi_2|||_s + |||\Phi_1^n|||_t |||\Phi_2 - \Phi_2^m|||_s\right].
\end{aligned}$$

Since $\{\Phi_1^n\}_{n=1}^\infty$ is convergent in $\mathcal{N}_W^2(0,T)$, it is bounded, i.e., there exists $M > 0$ such that $|||\Phi_1^n|||_T \le M$ for all $n \in \mathbb{N}$. Thus

$$\|\operatorname{Cor}(\Phi_1 \cdot W(t), \Phi_2 \cdot W(s)) - \operatorname{Cor}(\Phi_1^n \cdot W(t), \Phi_2^m \cdot W(s))\|_{B(H)} \longrightarrow 0$$

as $m, n \to \infty$. On the other hand, for all $h \in H$

$$\begin{aligned}
&\left\|\mathbb{E}\int_0^{s\wedge t} (\Phi_1(r) Q^{1/2})(\Phi_2(r) Q^{1/2})^* h \, dr - \mathbb{E}\int_0^{s\wedge t} (\Phi_1^n(r) Q^{1/2})(\Phi_2^m(r) Q^{1/2})^* h \, dr\right\|_H \\
&\le \|h\|_H \mathbb{E}\int_0^{s\wedge t} \left\|(\Phi_1(r) Q^{1/2})(\Phi_2(r) Q^{1/2})^* - (\Phi_1^n(r) Q^{1/2})(\Phi_2^m(r) Q^{1/2})^*\right\|_{B(H)} dr \\
&\le \|h\|_H \mathbb{E}\int_0^{s\wedge t} \left\|(\Phi_1(r) Q^{1/2})(\Phi_2(r) Q^{1/2})^* - (\Phi_1^n(r) Q^{1/2})(\Phi_2^m(r) Q^{1/2})^*\right\|_{B_1(H)} dr \\
&\le \|h\|_H \mathbb{E}\int_0^{s\wedge t} \left\|\left[(\Phi_1(r) - \Phi_1^n(r)) Q^{1/2}\right] (\Phi_2(r) Q^{1/2})^*\right\|_{B_1(H)} dr + \\
&\quad + \|h\|_H \mathbb{E}\int_0^{s\wedge t} \left\|(\Phi_1^n(r) Q^{1/2}) \left[(\Phi_2(r) - \Phi_2^m(r)) Q^{1/2}\right]^*\right\|_{B_1(H)} dr \\
&\le \|h\|_H \left[|||\Phi_1 - \Phi_1^n|||_T |||\Phi_2|||_T + |||\Phi_1^n|||_T |||\Phi_2 - \Phi_2^m|||_T\right] \longrightarrow 0
\end{aligned}$$

as $m, n \to \infty$ by Inequality (4.13). Thus

$$\mathrm{Cor}(\Phi_1 \cdot W(t), \Phi_2 \cdot W(s)) = \mathbb{E} \int_0^{s \wedge t} (\Phi_1(r) Q^{1/2})(\Phi_2(r) Q^{1/2})^* \, dr$$

for all $\Phi_1, \Phi_2 \in \mathcal{N}_W^2(0, T)$ and $s, t \in [0, T]$. □

The Gaussianity of the Wiener process is inherited to the stochastic integral if the integrand is deterministic.

Lemma 4.42. *If $\Phi \in L^2(0, T; B_2(U_0, H))$, then $\Phi \cdot W$ is a Gaussian process in H. The covariance operator of $\Phi \cdot W(t)$ is*

$$\mathrm{Cov}(\Phi \cdot W(t)) = \int_0^t (\Phi(s) Q^{1/2})(\Phi(s) Q^{1/2})^* \, ds$$

for all $t \in [0, T]$. Furthermore, $\Phi \cdot W(t)$ is independent of $\mathcal{F}_0$ for all $t \in [0, T]$.

Proof. If $\Phi \in L^2(0, T; B_2(U_0, H))$, then Φ is a deterministic $B_2(U_0, H)$-valued predictable process such that $|||\Phi|||_T = \|\Phi\|_{L^2(0,T;B_2(U_0,H))}$. By Proposition 4.41 the covariance of $\Phi \cdot W(t)$ is

$$\mathrm{Cov}(\Phi \cdot W(t)) = \mathbb{E} \int_0^t (\Phi(s) Q^{1/2})(\Phi(s) Q^{1/2})^* \, ds = \int_0^t (\Phi(s) Q^{1/2})(\Phi(s) Q^{1/2})^* \, ds$$

for all $t \in [0, T]$.

Let Φ be a deterministic elementary process, i.e., there exist a sequence $0 = t_0 < t_1 < \ldots < t_k = T$ and a sequence $\{\Phi_m\}_{m=0}^{k-1}$ of bounded linear operators such that $\Phi(t) = \Phi_m$ if $t \in (t_m, t_{m+1}]$ for all $m = 0, \ldots, k-1$. Then for all $t \in [0, T]$

$$\Phi \cdot W(t) = \sum_{m=0}^{k-1} \Phi_m (W(t_{m+1} \wedge t) - W(t_m \wedge t)).$$

Since $W(t) - W(s)$ is independent of $\mathcal{F}_s$ for all $0 \le s < t \le T$, the stochastic integral $\Phi \cdot W(t)$ is independent of $\mathcal{F}_0$ for all $t \in [0, T]$. We want to show that for all $l \in \mathbb{N}$ and $s_1, \ldots, s_l \in [0, T]$ the H^l-valued random variable $(\Phi \cdot W(s_1), \ldots, \Phi \cdot W(s_l))$ is Gaussian. Let $h_1, \ldots, h_l \in H$. We need to prove that

$$((\Phi \cdot W(s_1), \ldots, \Phi \cdot W(s_l)), (h_1, \ldots, h_l))_{H^l} := \sum_{i=1}^{l} (\Phi \cdot W(s_i), h_i)_H$$

is a real valued Gaussian random variable. We may assume that $0 \le s_1 < \ldots < s_l \le T$. We combine $\{t_m\}_{m=0}^{k}$ and $\{s_i\}_{i=1}^{l}$ to be a partition $\{r_j\}_{j=1}^{k+l+1}$ of the interval $[0, T]$. Thus

$$\begin{aligned} \sum_{i=1}^{l} (\Phi \cdot W(s_i), h_i)_H &= \sum_{i=1}^{l} \left(\sum_{j=1}^{k+l+1} \tilde{\Phi}_j (W(r_{j+1} \wedge s_i) - W(r_j \wedge s_i)), h_i \right)_H \\ &= \sum_{i=1}^{l} \sum_{j: s_{i-1} \le r_j < s_i} \left(\tilde{\Phi}_j (W(r_{j+1}) - W(r_j)), \sum_{n=i}^{l} h_n \right)_H \end{aligned}$$

where $\tilde{\Phi}_j = \Phi_m$ if $r_j = t_m$ or $r_j = s_i$ and $t_m < s_i < t_{m+1}$ for all $j = 1, \ldots, k + l + 1$. Since $\tilde{\Phi}_j(W(r_{j+1}) - W(r_j))$ is a $\mathcal{F}_{r_{j+1}}$-measurable Gaussian random variable independent of $\mathcal{F}_{r_j}$ for all $j = 1, \ldots, k + l + 1$ and the sum of mutually independent real valued Gaussian random variables is Gaussian, $\Phi \cdot W$ is a Gaussian process.

Let $\Phi \in L^2(0, T; B_2(U_0, H))$. Then there exists a sequence $\{\Phi_n\}_{n=1}^{\infty}$ of elementary processes such that $|||\Phi - \Phi_n|||_T \to 0$ as $n \to \infty$ by Proposition 4.37. The sequence can be chosen such a way that Φ_n are deterministic since Φ is deterministic. Thus $\Phi_n \cdot W$ is a Gaussian process. We want to show that for all $k \in \mathbb{N}$ and $t_1, \ldots, t_k \in [0, T]$ the H^k-valued random variable $(\Phi \cdot W(t_1), \ldots, \Phi \cdot W(t_k))$ is Gaussian. Let $h_1, \ldots, h_l \in H$. Since

$$\begin{aligned}
&\left(\mathbb{E} \left| \sum_{i=1}^{k} (\Phi \cdot W(t_i), h_i)_H - \sum_{i=1}^{k} (\Phi_n \cdot W(t_i), h_i)_H \right|^2 \right)^{\frac{1}{2}} \\
&\leq \sum_{i=1}^{k} \left(\mathbb{E}((\Phi - \Phi_n) \cdot W(t_i), h_i)_H^2 \right)^{\frac{1}{2}} \leq \sum_{i=1}^{k} \|h_i\|_H \left(\mathbb{E}\|(\Phi - \Phi_n) \cdot W(t_i)\|_H^2 \right)^{\frac{1}{2}} \\
&= \sum_{i=1}^{k} \|h_i\|_H |||\Phi - \Phi_n|||_{t_i} \leq |||\Phi - \Phi_n|||_T \sum_{i=1}^{k} \|h_i\|_H \longrightarrow 0
\end{aligned}$$

as $n \to \infty$ and the limit of real valued Gaussian random variables in $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ is Gaussian [21, Theorem A.7],

$$\sum_{i=1}^{k} (\Phi \cdot W(t_i), h_i)_H = ((\Phi \cdot W(t_1), \ldots, \Phi \cdot W(t_k)), (h_1, \ldots, h_k))_{H^k}$$

is a real valued Gaussian random variable. Hence $\Phi \cdot W$ is a Gaussian process.

Since $\Phi_n \cdot W(t)$ is independent of $\mathcal{F}_0$, for all $A \in \mathcal{F}_0$, $h \in H$, $n \in \mathbb{N}$ and $t \in [0, T]$

$$\mathbb{E}\left[e^{i(\Phi_n \cdot W(t), h)_H} \chi_A \right] = \mathbb{P}(A) \mathbb{E}\left[e^{i(\Phi_n \cdot W(t), h)_H} \right].$$

Since $\Phi_n \cdot W \to \Phi \cdot W$ in $\mathcal{M}_T^2(H)$ as $n \to \infty$, then $\Phi_n \cdot W(t) \to \Phi \cdot W(t)$ in $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$ as $n \to \infty$ for all $t \in [0, T]$. Thus for all $t \in [0, T]$ there exists a subsequence $\{\Phi_{n_k}\}_{k=0}^{\infty}$ such that $\Phi_{n_k} \cdot W(t)$ converges pointwise to $\Phi \cdot W(t)$ almost surely. Hence by Lebesgue's dominated convergence theorem,

$$\begin{aligned}
&\left| \mathbb{E}\left[e^{i(\Phi_{n_k} \cdot W(t), h)_H} \chi_A \right] - \mathbb{E}\left[e^{i(\Phi \cdot W(t), h)_H} \chi_A \right] \right| \\
&\leq \mathbb{E}\left[\left| e^{i(\Phi_{n_k} \cdot W(t), h)_H} - e^{i(\Phi \cdot W(t), h)_H} \right| \chi_A \right] \longrightarrow 0
\end{aligned}$$

as $k \to \infty$ for all $A \in \mathcal{F}_0$ and $h \in H$. Therefore

$$\mathbb{E}\left[e^{i(\Phi \cdot W(t), h)_H} \chi_A \right] = \mathbb{P}(A) \mathbb{E}\left[e^{i(\Phi \cdot W(t), h)_H} \right]$$

for all $A \in \mathcal{F}_0$, $h \in H$ and $t \in [0, T]$. Thus

$$\mathbb{E}\left[e^{i(\Phi \cdot W(t), h)_H} | \mathcal{F}_0 \right] = \mathbb{E}\left[e^{i(\Phi \cdot W(t), h)_H} \right]$$

almost surely for all $h \in H$ and $t \in [0, T]$. Since

$$\mathbb{E}\left[e^{i(\Phi \cdot W(t), h)_H} \right] = \int_{\Omega} e^{i(\Phi \cdot W(t), h)_H} \, d\mathbb{P} = \int_H e^{i(x, h)_H} \, \mathcal{L}(\Phi \cdot W(t))(dx)$$

for all $h \in H$, by Lemma 4.13 for all $B \in \mathcal{B}(H)$ and $t \in [0,T]$

$$\mathbb{P}(\Phi \cdot W(t) \in B|\mathcal{F}_0) = \mathcal{L}(\Phi \cdot W(t))(B) = \mathbb{P}(\Phi \cdot W(t) \in B).$$

If $A \in \mathcal{F}_0$ and $B \in \mathcal{B}(H)$, for all $t \in [0,T]$

$$\begin{aligned}\mathbb{P}(A \cap \{\Phi \cdot W(t) \in B\}) = \mathbb{E}(\chi_A \chi_{\{\Phi \cdot W(t) \in B\}}) &= \int_A \chi_{\{\Phi \cdot W(t) \in B\}} \, d\mathbb{P} \\ &= \int_A \mathbb{P}(\Phi \cdot W(t) \in B|\mathcal{F}_0) \, d\mathbb{P} = \mathbb{P}(A)\mathbb{P}(\Phi \cdot W(t) \in B).\end{aligned}$$

Thus $\Phi \cdot W(t)$ is independent of $\mathcal{F}_0$ for all $t \in [0,T]$. □

4.5.2 The Ito Formula

Let $\varphi \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)$, $\Phi \in \mathcal{N}_W^2(0,T)$ and X_0 be an $\mathcal{F}_0$-measurable H-valued random variable. Then the process

$$X(t) = X_0 + \int_0^t \varphi(s) \, ds + \int_0^t \Phi(s) \, dW(s)$$

for all $t \in [0,T]$ is well defined since the trajectories of φ are Bochner integrable almost surely. The process X is continuous and has a predictable version by Lemma 4.26, Theorem 4.38 and Proposition 4.40. Let us assume that the function $F : [0,T] \times H \to \mathbb{R}$ is continuously differentiable with respect to t and twice continuously differentiable with respect to h. Moreover, we assume that F_h and F_{hh} are bounded, F_t and F_{hh} are uniformly continuous on bounded subsets of $[0,T] \times H$ and F_t is Lipschitz continuous with respect to h with integrable Lipschitz constant, i.e., for all $t \in [0,T]$ there exists $L(t) > 0$ such that

$$|F_t(t,h) - F_t(t,f)| \leq L(t)\|h - f\|_H$$

for all $f, h \in H$ and $L \in L^1(0,T)$.

Theorem 4.43. *Under the above conditions for all $t \in [0,T]$ the Ito formula*

$$\begin{aligned}F(t, X(t)) = F(0, X_0) + \int_0^t (F_h(s, X(s)), \Phi(s) dW(s))_H + \int_0^t F_t(s, X(s)) \, ds + \\ + \int_0^t (F_h(s, X(s)), \varphi(s))_H \, ds + \frac{1}{2} \int_0^t \operatorname{Tr}(F_{hh}(s, X(s))\Phi(s)Q\Phi^*(s)) \, ds\end{aligned}$$

is valid almost surely.

Proof. Let us assume that there exist $\varphi_0 \in H$ and $\Phi_0 \in B(U,H)$ such that $\varphi(t) = \varphi_0$ and $\Phi(t) = \Phi_0$ for all $t \in [0,T]$. Then $X(t) = X_0 + t\varphi_0 + \Phi_0 W(t)$ for all $t \in [0,T]$. Since the Wiener process is continuous and $W(t)$, $t \in [0,T]$, is $\mathcal{F}_t$-measurable, X is continuous and adapted. Let $s,t \in [0,T]$ such that $s \leq t$. Then

$$\begin{aligned}(\mathbb{E}\|X(t) - X(s)\|_H^2)^{\frac{1}{2}} &= (\mathbb{E}\|(t-s)\varphi_0 + \Phi_0(W(t) - W(s))\|_H^2)^{\frac{1}{2}} \\ &\leq (\mathbb{E}\|(t-s)\varphi_0\|_H^2)^{\frac{1}{2}} + (\mathbb{E}\|\Phi_0(W(t) - W(s))\|_H^2)^{\frac{1}{2}} \\ &\leq (t-s)\|\varphi_0\|_H + \|\Phi_0\|_{B(U,H)}(\mathbb{E}\|W(t) - W(s))\|_U^2)^{\frac{1}{2}} \\ &\leq (t-s)\|\varphi_0\|_H + \|\Phi_0\|_{B(U,H)}\sqrt{(t-s)\operatorname{Tr} Q}\end{aligned}$$

by Lemma 4.33. Hence X is mean square continuous. Therefore X has a predictable version by Lemma 4.20 and Proposition 4.24.

Let the points $0 = t_0 < t_1 < \ldots < t_k = t$ define a partition π of a fixed interval $[0, t] \subseteq [0, T]$. We denote $|\pi| := \max_{0 \le j \le k-1}(t_{j+1} - t_j)$. Then

$$\begin{aligned} &F(t, X(t)) - F(0, X_0) \\ &= \sum_{j=0}^{k-1}[F(t_{j+1}, X(t_{j+1})) - F(t_j, X(t_{j+1}))] + \sum_{j=0}^{k-1}[F(t_j, X(t_{j+1})) - F(t_j, X(t_j))]. \end{aligned}$$

Since F is continuously differentiable with respect to t and twice continuously differentiable with respect to h and X is continuous, by applying Taylor's formula for almost all $\omega \in \Omega$ there exist $\theta_0, \theta_1, \ldots, \theta_{k-1}, \vartheta_0, \vartheta_1, \ldots, \vartheta_{k-1} \in [0, 1]$ such that

$$\begin{aligned} &F(t, X(t)) - F(0, X_0) \\ &= \sum_{j=0}^{k-1} F_t(t_{j+1}, X(t_{j+1}))\Delta t_j + \sum_{j=0}^{k-1}[F_t(\tilde{t}_j, X(t_{j+1})) - F_t(t_{j+1}, X(t_{j+1}))]\Delta t_j + \\ &+ \sum_{j=0}^{k-1} (F_h(t_j, X(t_j)), \Delta X_j)_H + \frac{1}{2}\sum_{j=0}^{k-1} (F_{hh}(t_j, X(t_j))\Delta X_j, \Delta X_j)_H + \\ &+ \frac{1}{2}\sum_{j=0}^{k-1} \Big([F_{hh}(t_j, \tilde{X}_j) - F_{hh}(t_j, X(t_j))]\Delta X_j, \Delta X_j\Big)_H \end{aligned}$$

where $\Delta t_j := t_{j+1} - t_j$, $\tilde{t}_j := t_j + \theta_j(t_{j+1} - t_j)$, $\Delta X_j := X(t_{j+1}) - X(t_j)$ and $\tilde{X}_j := X(t_j) + \vartheta_j(X(t_{j+1}) - X(t_j))$ for all $j = 0, \ldots, k-1$. We examine the terms of the Taylor expansion one by one.

Since F_t is continuous from $[0, T] \times H$ to $\mathbb{R}$ and X is continuous,

$$\sum_{j=0}^{k-1} F_t(t_{j+1}, X(t_{j+1}))\Delta t_j$$

is an approximation of the Riemann integral $\int_0^t F_t(s, X(s))\, ds$ almost surely. Thus

$$\sum_{j=0}^{k-1} F_t(t_{j+1}, X(t_{j+1}))\Delta t_j \longrightarrow \int_0^t F_t(s, X(s))\, ds$$

as $|\pi| \to 0$ almost surely.

Since X is continuous, for almost all $\omega \in \Omega$ the set

$$A_\omega := \{h \in H : X(s, \omega) = h \text{ for some } s \in [0, T]\}$$

is bounded in H. Since F_t is uniformly continuous on bounded subsets of $[0, T] \times H$, for all $\varepsilon > 0$ there exists $\delta(\omega) > 0$ such that $|F_t(s, h) - F_t(r, h)| < \varepsilon$ for all $s, r \in [0, t]$, $h \in A_\omega$ and $|s - r| < \delta(\omega)$. Let $|\pi| < \delta(\omega)$. Then

$$\begin{aligned} &\left|\sum_{j=0}^{k-1}[F_t(\tilde{t}_j, X(t_{j+1})) - F_t(t_{j+1}, X(t_{j+1}))]\Delta t_j\right| \\ &\le \sum_{j=0}^{k-1} |F_t(\tilde{t}_j, X(t_{j+1})) - F_t(t_{j+1}, X(t_{j+1}))|\Delta t_j < \sum_{j=0}^{k-1} \varepsilon\Delta t_j = \varepsilon t \end{aligned}$$

since $|\tilde{t}_j - t_{j+1}| = (1-\theta_j)(t_{j+1}-t_j) < \delta(\omega)$ for all $j = 0, \dots, k-1$. Hence

$$\sum_{j=0}^{k-1}[F_t(\tilde{t}_j, X(t_{j+1})) - F_t(t_{j+1}, X(t_{j+1}))]\Delta t_j \longrightarrow 0$$

as $|\pi| \to 0$ almost surely.

The third term has to be divided in two parts

$$\begin{aligned}&\sum_{j=0}^{k-1}(F_h(t_j, X(t_j)), \Delta X_j)_H \\ &= \sum_{j=0}^{k-1}(F_h(t_j, X(t_j)), \varphi_0)_H \Delta t_j + \sum_{j=0}^{k-1}(F_h(t_j, X(t_j)), \Phi_0 \Delta W_j)_H\end{aligned}$$

where $\Delta W_j := W(t_{j+1}) - W(t_j)$ for all $j = 0, \dots, k-1$. Since F_h is continuous from $[0,T] \times H$ to H and X is continuous, $(F_h(s, X(s)), \varphi_0)_H$, $s \in [0,T]$, is continuous. Hence

$$\sum_{j=0}^{k-1}(F_h(t_j, X(t_j)), \varphi_0)_H \Delta t_j \longrightarrow \int_0^t (F_h(s, X(s)), \varphi(s))_H \, ds$$

as $|\pi| \to 0$ almost surely. Since F_h is bounded, there exists $C > 0$ such that $\|F_h(s,h)\|_H \le C$ for all $s \in [0,T]$ and $h \in H$. Since $(F_h(s,h), \Phi_0 \cdot)_H : U \to \mathbb{R}$ is a bounded linear operator, $(F_h(s,h), \Phi_0 \cdot)_H \in B_2(U_0, \mathbb{R})$ for all $s \in [0,T]$ and $h \in H$. Furthermore,

$$\|(F_h(s,h), \Phi_0 \cdot)_H\|_{B_2(U_0,\mathbb{R})} \le \|F_h(s,h)\|_H \|\Phi\|_{B_2(U_0,H)} \le C\|\Phi\|_{B(U,H)} \sqrt{\operatorname{Tr} Q}.$$

Since X has a predictable version and F_h is continuous, $(F_h(s, X(s)), \Phi_0 \cdot)$ has a predictable version with values in $B_2(U_0, \mathbb{R})$. Since

$$\begin{aligned}|||(F_h(\cdot, X(\cdot)), \Phi_0 \cdot)_H|||_t^2 &= \mathbb{E} \int_0^t \|(F_h(s, X(s)), \Phi_0 \cdot)_H\|^2_{B_2(U_0,\mathbb{R})} \, ds \\ &\le tC^2 \|\Phi\|^2_{B(U,H)} \operatorname{Tr} Q,\end{aligned}$$

$(F_h(\cdot, X(\cdot)), \Phi_0 \cdot)_H \in \mathcal{N}_W^2(0,t; B_2(U_0, \mathbb{R}))$. Hence $\int_0^s (F_h(r, X(r)), \Phi(r) dW(r))_H \in \mathcal{M}_t^2(\mathbb{R})$. On the other hand,

$$G_\pi(s) := \sum_{j=0}^{k-1}(F_h(t_j, X(t_j)), \Phi_0 \cdot)_H \chi_{(t_j, t_{j+i}]}(s)$$

is a $B(U,H)$-valued predictable process for all $s \in [0,t]$ since $X(t_j)$ is $\mathcal{F}_{t_j}$-measurable for all $j = 0, \dots, k-1$, and

$$|||G_\pi|||_t^2 = \mathbb{E} \sum_{j=0}^{k-1} \|(F_h(t_j, X(t_j)), \Phi_0 \cdot)_H\|^2_{B_2(U_0,\mathbb{R})} \Delta t_j \le tC^2 \|\Phi\|^2_{B(U,H)} \operatorname{Tr} Q.$$

Hence $G_\pi \in \mathcal{N}_W^2(0,t;B_2(U_0,\mathbb{R}))$. Then

$$
\begin{aligned}
&\mathbb{E}\left|\int_0^t (F_h(s,X(s)),\Phi(s)dW(s))_H - \sum_{j=0}^{k-1}(F_h(t_j,X(t_j)),\Phi_0\Delta W_j)_H\right|^2 \\
&= \mathbb{E}\int_0^t \|(F_h(s,X(s)),\Phi_0\cdot)_H - G_\pi(s)\|^2_{B_2(U_0,\mathbb{R})}\, ds \\
&= \mathbb{E}\int_0^t \sum_{j=0}^{k-1}\|(F_h(s,X(s)),\Phi_0\cdot)_H - (F_h(t_j,X(t_j)),\Phi_0\cdot)_H\|^2_{B_2(U_0,\mathbb{R})}\chi_{(t_j,t_{j+i}]}(s)\, ds \\
&\le \|\Phi\|^2_{B(U,H)}\operatorname{Tr} Q\,\mathbb{E}\int_0^t\sum_{j=0}^{k-1}\|F_h(s,X(s)) - F_h(t_j,X(t_j))\|^2_H\chi_{(t_j,t_{j+i}]}(s)\, ds.
\end{aligned}
$$

Since F_h is continuous and X is continuous, $F_h(s,X(s))$, $s\in[0,T]$, is continuous. Thus

$$
\sum_{j=0}^{k-1}\|F_h(s,X(s)) - F_h(t_j,X(t_j))\|^2_H\chi_{(t_j,t_{j+i}]}(s) \longrightarrow 0
$$

as $|\pi|\to 0$ almost surely. Since

$$
\sum_{j=0}^{k-1}\|F_h(s,X(s)) - F_h(t_j,X(t_j))\|^2_H\chi_{(t_j,t_{j+i}]}(s) \le 4C^2
$$

for all $s\in[0,T]$, by Lebesgue's dominated convergence theorem,

$$
\mathbb{E}\left|\int_0^t (F_h(s,X(s)),\Phi(s)dW(s))_H - \sum_{j=0}^{k-1}(F_h(t_j,X(t_j)),\Phi_0\Delta W_j)_H\right|^2 \longrightarrow 0
$$

as $|\pi|\to 0$. Hence there exists a subsequence such that

$$
\sum_{j=0}^{k-1}(F_h(t_j,X(t_j)),\Phi_0\Delta W_j)_H \longrightarrow \int_0^t (F_h(s,X(s)),\Phi(s)dW(s))_H
$$

as $|\pi|\to 0$ almost surely.

The fourth term has to be divided in three parts

$$
\begin{aligned}
&\sum_{j=0}^{k-1}(F_{hh}(t_j,X(t_j))\Delta X_j,\Delta X_j)_H \\
&= \sum_{j=0}^{k-1}(F_{hh}(t_j,X(t_j))\varphi_0,\varphi_0)_H(\Delta t_j)^2 + \sum_{j=0}^{k-1}(F_{hh}(t_j,X(t_j))\Phi_0\Delta W_j,\Phi_0\Delta W_j)_H + \\
&\quad + \sum_{j=0}^{k-1}[(F_{hh}(t_j,X(t_j))\varphi_0,\Phi_0\Delta W_j)_H + (F_{hh}(t_j,X(t_j))\Phi_0\Delta W_j,\varphi_0)_H]\Delta t_j.
\end{aligned}
$$

Since F_{hh} is a bounded function from $[0,T]\times H$ to $B(H)$, i.e., there exists $D>0$ such that $\|F_{hh}(t,h)\|_{B(H)}\le D$ for all $(t,h)\in[0,T]\times H$,

$$
\left|\sum_{j=0}^{k-1}(F_{hh}(t_j,X(t_j))\varphi_0,\varphi_0)_H(\Delta t_j)^2\right| \le D\|\varphi_0\|^2_H\sum_{j=0}^{k-1}(\Delta t_j)^2 \longrightarrow 0
$$

as $|\pi| \to 0$. Thus

$$\sum_{j=0}^{k-1} (F_{hh}(t_j, X(t_j))\varphi_0, \varphi_0)_H (\Delta t_j)^2 \longrightarrow 0$$

as $|\pi| \to 0$ for all $\omega \in \Omega$. Since W is continuous, for almost all $\omega \in \Omega$ and all $\varepsilon > 0$ there exists $\delta(\omega) > 0$ such that $\|W(t) - W(s)\|_U < \varepsilon$ for all $s, t \in [0, T]$ and $|t - s| < \delta(\omega)$. Let $|\pi| < \delta(\omega)$. Then

$$\begin{aligned}
&\left| \sum_{j=0}^{k-1} [(F_{hh}(t_j, X(t_j))\varphi_0, \Phi_0 \Delta W_j)_H + (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \varphi_0)_H] \Delta t_j \right| \\
&\leq 2D\|\varphi_0\|_H \|\Phi_0\|_{B(U,H)} \sum_{j=0}^{k-1} \|W(t_{j+1}) - W(t_j)\|_U \Delta t_j < 2tD\|\varphi_0\|_H \|\Phi_0\|_{B(U,H)} \varepsilon.
\end{aligned}$$

Hence

$$\sum_{j=0}^{k-1} [(F_{hh}(t_j, X(t_j))\varphi_0, \Phi_0 \Delta W_j)_H + (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \varphi_0)_H] \Delta t_j \longrightarrow 0$$

as $|\pi| \to 0$ almost surely. Let $\{f_l\}_{l=1}^{\infty}$ be an orthonormal basis in H. Then by the Lebesgue's dominated convergence theorem,

$$\begin{aligned}
&\mathbb{E}[\chi_A (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H] \\
&= \mathbb{E}\left[\chi_A \sum_{l=1}^{\infty} (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, f_l)_H (\Phi_0 \Delta W_j, f_l)_H \right] \\
&= \sum_{l=1}^{\infty} \mathbb{E}[\chi_A (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, f_l)_H (\Phi_0 \Delta W_j, f_l)_H]
\end{aligned}$$

for all $A \in \mathcal{F}_{t_j}$ and $j = 0, \ldots, k-1$ because for all $n \in \mathbb{N}$

$$\begin{aligned}
&\sum_{l=1}^{n} |(F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, f_l)_H (\Phi_0 \Delta W_j, f_l)_H| \\
&\leq \|F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j\|_H \|\Phi_0 \Delta W_j\|_H \leq D\|\Phi_0\|_{B(U,H)}^2 \|\Delta W_j\|_U^2
\end{aligned}$$

and $\mathbb{E}\|\Delta W_j\|_U^2 = \Delta t_j \operatorname{Tr} Q$ for all $j = 0, \ldots, k-1$. Since ΔW_j is independent of $\mathcal{F}_{t_j}$ and $\Phi_0^* F_{hh}^*(t_j, X(t_j)) f_l$ is $\mathcal{F}_{t_j}$-measurable, according to Proposition 4.10,

$$\begin{aligned}
&\mathbb{E}[\chi_A (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, f_l)_H (\Phi_0 \Delta W_j, f_l)_H] \\
&= \mathbb{E}[\chi_A (\Delta W_j, \Phi_0^* F_{hh}^*(t_j, X(t_j)) f_l)_H (\Delta W_j, \Phi_0^* f_l)_H] \\
&= \mathbb{E}[\chi_A \mathbb{E}((\Delta W_j, \Phi_0^* F_{hh}^*(t_j, X(t_j)) f_l)_H (\Delta W_j, \Phi_0^* f_l)_H | \mathcal{F}_{t_j})] \\
&= \mathbb{E}[\chi_A (\Delta t_j Q \Phi_0^* f_l, \Phi_0^* F_{hh}^*(t_j, X(t_j)) f_l)_H] \\
&= \Delta t_j \mathbb{E}[\chi_A (F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^* f_l, f_l)_H]
\end{aligned}$$

for all $A \in \mathcal{F}_{t_j}$ and $j = 0, \ldots, k-1$ and $l \in \mathbb{N}$. Thus

$$\begin{aligned}
&\mathbb{E}[\chi_A (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H] \\
&= \sum_{l=1}^{\infty} \Delta t_j \mathbb{E}[\chi_A (F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^* f_l, f_l)_H]
\end{aligned}$$

for all $A \in \mathcal{F}_{t_j}$ and $j = 0, \dots, k-1$. Since

$$\sum_{l=1}^{n} |(F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^* f_l, f_l)_H| \le \|F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^*\|_{B_1(H)} \\ \le D\|\Phi_0\|^2_{B(U,H)} \operatorname{Tr} Q$$

for all $n \in \mathbb{N}$, by Lebesgue's dominated convergence theorem,

$$\begin{aligned} &\mathbb{E}[\chi_A (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H] \\ &= \Delta t_j \mathbb{E}[\chi_A \sum_{l=1}^{\infty} (F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^* f_l, f_l)_H] \\ &= \Delta t_j \mathbb{E}[\chi_A \operatorname{Tr}(F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^*)] \end{aligned}$$

for all $A \in \mathcal{F}_{t_j}$ and $j = 0, \dots, k-1$. Hence

$$\mathbb{E}[(F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H | \mathcal{F}_{t_j}] = \Delta t_j \operatorname{Tr}(F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^*)$$

almost surely for all $j = 0, \dots, k-1$. Since $(F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H$ is $\mathcal{F}_{t_i}$-measurable for all $0 \le j < i \le k-1$ and by Lemma 4.33,

$$\begin{aligned} \mathbb{E}(F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H^2 &\le D^2 \|\Phi_0\|^4_{B(U,H)} \mathbb{E}\|\Delta W_j\|_U^4 \\ &\le 3D^2 \|\Phi_0\|^4_{B(U,H)} (\operatorname{Tr} Q)^2 (\Delta t_j)^2, \end{aligned}$$

we can use Lemma 4.11. Thus

$$\begin{aligned} &\mathbb{E}\left(\sum_{j=0}^{k-1} (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H - \sum_{j=0}^{k-1} \operatorname{Tr}(F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^*)\Delta t_j\right)^2 \\ &= \sum_{j=0}^{k-1} \Big[\mathbb{E}(F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H^2 - \mathbb{E}(\operatorname{Tr}(F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^*)\Delta t_j)^2\Big] \\ &\le \sum_{j=0}^{k-1} \mathbb{E}(F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H^2 \\ &\le 3D^2 \|\Phi_0\|^4_{B(U,H)} (\operatorname{Tr} Q)^2 \sum_{j=0}^{k-1} (\Delta t_j)^2 \longrightarrow 0 \end{aligned}$$

as $|\pi| \to 0$. Hence there exists a subsequence such that

$$\sum_{j=0}^{k-1} (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H - \sum_{j=0}^{k-1} \operatorname{Tr}(F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^*)\Delta t_j \longrightarrow 0$$

as $|\pi| \to 0$ almost surely. Since F_{hh} is continuous, X is continuous and $\operatorname{Tr} A - \operatorname{Tr} B = \operatorname{Tr}(A - B)$ for all $A, B \in B_1(H)$, the process $\operatorname{Tr}(F_{hh}(s, X(s))\Phi_0 Q \Phi_0^*)$ is continuous on $[0, T]$. Therefore

$$\sum_{j=0}^{k-1} \operatorname{Tr}(F_{hh}(t_j, X(t_j))\Phi_0 Q \Phi_0^*)\Delta t_j \longrightarrow \int_0^t \operatorname{Tr}(F_{hh}(s, X(s))\Phi_0 Q \Phi_0^*)\, ds$$

as $|\pi| \to 0$ almost surely. Thus

$$\sum_{j=0}^{k-1} (F_{hh}(t_j, X(t_j))\Phi_0 \Delta W_j, \Phi_0 \Delta W_j)_H \longrightarrow \int_0^t \operatorname{Tr}(F_{hh}(s, X(s))\Phi_0 Q \Phi_0^*)\, ds$$

as $|\pi| \to 0$ almost surely.

Since F_{hh} is uniformly continuous on bounded subsets of $[0,T] \times H$, for all $\varepsilon > 0$ and almost all $\omega \in \Omega$ there exists $\delta(\omega) > 0$ such that $\|F_{hh}(s,h) - F_{hh}(s,f)\|_{B(H)} < \varepsilon$ for all $s \in [0,t]$, $f, h \in A_\omega$ and $\|h - f\|_H < \delta(\omega)$. Since X is continuous, for almost all $\omega \in \Omega$ there exists $\rho(\omega) > 0$ such that $\|X(s) - X(r)\|_H < \delta(\omega)$ for all $r, s \in [0,t]$ and $|s - r| < \rho(\omega)$. Let $|\pi| < \rho(\omega)$. Then

$$\left| \sum_{j=0}^{k-1} \Big([F_{hh}(t_j, \tilde{X}_j) - F_{hh}(t_j, X(t_j))]\Delta X_j, \Delta X_j \Big)_H \right| < \varepsilon \sum_{j=0}^{k-1} (\Delta X_j, \Delta X_j)_H$$

since $\|\tilde{X}_j - X(t_j)\|_H = \vartheta_j \|X(t_{j+1}) - X(t_j)\|_H < \delta(\omega)$ for all $j = 0, \dots, k-1$. The examination of the fourth term showed that there exists a subsequence such that

$$\sum_{j=0}^{k-1} (\Delta X_j, \Delta X_j)_H \longrightarrow \int_0^t \operatorname{Tr}(\Phi_0 Q \Phi_0^*)\, ds = t \operatorname{Tr}(\Phi_0 Q \Phi_0^*)$$

as $|\pi| \to 0$ almost surely. Thus there exists a subsequence such that $\{\sum_{j=0}^{k-1} \|\Delta X_j\|_H^2\}$ is a bounded sequence almost surely. By using the subsequence

$$\sum_{j=0}^{k-1} \Big([F_{hh}(t_j, \tilde{X}_j) - F_{hh}(t_j, X(t_j))]\Delta X_j, \Delta X_j \Big)_H \longrightarrow 0$$

as $|\pi| \to 0$ almost surely.

Therefore, if $X(t) = X_0 + t\varphi_0 + \Phi_0 W(t)$, for all $t \in [0,T]$

$$\begin{aligned} &F(t, X(t)) - F(0, X_0) \\ &= \int_0^t F_t(s, X(s))\, ds + \int_0^t (F_h(s, X(s)), \varphi_0)_H\, ds + \\ &\quad + \int_0^t (F_h(s, X(s)), \Phi_0 dW(s))_H + \frac{1}{2} \int_0^t \operatorname{Tr}(F_{hh}(s, X(s))\Phi_0 Q \Phi_0^*)\, ds \end{aligned}$$

almost surely.

If φ and Φ are elementary processes, there exist a sequence $0 = t_0 < t_1 < \ldots < t_k = T$ and a number $n_m \in \mathbb{N}$ and sets $\{A_m^l\}_{l=1}^{n_m} \subset \mathcal{F}_{t_m}$ for all $m = 0, \dots, k-1$ such that

$$\begin{aligned} \varphi(t,\omega) &= \sum_{m=0}^{k-1} \sum_{l=1}^{n_m} \varphi_m^l \chi_{A_m^l}(\omega) \chi_{(t_m, t_{m+1}]}(t), \\ \Phi(t,\omega) &= \sum_{m=0}^{k-1} \sum_{l=1}^{n_m} \Phi_m^l \chi_{A_m^l}(\omega) \chi_{(t_m, t_{m+1}]}(t) \end{aligned}$$

for all $(t,\omega) \in \Omega_T$ where $\varphi_m^l \in H$ and $\Phi_m^l \in B(U,H)$ for all $m = 0,\dots,k-1$ and $l = 1,\dots,n_m$. Thus

$$\begin{aligned} X(t) = X_0 + \sum_{m=0}^{k-1}\sum_{l=1}^{n_m} \varphi_m^l \chi_{A_m^l}(t_{m+1}\wedge t - t_m \wedge t) + \\ + \sum_{m=0}^{k-1}\sum_{l=1}^{n_m} \Phi_m^l \chi_{A_m^l}(W(t_{m+1}\wedge t) - W(t_m\wedge t)) \end{aligned}$$

for all $t \in [0,T]$. Hence

$$\begin{aligned} X(t) &= X(t_m) + \varphi_m^l(t - t_m) + \Phi_m^l(W(t) - W(t_m)) \\ &= X(t_m) - \varphi_m^l t_m - \Phi_m^l W(t_m) + \varphi_m^l t + \Phi_m^l W(t) \end{aligned}$$

if $t \in [t_m, t_{m+1}]$ and $\omega \in A_m^l$ for some $m = 0,\dots,k-1$ and $l = 1,\dots,n_m$. Since $X(t_m) - \varphi_m^l t_m - \Phi_m^l W(t_m)$ is $\mathcal{F}_{t_m}$-measurable and the Ito formula is valid pointwise almost surely for processes of the form $X(t) = X_0 + t\varphi_0 + \Phi_0 W(t)$, the Ito formula is valid also almost surely for elementary processes φ and Φ.

If $\varphi \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)$ and $\Phi \in \mathcal{N}_W^2(0,T)$, by Propositions 4.25 and 4.37 there exist a sequence $\{\varphi_n\}_{n=1}^\infty$ of H-valued elementary processes and a sequence $\{\Phi_m\}_{m=1}^\infty$ of $B(U,H)$-valued elementary processes such that $|||\Phi_m - \Phi|||_T \to 0$ and $\|\varphi_n - \varphi\|_{L^1(\Omega_T,\mathcal{P}_T,\mathbb{P}_T;H)} \to 0$ as $m,n \to \infty$. We define the processes X and $X_{n,m}$ by

$$\begin{aligned} X(t) &:= X_0 + \int_0^t \varphi(s)\, ds + \int_0^t \Phi(s)\, dW(s) \\ X_{n,m}(t) &:= X_0 + \int_0^t \varphi_n(s)\, ds + \int_0^t \Phi_m(s)\, dW(s) \end{aligned}$$

for all $t \in [0,T]$ and $m,n \in \mathbb{N}$. Then X and $X_{n,m}$ for all $m,n \in \mathbb{N}$ have a predictable version. Furthermore, by the Fubini theorem,

$$\begin{aligned} &\int_{\Omega_T} \|X(t,\omega) - X_{n,m}(t,\omega)\|_H \,\mathbb{P}_T(dt, d\omega) \\ &= \int_0^T \mathbb{E}\|X(t) - X_{n,m}(t)\|_H \, dt \\ &\le \int_0^T \left[\mathbb{E}\int_0^t \|\varphi(s) - \varphi_n(s)\|_H\, ds + \mathbb{E}\|(\Phi - \Phi_m)\cdot W(s)\|_H\right] dt \\ &\le \int_0^T \left[\|\varphi_n - \varphi\|_{L^1(\Omega_T,\mathcal{P}_T,\mathbb{P}_T;H)} + \left(\mathbb{E}\|(\Phi - \Phi_m)\cdot W(s)\|_H^2\right)^{\frac{1}{2}}\right] dt \\ &= \int_0^T \left[\|\varphi_n - \varphi\|_{L^1(\Omega_T,\mathcal{P}_T,\mathbb{P}_T;H)} + |||\Phi_m - \Phi|||_t\right] dt \\ &\le T\left[\|\varphi_n - \varphi\|_{L^1(\Omega_T,\mathcal{P}_T,\mathbb{P}_T;H)} + |||\Phi_m - \Phi|||_T\right] \longrightarrow 0 \end{aligned}$$

as $m,n \to \infty$. Hence $X_{n,m} \to X$ in $L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)$ as $m,n \to \infty$. Therefore there exists a subsequence such that $X_{n,m}(t,\omega)$ converges pointwise to $X(t,\omega)$ for almost all $(t,\omega) \in \Omega_T$. Thus $F(t, X_{n,m}(t)) \to F(t, X(t))$ as $m,n \to \infty$ almost surely for almost all $t \in [0,T]$ because F is continuous.

Since F_t is Lipschitz continuous with respect to h with integrable Lipschitz constant, i.e., for all $t \in [0,T]$ there exists $L(t) > 0$ such that $|F_t(t,h) - F_t(t,f)| \le L(t)\|h - f\|_H$

for all $f, h \in H$ and $L \in L^1(0,T)$, for all $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\left|\int_0^t F_t(s, X_{n,m}(s))\, ds - \int_0^t F_t(s, X(s))\, ds\right| \\
&\leq \mathbb{E}\int_0^t L(s)\|X_{n,m}(s) - X(s)\|_H\, ds \\
&= \int_0^t L(s)\mathbb{E}\|X_{n,m}(s) - X(s)\|_H\, ds \\
&\leq \|L\|_{L^1(0,T)}\left[\|\varphi_n - \varphi\|_{L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)} + |||\Phi_m - \Phi|||_T\right] \longrightarrow 0
\end{aligned}$$

as $m, n \to \infty$. Hence for all $t \in [0,T]$ there exists a subsequence such that

$$\int_0^t F_t(s, X_{n,m}(s))\, ds \longrightarrow \int_0^t F_t(s, X(s))\, ds$$

as $m, n \to \infty$ almost surely.

Since F_h is bounded, for all $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\left|\int_0^t (F_h(s, X_{n,m}(s)), \varphi_n(s))_H\, ds - \int_0^t (F_h(s, X(s)), \varphi(s))_H\, ds\right| \\
&\leq \mathbb{E}\int_0^t |(F_h(s, X_{n,m}(s)), \varphi_n(s) - \varphi(s))_H|\, ds + \\
&\quad + \mathbb{E}\int_0^t |(F_h(s, X_{n,m}(s)) - F_h(s, X(s)), \varphi(s))_H|\, ds \\
&\leq \mathbb{E}\int_0^t \|F_h(s, X_{n,m}(s))\|_H \|\varphi_n(s) - \varphi(s)\|_H\, ds + \\
&\quad + \mathbb{E}\int_0^t \|F_h(s, X_{n,m}(s)) - F_h(s, X(s))\|_H \|\varphi(s)\|_H\, ds \\
&\leq C\|\varphi_n - \varphi\|_{L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)} + \mathbb{E}\int_0^t \|F_h(s, X_{n,m}(s)) - F_h(s, X(s))\|_H \|\varphi(s)\|_H\, ds.
\end{aligned}$$

Since F_h is continuous and $\varphi \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)$, by Lebesgue's dominated convergence theorem for all $t \in [0,T]$

$$\mathbb{E}\left|\int_0^t (F_h(s, X_{n,m}(s)), \varphi_n(s))_H\, ds - \int_0^t (F_h(s, X(s)), \varphi(s))_H\, ds\right| \longrightarrow 0$$

as $m, n \to \infty$. Hence for all $t \in [0,T]$ there exists a subsequence such that

$$\int_0^t (F_h(s, X_{n,m}(s)), \varphi_n(s))_H\, ds \longrightarrow \int_0^t (F_h(s, X(s)), \varphi(s))_H\, ds$$

as $m, n \to \infty$ almost surely.

Since F_h is bounded, for all $t \in [0,T]$

$$
\begin{aligned}
&\left(\mathbb{E}\left|\int_0^t (F_h(s,X_{n,m}(s)),\Phi_m(s)dW(s))_H - \int_0^t (F_h(s,X(s)),\Phi(s)dW(s))_H\right|^2\right)^{\frac{1}{2}} \\
&= \left(\left|\mathbb{E}\int_0^t (\Phi_m^*(s)F_h(s,X_{n,m}(s)) - \Phi^*(s)F_h(s,X(s)),dW(s))_U\right|^2\right)^{\frac{1}{2}} \\
&= \left(\mathbb{E}\int_0^t \|(\Phi_m^*(s)F_h(s,X_{n,m}(s)) - \Phi^*(s)F_h(s,X(s)),\cdot)_U\|^2_{B_2(U_0,\mathbb{R})}\, ds\right)^{\frac{1}{2}} \\
&\le \left(\mathbb{E}\int_0^t \|(F_h(s,X_{n,m}(s)),(\Phi_m(s)-\Phi(s))\cdot)_U\|^2_{B_2(U_0,\mathbb{R})}\, ds\right)^{\frac{1}{2}} + \\
&\quad + \left(\mathbb{E}\int_0^t \|([F_h(s,X_{n,m}(s)) - F_h(s,X(s))],\Phi(s)\cdot)_U\|^2_{B_2(U_0,\mathbb{R})}\, ds\right)^{\frac{1}{2}} \\
&\le \left(\mathbb{E}\int_0^t \|F_h(s,X_{n,m}(s))\|^2_H\|\Phi_m(s)-\Phi(s)\|^2_{B_2(U_0,H)}\, ds\right)^{\frac{1}{2}} + \\
&\quad + \left(\mathbb{E}\int_0^t \|F_h(s,X_{n,m}(s)) - F_h(s,X(s))\|^2_H\|\Phi(s)\|^2_{B_2(U_0,H)}\, ds\right)^{\frac{1}{2}} \\
&\le C|||\Phi_m - \Phi|||_T + \left(\mathbb{E}\int_0^t \|F_h(s,X_{n,m}(s)) - F_h(s,X(s))\|^2_H\|\Phi(s)\|^2_{B_2(U_0,H)}\, ds\right)^{\frac{1}{2}}.
\end{aligned}
$$

Since F_h is continuous and $\Phi \in \mathcal{N}^2_W(0,T)$, by Lebesgue's dominated convergence theorem for all $t \in [0,T]$

$$
\mathbb{E}\left|\int_0^t (F_h(s,X_{n,m}(s)),\Phi_m(s)dW(s))_H - \int_0^t (F_h(s,X(s)),\Phi(s)dW(s))_H\right|^2 \longrightarrow 0
$$

as $m,n \to \infty$. Hence for all $t \in [0,T]$ there exists a subsequence such that

$$
\int_0^t (F_h(s,X_{n,m}(s)),\Phi_m(s)dW(s))_H \longrightarrow \int_0^t (F_h(s,X(s)),\Phi(s)dW(s))_H
$$

as $m,n \to \infty$ almost surely.

Since $F_{hh}(s,X_{n,m}(s))\Phi_m(s)Q\Phi_m^*(s)$ and $F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s)$ are nuclear operators for all $s \in [0,T]$,

$$
\begin{aligned}
&|\operatorname{Tr}[F_{hh}(s,X_{n,m}(s))\Phi_m(s)Q\Phi_m^*(s)] - \operatorname{Tr}[F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s)]| \\
&= |\operatorname{Tr}[F_{hh}(s,X_{n,m}(s))\Phi_m(s)Q\Phi_m^*(s) - F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s)]| \\
&\le \|F_{hh}(s,X_{n,m}(s))\Phi_m(s)Q\Phi_m^*(s) - F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s)\|_{B_1(H)}.
\end{aligned}
$$

Since F_{hh} is bounded, for all $t \in [0,T]$

$$
\begin{aligned}
&\mathbb{E}\left|\int_0^t \operatorname{Tr}(F_{hh}(s,X_{n,m}(s))\Phi_m(s)Q\Phi_m^*(s))\,ds - \int_0^t \operatorname{Tr}(F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s))\,ds\right| \\
&\le \mathbb{E}\int_0^t \|F_{hh}(s,X_{n,m}(s))\Phi_m(s)Q\Phi_m^*(s) - F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s)\|_{B_1(H)}\,ds \\
&\le \mathbb{E}\int_0^t \|[F_{hh}(s,X_{n,m}(s)) - F_{hh}(s,X(s))]\Phi(s)Q\Phi^*(s)\|_{B_1(H)}\,ds+ \\
&\quad + \mathbb{E}\int_0^t \|F_{hh}(s,X_{n,m}(s))\Phi(s)Q(\Phi_m^*(s) - \Phi^*(s))\|_{B_1(H)}\,ds+ \\
&\quad + \mathbb{E}\int_0^t \|F_{hh}(s,X_{n,m}(s))(\Phi_m(s) - \Phi(s))Q\Phi_m^*(s)\|_{B_1(H)}\,ds \\
&\le \mathbb{E}\int_0^t \|F_{hh}(s,X_{n,m}(s)) - F_{hh}(s,X(s))\|_{B(H)}\|\Phi(s)\|^2_{B_2(U_0,H)}\,ds+ \\
&\quad + \mathbb{E}\int_0^t \|F_{hh}(s,X_{n,m}(s))\|_{B(H)}\|\Phi(s)\|_{B_2(U_0,H)}\|\Phi_m(s) - \Phi(s)\|_{B_2(U_0,H)}\,ds+ \\
&\quad + \mathbb{E}\int_0^t \|F_{hh}(s,X_{n,m}(s))\|_{B(H)}\|\Phi_m(s) - \Phi(s)\|_{B_2(U_0,H)}\|\Phi_m(s)\|_{B_2(U_0,H)}\,ds \\
&\le \mathbb{E}\int_0^t \|F_{hh}(s,X_{n,m}(s)) - F_{hh}(s,X(s))\|_{B(H)}\|\Phi(s)\|^2_{B_2(U_0,H)}\,ds+ \\
&\quad + D(|||\Phi|||_T + |||\Phi_m|||_T)|||\Phi_m - \Phi|||_T.
\end{aligned}
$$

Since $\{\Phi_m\}_{m=1}^{\infty}$ is a convergent sequence in $\mathcal{N}_W^2(0,T)$, it is also bounded. Therefore by Lebesgue's dominated convergence theorem for all $t \in [0,T]$

$$
\mathbb{E}\left|\int_0^t [\operatorname{Tr}(F_{hh}(s,X_{n,m}(s))\Phi_m(s)Q\Phi_m^*(s)) - \operatorname{Tr}(F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s))]\,ds\right| \longrightarrow 0
$$

as $m,n \to \infty$ since F_{hh} is continuous and $\Phi \in \mathcal{N}_W^2(0,T)$. Hence for all $t \in [0,T]$ there exists a subsequence such that

$$
\int_0^t \operatorname{Tr}(F_{hh}(s,X_{n,m}(s))\Phi_m(s)Q\Phi_m^*(s))\,ds \longrightarrow \int_0^t \operatorname{Tr}(F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s))\,ds
$$

as $m,n \to \infty$ almost surely.

Therefore

$$
\begin{aligned}
&F(t,X(t)) - F(0,X_0) \\
&= \int_0^t F_t(s,X(s))\,ds + \int_0^t (F_h(s,X(s)),\varphi(s))_H\,ds+ \\
&\quad + \int_0^t (F_h(s,X(s)),\Phi(s)dW(s))_H + \frac{1}{2}\int_0^t \operatorname{Tr}(F_{hh}(s,X(s))\Phi(s)Q\Phi^*(s))\,ds
\end{aligned}
$$

for almost all $(t,\omega) \in \Omega_T$. Since both sides are continuous with respect to t, the Ito formula is valid almost surely for all $t \in [0,T]$. □

4.6 Linear Equation with Additive Noise

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space and $\{\mathcal{F}_t\}_{t\ge 0}$ a normal filtration. Let $(H,(\cdot,\cdot)_H)$ and $(U,(\cdot,\cdot)_U)$ be real separable Hilbert spaces and $Q \in B(U)$ a positive self-adjoint trace class operator with $\operatorname{Ker} Q = \{0\}$. Let $W(t)$, $t \geq 0$, be a Q-Wiener process in $(\Omega, \mathcal{F}, \mathbb{P})$ with values in U with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. We consider the linear equation

$$\begin{cases} dX(t) = [AX(t) + f(t)]dt + BdW(t), \\ X(0) = X_0 \end{cases} \tag{4.14}$$

where $A : \mathcal{D}(A) \subset H \to H$ and $B : U \to H$ are linear operators and f is an H-valued stochastic process. We assume that A is sectorial and hence generates an analytic semigroup $\{U(t)\}_{t\geq 0}$ in H. In addition, $\mathcal{D}(A)$ is dense in H. Therefore the semigroup $U(t)$ is strongly continuous. The operator B is assumed to be bounded. It is natural to require that $f \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)$ for some $T > 0$, i.e., f is an integrable H-valued predictable process and X_0 is $\mathcal{F}_0$-measurable.

Definition 4.44. *An H-valued predictable process $X(t)$, $t \in [0, T]$, is said to be a (strong) solution to the stochastic initial value problem (4.14) if $X(t, \omega) \in \mathcal{D}(A)$ for almost all $(t, \omega) \in \Omega_T$, $AX \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)$ and for all $t \in [0, T]$*

$$X(t) = X_0 + \int_0^t [AX(s) + f(s)]\, ds + BW(t)$$

almost surely.

A strong solution has a continuous modification by Lemma 4.26 and Theorem 4.38.

We denote

$$f_A(t) := \int_0^t U(t-s)f(s)\, ds \quad \text{and} \quad W_A(t) := \int_0^t U(t-s)B\, dW(s)$$

for all $t \in [0, T]$. The processes f_A and W_A have a great importance in our study of linear equations. The following lemma and proposition present the basic properties of f_A and W_A.

Lemma 4.45. *The process f_A has a predictable version.*

Proof. Since $U(t)$ is strongly continuous, it is measurable from $[0, T]$ to $B(H)$. By Proposition 2.2 there exist $\theta \in \mathbb{R}$ and $M > 0$ such that $\|U(t)\|_{B(H)} \leq Me^{\theta t}$ for all $t > 0$. Thus $U(t - \cdot)f$ is a predictable process on Ω_t for all $t \in [0, T]$ and

$$\begin{aligned} \mathbb{E}\int_0^t \|U(t-s)f(s)\|_H\, ds &\leq \mathbb{E}\int_0^t \|U(t-s)\|_{B(H)}\|f(s)\|_H\, ds \\ &\leq \max\{1, M, Me^{\theta T}\}\mathbb{E}\int_0^t \|f(s)\|_H\, ds \\ &= \max\{1, M, Me^{\theta T}\}\|f\|_{L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)}. \end{aligned}$$

Hence the process f_A is well-defined because the trajectories of $U(t-\cdot)f$ are Bochner integrable almost surely. Furthermore, f_A is adapted. Let $0 \leq s < t \leq T$. Then

$$\begin{aligned} &\mathbb{E}\|f_A(t) - f_A(s)\|_H \\ &= \mathbb{E}\left\|\int_0^s (U(t-r) - U(s-r))f(r)\, dr + \int_s^t U(t-r)f(r)\, dr\right\|_H \\ &\leq \mathbb{E}\int_0^T \chi_{[0,s]}(r)\|U(t-r) - U(s-r)\|_{B(H)}\|f(r)\|_H\, dr + \\ &\quad + \mathbb{E}\int_0^T \chi_{[s,t]}(r)\|U(t-r)\|_{B(H)}\|f(r)\|_H\, dr. \end{aligned}$$

Since U is strongly continuous and $\|U(t)\|_{B(H)} \leq \max(1, M, Me^{\theta T})$ for all $t \in [0,T]$ and $f \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; H)$, by Lebesgue's dominated convergence theorem $\mathbb{E}\|f_A(t) - f_A(s)\|_H \to 0$ as $|t-s| \to 0$. Therefore f_A is stochastically continuous since for all $\varepsilon > 0$ and $\delta > 0$ there exists $\rho > 0$ such that $\mathbb{E}\|f_A(t) - f_A(s)\|_H < \varepsilon\delta$ if $|t-s| < \rho$, and hence

$$\mathbb{P}(\|f_A(t) - f_A(s)\|_H \geq \varepsilon) \leq \frac{\mathbb{E}\|f_A(t) - f_A(s)\|_H}{\varepsilon} < \delta$$

if $|t-s| < \rho$. Therefore f_A has a predictable version by Proposition 4.24. □

The process W_A is called a *stochastic convolution.*

Proposition 4.46. *The process W_A is Gaussian, continuous in mean square and has a predictable version. In addition,*

$$\operatorname{Cov} W_A(t) = \int_0^t U(t-s)BQB^*U^*(t-s)\,ds$$

for all $t \in [0,T]$.

Proof. Since $U(t)$ is strongly continuous, it is measurable from $[0,T]$ to $B(H)$. Furthermore, for all $t \in [0,T]$

$$\begin{aligned}\int_0^t \|U(t-s)B\|^2_{B_2(U_0,H)}\,ds &\leq \int_0^t \|U(t-s)\|^2_{B(H)}\|B\|^2_{B(U,H)} \operatorname{Tr} Q\,ds \\ &\leq \|B\|^2_{B(U,H)} \operatorname{Tr} Q \int_0^t M^2 e^{2\theta(t-s)}\,ds \\ &\leq -\frac{M^2}{2\theta}(1 - e^{2\theta t})\|B\|^2_{B(U,H)} \operatorname{Tr} Q.\end{aligned}$$

Hence $U(t-\cdot)B \in L^2(0,t; B_2(U_0,H)$ for all $t \in [0,T]$. Thus the process W_A is well defined and adapted. Let $0 \leq s < t \leq T$. Then

$$\begin{aligned}&W_A(t) - W_A(s) \\ &= \int_0^s (U(t-r) - U(s-r))B\,dW(r) + \int_s^t U(t-r)B\,dW(r) \\ &= \int_0^T \chi_{[0,s]}(r)(U(t-r) - U(s-r))B\,dW(r) + \int_0^T \chi_{[s,t]}(r)U(t-r)B\,dW(r).\end{aligned}$$

Thus

$$
\begin{aligned}
&\left(\mathbb{E}\|W_A(t) - W_A(s)\|_H^2\right)^{\frac{1}{2}} \\
&\le \left(\mathbb{E}\left\|\int_0^T \chi_{[0,s]}(r)(U(t-r) - U(s-r))B\, dW(r)\right\|_H^2\right)^{\frac{1}{2}} + \\
&\quad + \left(\mathbb{E}\left\|\int_0^T \chi_{[s,t]}(r)U(t-r)B\, dW(r)\right\|_H^2\right)^{\frac{1}{2}} \\
&= \left(\mathbb{E}\int_0^T \|\chi_{[0,s]}(r)(U(t-r) - U(s-r))B\|_{B_2(U_0,H)}^2\, ds\right)^{\frac{1}{2}} + \\
&\quad + \left(\mathbb{E}\int_0^T \|\chi_{[s,t]}(r)U(t-r)B\|_{B_2(U_0,H)}^2\, ds\right)^{\frac{1}{2}} \\
&\le \|B\|_{B(U,H)}\sqrt{\operatorname{Tr} Q}\left(\int_0^T \chi_{[0,s]}(r)\|U(t-r) - U(s-r)\|_{B(H)}^2\, ds\right)^{\frac{1}{2}} + \\
&\quad + \|B\|_{B(U,H)}\sqrt{\operatorname{Tr} Q}\left(\int_0^T \chi_{[s,t]}(r)\|U(t-r)\|_{B(H)}^2\, ds\right)^{\frac{1}{2}}.
\end{aligned}
$$

Since $\|U(t)\|_{B(H)} \le \max(1, M, Me^{\theta T})$ for all $t \in [0,T]$ and U is strongly continuous, by Lebesgue's dominated convergence theorem $\mathbb{E}\|W_A(t) - W_A(s)\|_H^2 \to 0$ as $|t-s| \to 0$. Therefore W_A is mean square continuous. Hence W_A has a predictable version by Lemma 4.20 and Proposition 4.24.

We want to show that for all $n \in \mathbb{N}$ and $t_1, \dots, t_n \in [0,T]$ the H^n-valued random variable $(W_A(t_1), \dots, W_A(t_n))$ is Gaussian. Let $h_1, \dots, h_n \in H$. We need to prove that

$$
((W_A(t_1), \dots, W_A(t_n)), (h_1, \dots, h_n))_{H^n} := \sum_{i=1}^n (W_A(t_i), h_i)_H
$$

is a real valued Gaussian random variable. We may assume that $0 \le t_1 < \ldots < t_n \le T$. Then

$$
\begin{aligned}
\sum_{i=1}^n (W_A(t_i), h_i)_H &= \sum_{i=1}^n \left(\int_0^{t_i} U(t_i - s)B\, dW(s), h_i\right)_H \\
&= \sum_{i=1}^n \left(\sum_{j=1}^i \int_{t_{j-1}}^{t_j} U(t_i - s)B\, dW(s), h_i\right)_H \\
&= \sum_{j=1}^n \left(\int_{t_{j-1}}^{t_j} U(t_j - s)B\, dW(s), \sum_{i=j}^n U^*(t_i - t_j)h_i\right)_H.
\end{aligned}
$$

Since $U(t-\cdot)B \in L^2(0,t; B_2(U_0,H))$ for all $t \in [0,T]$,

$$
\int_s^t U(t-r)B\, dW(r)
$$

is a Gaussian $\mathcal{F}_t$-measurable random variable independent of $\mathcal{F}_s$ for all $0 \le s < t \le T$ by Lemma 4.42. The sum of mutually independent real valued Gaussian random variables is Gaussian. Hence W_A is a Gaussian process. By Lemma 4.42 the covariance of $W_A(t)$ is as claimed. $\square$

Let $X(t)$, $t \in [0,T]$, be a strong solution to the stochastic initial value problem (4.14) and $t \in [0,T]$. Then there exists a sequence $\{t_n\}_{n=1}^{\infty}$ such that $t_n < t$ for all $n \in \mathbb{N}$ and $t_n \to t$ as $n \to \infty$. Let $h \in H$ and $n \in \mathbb{N}$. We define the function $F : [0,t_n] \times H \to \mathbb{R}$ by $F(s,x) = (U(t-s)x, h)_H$. Then F is continuously differentiable with respect to s and twice continuously differentiable with respect to x and

$$\begin{cases} F_s(s,x) = (-AU(t-s)x, h)_H, \\ F_x(s,x) = U^*(t-s)h, \\ F_{xx}(s,x) = 0 \end{cases}$$

since $U(t)$ is strongly continuous, $AU(t)$ is continuous on $(0,\infty)$ and for all $t > 0$

$$\begin{cases} \|U(t)\|_{B(H)} \le Me^{\theta t}, \\ \|AU(t)\|_{B(H)} \le Ct^{-1}e^{(\theta+1)t} \end{cases}$$

for some $\theta \in \mathbb{R}$, $M > 0$ and $C > 0$ according to Proposition 2.2. Furthermore, F_s is uniformly continuous on bounded subsets of $[0,t_n] \times H$ and Lipschitz continuous with respect to x with Lipschitz constant

$$L(s) = C\|h\|_H e^{(\theta+1)(t-s)}(t-s)^{-1},$$

which is integrable on $[0,t_n]$, and F_x is bounded. Then by the Ito formula,

$$\begin{aligned} F(t_n, X(t_n)) = F(0,X_0) &+ \int_0^{t_n} (F_x(s,X(s)), BdW(s))_H + \\ &+ \int_0^{t_n} F_s(s,X(s))\,ds + \int_0^{t_n} (F_x(s,X(s)), AX(s) + f(s))_H\,ds \\ &= (U(t)X_0, h)_H + \int_0^{t_n} (U^*(t-s)h, AX(s) + f(s))_H\,ds + \\ &+ \int_0^{t_n} (-AU(t-s)X(s), h)_H\,ds + \int_0^{t_n} (U^*(t-s)h, BdW(s))_H \\ &= (U(t)X_0, h)_H + \int_0^{t_n} (U(t-s)AX(s) - AU(t-s)X(s), h)_H\,ds + \\ &+ \int_0^{t_n} (U(t-s)f(s), h)_H\,ds + \int_0^{t_n} (U(t-s)BdW(s), h)_H \end{aligned}$$

almost surely. Since $AU(t)x = U(t)Ax$ for all $x \in \mathcal{D}(A)$ and $X(t,\omega) \in \mathcal{D}(A)$ for almost all $(t,\omega) \in \Omega_T$,

$$\begin{aligned} &(U(t-t_n)X(t_n), h)_H \\ &= \left(U(t)X_0 + \int_0^{t_n} U(t-s)f(s)\,ds + \int_0^{t_n} U(t-s)BdW(s), h \right)_H \end{aligned}$$

almost surely. Thus

$$U(t-t_n)X(t_n) = U(t)X_0 + \int_0^{t_n} U(t-s)f(s)\,ds + \int_0^{t_n} U(t-s)BdW(s)$$

almost surely. Since the strong solution has a continuous modification, the analytic semigroup is strongly continuous and the integrals are continuous processes by Lemma 4.26 and Theorem 4.38,

$$X(t) = U(t)X_0 + \int_0^t U(t-s)f(s)\,ds + \int_0^t U(t-s)BdW(s)$$

for all $t \in [0,T]$ almost surely.

Theorem 4.47. *Under the above assumptions if the stochastic initial value problem (4.14) has a strong solution, it is given by the formula*

$$X(t) = U(t)X_0 + \int_0^t U(t-s)f(s)\,ds + \int_0^t U(t-s)B\,dW(s) \tag{4.15}$$

for all $t \in [0,T]$ almost surely.

By Lemma 4.45 and Proposition 4.46 the right hand side of (4.15) has a predictable modification. It is natural to consider Process (4.15) as a generalized solution to the stochastic initial value problem (4.14) even if it is not the strong solution in the sense of Definition 4.44.

Definition 4.48. *The predictable process given by the formula*

$$X(t) = U(t)X_0 + \int_0^t U(t-s)f(s)\,ds + \int_0^t U(t-s)B\,dW(s)$$

for all $t \in [0,T]$ almost surely is called the weak solution to the stochastic initial value problem (4.14).

Chapter 5

Complete Electrode Model

In electrical impedance tomography (EIT) electric currents are applied to electrodes on the surface of an object and the resulting voltages are measured using the same electrodes. If the conductivity distribution inside the object is known, the forward problem of EIT is to calculate the electrode potentials corresponding to given electrode currents. In this chapter we introduce the most realistic model for the EIT, the complete electrode model (CEM). It takes into account the electrodes on the surface of the object as well as contact impedances between the object and electrodes. The existence and uniqueness of the weak solution to the complete electrode model in bounded domains has been shown in the article [48]. Usually in applications the requirement of the boundedness of the object is fulfilled. Since we are interested in electrical impedance process tomography and assume that the pipeline is infinitely long, we need the analogous result in unbounded domains. Because of the state estimation approach to the electrical impedance process tomography problem we examine the Fréchet differentiability of the electrode potentials with respect to the conductivity distribution. The results concerning unbounded domains are made by the author.

5.1 Complete Electrode Model in Bounded Domains

Let D be a bounded domain in $\mathbb{R}^n$, $n \geq 2$, with a smooth boundary ∂D and σ a conductivity distribution in D. We assume that $\sigma \in L^\infty(\bar{D})$, i.e., σ is essentially bounded in the domain D up to the boundary. To the surface of the body D we attach L electrodes. We identify the electrode with the part of the surface it contacts. These subsets of ∂D we denote by e_l for all $1 \leq l \leq L$. The electrodes e_l are assumed to be open connected subsets of ∂D for all $1 \leq l \leq L$ whose closures are disjoint. In the case $n \geq 3$ we assume that the boundaries of electrodes are smooth curves on ∂D. Through these electrodes we inject current into the body and on the same electrodes we measure the resulting voltages. The current applied to the electrode e_l is marked with I_l for all $1 \leq l \leq L$. We call a vector $I := (I_1, \ldots, I_L)^T$ of L currents a *current pattern* if it satisfies the *conservation of charge condition* $\sum_{l=1}^{L} I_l = 0$. The corresponding *voltage pattern* we denote by $U := (U_1, \ldots, U_L)^T$. We choose the ground or reference potential so that $U_1 = 0$. If the voltage pattern U instead of the current pattern I were given, the electric potential u in the interior of the domain

D would satisfy the boundary value problem

$$\nabla \cdot \sigma \nabla u = 0 \quad \text{in } D, \tag{5.1}$$

$$u + z_l \sigma \frac{\partial u}{\partial \nu} = U_l \quad \text{on } e_l,\ 1 \leq l \leq L, \tag{5.2}$$

$$\sigma \frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial D \setminus \cup_{l=1}^{L} e_l \tag{5.3}$$

where $z_l \in \mathbb{R}_+$ is the contact impedance on the electrode e_l for all $1 \leq l \leq L$ and ν is the exterior unit normal on ∂D. We denote $z := (z_1, \dots, z_L)^T$. The weak solution to the boundary value problem (5.1)–(5.3) is defined to be the solution $u \in H^1(D)$ to the variational problem

$$\int_D \sigma(x) \nabla u(x) \cdot \nabla v(x)\, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} u(x) v(x)\, dS(x) = \sum_{l=1}^{L} \frac{1}{z_l} U_l \int_{e_l} v(x)\, dS(x)$$

for all $v \in H^1(D)$ with appropriate assumptions on the conductivity σ and contact impedances z. The corresponding current pattern would be given by

$$I_l = \int_{e_l} \sigma \frac{\partial u}{\partial \nu}\, dS$$

for all $1 \leq l \leq L$. Since we want to inject current and measure voltage, the boundary value problem we are interested in is

$$\nabla \cdot \sigma \nabla u = 0 \quad \text{in } D, \tag{5.4}$$

$$\sigma \frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial D \setminus \cup_{l=1}^{L} e_l, \tag{5.5}$$

$$\int_{e_l} \sigma \frac{\partial u}{\partial \nu}\, dS = I_l, \quad 1 \leq l \leq L \tag{5.6}$$

when the current pattern I is known. Since the boundary value problem (5.4)–(5.6) does not have a unique solution, we add an extra boundary condition, namely

$$u + z_l \sigma \frac{\partial u}{\partial \nu} = U_l \quad \text{on } e_l,\ 1 \leq l \leq L. \tag{5.7}$$

The boundary value problem (5.4)–(5.7) is called the *complete electrode model*. We assume that the conductivity distribution and contact impedances are known. For a given current pattern I the solution to the complete electrode model contains the electric potential u in the interior of the body as well as L surface potentials U. We are looking for the solution from the space $H := H^1(D) \oplus \mathbb{R}^L$. In the article [48] it has been shown that the complete electrode model has the variational formulation

$$B((u,U),(v,V)) = \sum_{l=1}^{L} I_l V_l \tag{5.8}$$

for all $(v, V) \in H$ where $B : H \times H \to \mathbb{R}$ is the bilinear form

$$B((u,U),(v,V)) := \int_D \sigma(x) \nabla u(x) \cdot \nabla v(x)\, dx + \sum_{l=1}^{L} \frac{1}{z_l} \int_{e_l} (u(x) - U_l)(v(x) - V_l)\, dS(x)$$

for all $(u, U), (v, V) \in H$. We notice that if $B((u, U), (u, U)) = 0$, then $u = U_1 = \ldots = U_L =$ constant. Hence the variational problem (5.8) for all $(v, V) \in H$ cannot have a unique solution in H. We can always add a constant to the solution. Thus we need to choose the ground potential.

Theorem 5.1. [48, Theorem 3.3 and Corollary 3.4] *Let us assume that there are strictly positive constants σ_0, σ_1 and Z such that $\sigma_0 \leq \sigma(x) \leq \sigma_1$ for almost all $x \in \bar{D}$ and $z_l > Z$ for all $1 \leq l \leq L$. Then for a given current pattern $(I_l)_{l=1}^L \in \mathbb{R}^L$ there exists a unique $(u, U) \in H$ satisfying*

$$B((u,U),(v,V)) = \sum_{l=1}^{L} I_l V_l$$

for all $(v, V) \in H$ if the ground potential is chosen such a way that $U_1 = 0$.

In the article [48] it is assumed that σ is continuously differentiable in the domain D up to the boundary, i.e., $\sigma \in C^1(\bar{D})$ and the ground potential is chosen such a way that $\sum_{l=1}^L U_l = 0$. Nevertheless, in the proof of Theorem 5.1 the assumption of continuous differentiability is not required and any appropriate choice of the ground potential ensures the uniqueness. Hence by Theorem 5.1 for all current patterns the complete electrode model has a unique weak solution in H if it is assumed that the ground potential is chosen such a way that $U_1 = 0$.

5.2 Complete Electrode Model in Unbounded Domains

Let D be an unbounded domain in $\mathbb{R}^n$, $n \geq 2$. We use the same notation as above. All assumptions made in Section 5.1 are expected to be valid also in this section. In addition, we suppose that the electrode e_l is a bounded subset of ∂D for all $1 \leq l \leq L$. We are interested in such a weak solution to the complete electrode model that the electric potential u is locally square integrable function and its weak derivatives are square integrable, i.e., $u \in L^2_{\text{loc}}(D)$ and $\nabla u \in L^2(D; \mathbb{R}^n)$. The local integrability is needed in unbounded domains since we allow the electric potential to be constant and the only square integrable constant function in unbounded domains is the zero function. Our aim is to prove that the complete electrode model has a unique weak solution. We need to modify the definition of the solution space H. Let K be a bounded connected open subset of D such that ∂K is smooth and $\cup_{l=1}^L e_l \subset \partial K$. We define the norm $\|\cdot\|_{\mathcal{H}^1_K(D)}$ by

$$\|\varphi\|_{\mathcal{H}^1_K(D)} := \left(\int_D \|\nabla\varphi(x)\|^2_{\mathbb{R}^n}\, dx + \int_K |\varphi(x)|^2\, dx \right)^{\frac{1}{2}}$$

for all $\varphi \in C_0^\infty(\mathbb{R}^n)$ and the space $\mathcal{H}^1_K(D)$ to be the closure of the set $\{\varphi|_D : \varphi \in C_0^\infty(\mathbb{R}^n)\}$ in the norm $\|\cdot\|_{\mathcal{H}^1_K(D)}$. Then $u \in L^2(K)$ and $\nabla u \in L^2(D; \mathbb{R}^n)$ for all $u \in \mathcal{H}^1_K(D)$. We denote $\mathcal{H} := \mathcal{H}^1_K(D) \oplus \mathbb{R}^L$ and

$$\|(u,U)\|_{\mathcal{H}} := \left(\|u\|^2_{\mathcal{H}^1_K(D)} + \|U\|^2_{\mathbb{R}^L} \right)^{\frac{1}{2}}$$

for all $(u, U) \in \mathcal{H}$. We define another norm in $\mathcal{H}$ by

$$\|(u,U)\|_\circ := \left(\|\nabla u\|^2_{L^2(D;\mathbb{R}^n)} + \sum_{l=1}^{L} \int_{e_l} |u(x) - U_l|^2\, dS(x) + |U_1|^2 \right)^{\frac{1}{2}}$$

for all $(u, U) \in \mathcal{H}$. The norms $\|\cdot\|_{\mathcal{H}}$ and $\|\cdot\|_\circ$ are equivalent.

Lemma 5.2. *There exist constants* $0 < \lambda \leq \Lambda < \infty$ *such that*

$$\lambda \|(u, U)\|_\circ \leq \|(u, U)\|_\mathcal{H} \leq \Lambda \|(u, U)\|_\circ$$

for all $(u, U) \in \mathcal{H}$.

Proof. Let $(u, U) \in \mathcal{H}$. By examining the norm $\|\cdot\|_\circ$ we obtain that

$$\begin{aligned}
\|(u, U)\|_\circ^2 &= \|\nabla u\|^2_{L^2(D;\mathbb{R}^n)} + \sum_{l=1}^{L} \int_{e_l} |u(x) - U_l|^2 \, dS(x) + |U_1|^2 \\
&\leq \|\nabla u\|^2_{L^2(D;\mathbb{R}^n)} + 2\sum_{l=1}^{L} \int_{e_l} |u(x)|^2 \, dS(x) + 2\sum_{l=1}^{L} m(e_l)|U_l|^2 + |U_1|^2 \\
&\leq \|\nabla u\|^2_{L^2(D;\mathbb{R}^n)} + 2\|u\|^2_{L^2(\partial K)} + \left(2 \max_{1 \leq l \leq L} (m(e_l)) + 1\right) \|U\|^2_{\mathbb{R}^L}
\end{aligned}$$

where $m(e_l)$ is the area of the l^{th} electrode for all $1 \leq l \leq L$. By the continuous imbedding $H^{\frac{1}{2}}(\partial K) \subset L^2(\partial K)$ and the trace theorem,

$$\|u\|_{L^2(\partial K)} \leq \|u\|_{H^{\frac{1}{2}}(\partial K)} \leq C\|u\|_{H^1(K)} \leq C\|u\|_{\mathcal{H}^1_K(D)}$$

for all $u \in \mathcal{H}^1_K(D)$. Hence

$$\|(u, U)\|_\circ^2 \leq C\left(\|u\|^2_{\mathcal{H}^1_K(D)} + \|U\|^2_{\mathbb{R}^L}\right) = C\|(u, U)\|^2_\mathcal{H}.$$

Thus the first part of the claim is proved.

Let us assume that there does not exist a constant $\Lambda > 0$ such that $\|(u, U)\|_\mathcal{H} \leq \Lambda\|(u, U)\|_\circ$ for all $(u, U) \in \mathcal{H}$. We pick a sequence $\{(u^n, U^n)\}_{n=1}^\infty \subset \mathcal{H}$ such that $\|(u^n, U^n)\|_\mathcal{H} = 1$ and $\|(u^n, U^n)\|_\circ < \frac{1}{n}$ for all $n \in \mathbb{N}$. Then $\{u^n\}_{n=1}^\infty$ is a bounded sequence in $H^1(K)$ since

$$\|u^n\|_{H^1(K)} \leq \|u^n\|_{\mathcal{H}^1_K(D)} \leq \|(u^n, U^n)\|_\mathcal{H} = 1$$

for all $n \in \mathbb{N}$. By the compact imbedding theorem of Sobolev spaces over bounded domains there exists a subsequence $\{u^{n_k}\}_{k=1}^\infty$ such that $u^{n_k} \to u$ in $L^2(K)$ as $k \to \infty$ for some $u \in L^2(K)$. However, because for all $n \in \mathbb{N}$

$$\|\nabla u^n\|_{L^2(K;\mathbb{R}^n)} \leq \|\nabla u^n\|_{L^2(D;\mathbb{R}^n)} \leq \|(u^n, U^n)\|_\circ < \frac{1}{n},$$

$\{u^{n_k}\}_{k=1}^\infty$ is a Cauchy sequence in $H^1(K)$. Hence $u^{n_k} \to u$ in $H^1(K)$ as $k \to \infty$. Since $\nabla u^n \to 0$ in $L^2(D;\mathbb{R}^n)$ as $n \to \infty$, the limit u satisfies $\nabla u = 0$ in K, i.e., $u = \text{constant} = c$ in K. If we define u to be the constant c in D, then $u^{n_k} \to u$ in $\mathcal{H}^1_K(D)$ as $k \to \infty$. In addition,

$$\begin{aligned}
\int_{e_l} |u^n(x) - U_l^n|^2 \, dS(x) &= \int_{e_l} |u^n(x) - c - (U_l^n - c)|^2 \, dS(x) \\
&= \int_{e_l} |u^n(x) - c|^2 \, dS(x) + |U_l^n - c|^2 m(e_l) + \\
&\quad - 2\left(U_l^n - c\right) \int_{e_l} \left(u^n(x) - c\right) \, dS(x) \\
&\geq -2|U_l^n - c| \int_{e_l} |u^n(x) - c| \, dS(x) + |U_l^n - c|^2 m(e_l)
\end{aligned}$$

for all $1 \leq l \leq L$ and $n \in \mathbb{N}$. Since $\|(u^n, U^n)\|_\circ < \frac{1}{n}$,

$$m(e_l)|U_l^n - c|^2 < \frac{1}{n^2} + 2|U_l^n - c| \int_{e_l} |u^n(x) - c| \, dS(x)$$

for all $1 \leq l \leq L$ and $n \in \mathbb{N}$. Since

$$\begin{aligned} |U_l^n - c| \int_{e_l} |u^n(x) - c| \, dS(x) &\leq (|U_l^n| + |c|) \sqrt{m(e_l)} \left(\int_{e_l} |u^n(x) - c|^2 \, dS(x) \right)^{\frac{1}{2}} \\ &\leq (1 + |c|) \sqrt{m(e_l)} \|u^n - c\|_{L^2(\partial K)} \\ &\leq C \, (1 + |c|) \sqrt{m(e_l)} \|u^n - c\|_{H^1(K)}, \end{aligned}$$

we get

$$m(e_l)|U_l^n - c|^2 < \frac{1}{n^2} + 2C \, (1 + |c|) \sqrt{m(e_l)} \|u^n - c\|_{H^1(K)}$$

for all $1 \leq l \leq L$ and $n \in \mathbb{N}$. Thus $U_l^{n_k}$ converges to c as $k \to \infty$ for all $1 \leq l \leq L$. Since $|U_1^n| \leq \|(u^n, U^n)\|_\circ < \frac{1}{n}$ for all $n \in \mathbb{N}$, we have $U_1^n \to 0$ as $n \to \infty$. Thus $c = 0$. This is a contradiction because

$$1 = \|(u^{n_k}, U^{n_k})\|_{\mathcal{H}}^2 = \|u^{n_k}\|_{\mathcal{H}_K^1(D)}^2 + \|U^{n_k}\|_{\mathbb{R}^L}^2 \longrightarrow 0$$

as $k \to \infty$. Hence the second part of the claim is valid. □

Since the norms $\|\cdot\|_{\mathcal{H}}$ and $\|\cdot\|_\circ$ are equivalent and the norm $\|\cdot\|_\circ$ does not depend on the set K, the Banach space $\mathcal{H}$ is independent of K, despite of the definition. If $(u, U) \in \mathcal{H}$, then u is square integrable in all bounded connected open subsets K of D such that ∂K is smooth and $\cup_{l=1}^L e_l \subset \partial K$. Therefore u is locally square integrable in D. Thus we may expect to find a weak solution to the complete electrode model from the space $\mathcal{H}$. Since we want to choose the ground potential such a way that $U_1 = 0$, we are looking for the weak solution to the complete electrode model from a subspace of $\mathcal{H}$, namely $\mathcal{H}_0 := \{(u, U) \in \mathcal{H} : U_1 = 0\}$.

Theorem 5.3. *Let us assume that there are strictly positive constants σ_0, σ_1 and Z such that $\sigma_0 \leq \sigma(x) \leq \sigma_1$ for almost all $x \in \bar{D}$ and $z_l > Z$ for all $1 \leq l \leq L$. Then for a given current pattern $(I_l)_{l=1}^L \in \mathbb{R}^L$ there exists a unique $(u, U) \in \mathcal{H}_0$ satisfying*

$$B((u, U), (v, V)) = \sum_{l=1}^L I_l V_l$$

for all $(v, V) \in \mathcal{H}_0$.

Proof. The norm $\|\cdot\|_\circ$ defines an inner product $(\cdot, \cdot)_\circ$ in $\mathcal{H}$ in a natural way. Hence $\mathcal{H}$ is a Hilbert space. Since the subspace $\mathcal{H}_0$ is closed, it is a Hilbert space. Furthermore,

$$\|(u, U)\|_\circ^2 = \|\nabla u\|_{L^2(D;\mathbb{R}^n)}^2 + \sum_{l=1}^L \int_{e_l} |u(x) - U_l|^2 \, dS(x)$$

for all $(u, U) \in \mathcal{H}_0$. We want to use the Lax-Milgram lemma [54, Theorem III.7] in $\mathcal{H}_0$. The form B is coercive in $\mathcal{H}_0$ since

$$\begin{aligned} |B((u, U), (u, U))| &= \int_D \sigma(x) \|\nabla u(x)\|_{\mathbb{R}^n}^2 \, dx + \sum_{l=1}^L \frac{1}{z_l} \int_{e_l} |u(x) - U_l|^2 \, dS(x) \\ &\geq \min\left(\sigma_0, \frac{1}{z_1}, \ldots, \frac{1}{z_L} \right) \|(u, U)\|_\circ^2 \end{aligned}$$

for all $(u, U) \in \mathcal{H}_0$. By Hölder's inequality,

$$
\begin{aligned}
&|B((u,U),(v,V))| \\
&\leq \int_D \sigma(x)|\nabla u(x)\cdot\nabla v(x)|\,dx + \sum_{l=1}^{L}\frac{1}{z_l}\int_{e_l}|u(x)-U_l||v(x)-V_l|\,dS(x) \\
&\leq \sigma_1\left(\int_D \|\nabla u(x)\|_{\mathbb{R}^n}^2\,dx\right)^{\frac{1}{2}}\left(\int_D \|\nabla v(x)\|_{\mathbb{R}^n}^2\,dx\right)^{\frac{1}{2}}+ \\
&\quad+\frac{1}{Z}\sum_{l=1}^{L}\left(\int_{e_l}|u(x)-U_l|^2\,dS(x)\right)^{\frac{1}{2}}\left(\int_{e_l}|v(x)-V_l|^2\,dS(x)\right)^{\frac{1}{2}} \qquad (5.9)\\
&\leq \sigma_1\|\nabla u\|_{L^2(D;\mathbb{R}^n)}\|\nabla v\|_{L^2(D;\mathbb{R}^n)}+ \\
&\quad+\frac{1}{Z}\left(\sum_{l=1}^{L}\int_{e_l}|u(x)-U_l|^2\,dS(x)\right)^{\frac{1}{2}}\left(\sum_{l=1}^{L}\int_{e_l}|v(x)-V_l|^2\,dS(x)\right)^{\frac{1}{2}} \\
&\leq \left(\sigma_1+\frac{1}{Z}\right)\|(u,U)\|_\circ\|(v,V)\|_\circ
\end{aligned}
$$

for all $(u, U), (v, V) \in \mathcal{H}_0$. Hence the bilinear form B is bounded in $\mathcal{H}_0$. Therefore the form B fulfills the assumptions of the Lax-Milgram lemma. We need to show that the right hand side of the variational formulation (5.8) is a linear continuous mapping for all current patterns. The linear mapping

$$f : \mathcal{H}_0 \to \mathbb{R}, \quad (v, V) \mapsto \sum_{l=1}^{L} I_l V_l$$

is well defined for all current patterns $(I_l)_{l=1}^L \in \mathbb{R}^L$. Let $(v, V) \in \mathcal{H}_0$. Then

$$|f(v, V)| \leq \|I\|_{\mathbb{R}^L}\|V\|_{\mathbb{R}^L} \leq \|I\|_{\mathbb{R}^L}\|(v, V)\|_{\mathcal{H}} \leq \Lambda\|I\|_{\mathbb{R}^L}\|(v, V)\|_\circ.$$

Hence the mapping f is continuous. Thus by the Lax-Milgram lemma there exists a unique element in $\mathcal{H}_0$ satisfying the variational formula (5.8) for all $(v, V) \in \mathcal{H}_0$. □

The form B is independent of K. Hence for all current patterns $(I_l)_{l=1}^L \in \mathbb{R}^L$ the unique $(u, U) \in \mathcal{H}_0$ satisfying the variational formulation (5.8) for all $(v, V) \in \mathcal{H}_0$ does not depend on the choice of K.

Corollary 5.4. *Let us assume that the hypotheses of Theorem 5.3 are satisfied. Then there exists a unique $(u, U) \in \mathcal{H}_0$ satisfying the variational formulation (5.8) for all $(v, V) \in \mathcal{H}$.*

Proof. Let (u, U) be the unique element in $\mathcal{H}_0$ satisfying the variational formulation (5.8) for all $(v, V) \in \mathcal{H}_0$ given by Theorem 5.3. For an arbitrary $(v, V) \in \mathcal{H}$ we define $(w, W) := (v - V_1, V - V_1)$. Then $(w, W) \in \mathcal{H}_0$. Hence $B((u, U), (w, W)) = \sum_{l=1}^L I_l W_l$. Since $\sum_{l=1}^L I_l = 0$ and $B((u, U), (c, (c, \ldots, c)^T)) = 0$ for all $c \in \mathbb{R}$ and $(u, U) \in \mathcal{H}$, we obtain $B((u, U), (v, V)) = \sum_{l=1}^L I_l V_l$ for all $(v, V) \in \mathcal{H}$. Since $B(u, U), (v, V)) = 0$ for all $(v, V) \in \mathcal{H}$ if and only if $u = U_1 = \ldots = U_L =$ constant, (u, U) is the unique element in $\mathcal{H}_0$ satisfying the variational formulation (5.8) for all $(v, V) \in \mathcal{H}$. □

We still need to prove that the solution to the variational problem (5.8) for all $(v, V) \in \mathcal{H}$ satisfies the complete electrode model.

Lemma 5.5. *If $(u, U) \in \mathcal{H}$ satisfies the variational formulation (5.8) for all $(v, V) \in \mathcal{H}$, then (u, U) also satisfies the complete electrode model (5.4)–(5.7).*

Proof. If $(u, U) \in \mathcal{H}$ satisfies the variational formulation (5.8) for all $(v, V) \in \mathcal{H}$, Equations (5.4)–(5.7) are obtained by considering particular choices of $(v, V) \in \mathcal{H}$. Let $v \in C_0^\infty(D)$ and $V = 0$. Then the variational formulation (5.8) is

$$\int_D \sigma(x)\nabla u(x) \cdot \nabla v(x)\, dx = 0.$$

Hence in the weak sense u satisfies $\nabla \cdot \sigma \nabla u = 0$ in D. Let $v \in C_0^\infty(\mathbb{R}^n)$ and $R > 0$ be such a constant that $\operatorname{supp} v \subset B(0, R)$. Then by Green's formula,

$$\begin{aligned}
0 &= \int_D v(x)\nabla \cdot \sigma(x)\nabla u(x)\, dx = \int_{D\cap B(0,R)} v(x)\nabla \cdot \sigma(x)\nabla u(x)\, dx \\
&= \int_{\partial(D\cap B(0,R))} \sigma(x)\frac{\partial u(x)}{\partial \nu} v(x)\, dS(x) - \int_{D\cap B(0,R)} \sigma(x)\nabla u(x) \cdot \nabla v(x)\, dx \\
&= \int_{\partial D} \sigma(x)\frac{\partial u(x)}{\partial \nu} v(x)\, dS(x) - \int_D \sigma(x)\nabla u(x) \cdot \nabla v(x)\, dx.
\end{aligned}$$

Therefore

$$\int_D \sigma(x)\nabla u(x) \cdot \nabla v(x)\, dx = \int_{\partial D} \sigma(x)\frac{\partial u(x)}{\partial \nu} v(x)\, dS(x)$$

for all $v \in C_0^\infty(\mathbb{R}^n)$. If $v \in C_0^\infty(\mathbb{R}^n)$ and $V = 0$,

$$\begin{aligned}
0 &= \int_D \sigma(x)\nabla u(x) \cdot \nabla v(x)\, dx + \sum_{l=1}^{L} \frac{1}{z_l}\int_{e_l} (u(x) - U_l)v(x)\, dS(x) \\
&= \int_{\partial D} \left(\sigma(x)\frac{\partial u(x)}{\partial \nu} + \sum_{l=1}^{L} \frac{1}{z_l}\chi_{e_l}(x)(u(x) - U_l)\right) v(x)\, dS(x).
\end{aligned}$$

Since $C_0^\infty(\partial D)$ is dense in $L^2(\partial D)$,

$$\sigma\frac{\partial u}{\partial \nu} + \sum_{l=1}^{L} \frac{1}{z_l}\chi_{e_l}(u - U_l) = 0 \tag{5.10}$$

on ∂D. Hence (u, U) satisfies the boundary conditions (5.5) and (5.7). Let $v \in C_0^\infty(\mathbb{R}^n)$ and $V \in \mathbb{R}^L$. Then

$$\begin{aligned}
\sum_{l=1}^{L} I_l V_l &= \int_{\partial D} \sigma(x)\frac{\partial u(x)}{\partial \nu} v(x)\, dS(x) + \sum_{l=1}^{L} \frac{1}{z_l}\int_{e_l} (u(x) - U_l)(v(x) - V_l)\, dS(x) \\
&= \int_{\partial D} \left(\sigma(x)\frac{\partial u(x)}{\partial \nu} + \sum_{l=1}^{L} \frac{1}{z_l}\chi_{e_l}(x)(u(x) - U_l)\right) v(x)\, dS(x) + \\
&\quad - \sum_{l=1}^{L} \frac{1}{z_l} V_l \int_{e_l} (u(x) - U_l)\, dS(x) \\
&= \sum_{l=1}^{L} \frac{1}{z_l}\left(m(e_l)U_l - \int_{e_l} u(x)\, dS(x)\right) V_l.
\end{aligned}$$

Since $V \in \mathbb{R}^L$ is arbitrary,

$$\frac{1}{z_l}\left(m(e_l)U_l - \int_{e_l} u(x)\, dS(x)\right) = I_l$$

for all $1 \le l \le L$. Then by the boundary condition (5.10),

$$\int_{e_l} \sigma \frac{\partial u}{\partial \nu}\, dS(x) = \frac{1}{z_l}\int_{e_l} (U_l - u(x))\, dS(x) = I_l$$

for all $1 \le l \le L$. Hence (u, U) satisfies the complete electrode model. □

Definition 5.6. *For all current patterns the weak solution to the complete electrode model (5.4)–(5.7) in unbounded domains is the solution to the variational problem (5.8) for all for all $(v, V) \in \mathcal{H}$ given by Corollary 5.4.*

In the article [48] the proof of the existence and uniqueness of the weak solution to the complete electrode model in bounded domains is done by using the quotient space $H/\mathbb{R}$. The same procedure would also work in unbounded domains by replacing H with $\mathcal{H}$. Since the choice of the ground potential is essential for the uniqueness, we wanted to restrict ourselves to the subspace $\mathcal{H}_0$ and hence avoid the quotient space $\mathcal{H}/\mathbb{R}$. It seems to be the natural way to solve the problem.

5.3 The Fréchet Differentiability of U

In the forward problem of EIT we are interested in the surface potentials U. The electric potential u in the interior of the domain D is needed in the mathematical formulation of the problem. By Corollary 5.4 there exists a function which maps the conductivity distribution, contact impedances and current pattern to the corresponding voltage pattern, i.e., $(\sigma, z, I) \mapsto U$ if the ground potential is chosen such a way that $U_1 = 0$. We want to show that this mapping is Fréchet differentiable with respect to the conductivity distribution σ. In Theorem 5.3 it is assumed that $\sigma \in L^\infty(\bar{D})$ and there are strictly positive constants σ_0, σ_1 and Z such that $\sigma_0 \le \sigma(x) \le \sigma_1$ for almost all $x \in \bar{D}$ and $z_l > Z$ for all $1 \le l \le L$. We define the subset $\Sigma(D)$ of $L^\infty(\bar{D})$ by

$$\begin{aligned}\Sigma(D) := \{\sigma \in L^\infty(\bar{D}) : &\text{ there are strictly positive constants } \sigma_0 \text{ and } \sigma_1 \\ &\text{ such that } \sigma_0 \le \sigma(x) \le \sigma_1 \text{ for almost all } x \in \bar{D}\}.\end{aligned}$$

Then $(\sigma, 1/z) \in \Sigma(D) \oplus \mathbb{R}^L_+$ if and only if σ and z satisfy the assumptions of Theorem 5.3. If $(\sigma, z) \in \Sigma(D) \oplus \mathbb{R}^L_+$, we denote

$$\begin{aligned}&B_{\sigma,z}((u, U), (v, V)) \\ &:= \int_D \sigma(x)\nabla u(x) \cdot \nabla v(x)\, dx + \sum_{l=1}^{L} z_l \int_{e_l} (u(x) - U_l)(v(x) - V_l)\, dS(x)\end{aligned}$$

for all $(u, U), (v, V) \in \mathcal{H}$.

Theorem 5.7. *Let $(I_l)_{l=1}^L \in \mathbb{R}^L$ be a current pattern. The mapping*

$$\mathcal{M} : \Sigma(D) \oplus \mathbb{R}^L_+ \to \mathcal{H}_0, \quad (\sigma, z) \mapsto (u, U)$$

where (u, U) is the solution the variational problem

$$B_{\sigma,z}((u,U),(v,V)) = \sum_{l=1}^{L} I_l V_l$$

for all $(v, V) \in \mathcal{H}$ is Fréchet differentiable. The derivative $\mathcal{M}'(\sigma, z)$ satisfies the following equation: Let $(s,\zeta) \in L^\infty(\bar{D}) \oplus \mathbb{R}^L$. Then $\mathcal{M}'(\sigma, z)(s,\zeta) =: (w, W) \in \mathcal{H}_0$ is the solution to the variational problem

$$\begin{aligned} &B_{\sigma,z}((w,W),(v,V)) \\ &= -\int_D s(x)\nabla u^0(x)\cdot\nabla v(x)\,dx - \sum_{l=1}^{L}\zeta_l\int_{e_l}(u^0(x)-U_l^0)(v(x)-V_l)\,dS(x) \end{aligned} \tag{5.11}$$

for all $(v, V) \in \mathcal{H}$ where $(u^0, U^0) := \mathcal{M}(\sigma, z)$.

The Fréchet differentiability of a mapping $(\sigma, z) \mapsto [u, U]$ where $[u, U] \in H/\mathbb{R}$ is the solution the variational problem

$$B_{\sigma,z}([u,U],[v,V]) = \sum_{l=1}^{L} I_l V_l$$

for all $[v, V] \in H/\mathbb{R}$ is shown in the article [18] with the assumption that the domain D is bounded and the conductivity distribution is piecewise continuous.

Proof of Theorem 5.7. If $(\sigma, z) \in \Sigma(D) \oplus \mathbb{R}^L_+$, by Corollary 5.4 the variational problem

$$B_{\sigma,z}((u,U),(v,V)) = \sum_{l=1}^{L} I_l V_l$$

for all $(v, V) \in \mathcal{H}$ has a unique solution $(u, U) \in \mathcal{H}_0$. Hence the mapping $\mathcal{M}$ is well defined.

Let $(\sigma, z) \in \Sigma(D) \oplus \mathbb{R}^L_+$ and $(s,\zeta) \in L^\infty(\bar{D}) \oplus \mathbb{R}^L$. We denote $(u^0, U^0) := \mathcal{M}(\sigma, z)$. We notice that

$$\begin{aligned} &\int_D s(x)\nabla u^0(x)\cdot\nabla v(x)\,dx + \sum_{l=1}^{L}\zeta_l\int_{e_l}(u^0(x)-U_l^0)(v(x)-V_l)\,dS(x) \\ &= B_{s,\zeta}((u^0,U^0),(v,V)) \end{aligned}$$

for all $(v, V) \in \mathcal{H}$. Since $(s,\zeta) \in L^\infty(\bar{D}) \oplus \mathbb{R}^L$, by Inequality (5.9),

$$|B_{s,\zeta}((u^0,U^0),(v,V))| \le \left(\|s\|_{L^\infty(\bar{D})} + \|\zeta\|_{l^\infty}\right)\left\|(u^0,U^0)\right\|_\circ \|(v,V)\|_\circ$$

for all $(v, V) \in \mathcal{H}_0$. Thus the right hand side of (5.11) is a continuous linear mapping from $\mathcal{H}_0$ to $\mathbb{R}$. By the Lax-Milgram lemma there exists a unique $(w, W) \in \mathcal{H}_0$ such that (w, W) satisfies the variational formulation (5.11) for all $(v, V) \in \mathcal{H}_0$. Similarly to the proof of Corollary 5.4 we can show that (w, W) satisfies the variational formulation (5.11) for all $(v, V) \in \mathcal{H}$. Thus the mapping $T_{\sigma,z} : (s,\zeta) \mapsto (w, W)$ is well defined. Obviously, $T_{\sigma,z}$ is linear. We define the norm in $L^\infty(\bar{D}) \oplus \mathbb{R}^L$ to be

$$\|(s,\zeta)\|_{L^\infty(\bar{D})\oplus\mathbb{R}^L} := \|s\|_{L^\infty(\bar{D})} + \|\zeta\|_{l^\infty}$$

for all $(s,\zeta) \in L^\infty(\bar{D}) \oplus \mathbb{R}^L$. Hence the operator $T_{\sigma,z}$ is also bounded since by the coercivity of the form $B_{\sigma,z}$ in $\mathcal{H}_0$,

$$\begin{aligned}
\|(w,W)\|_\circ^2 &\le C(\sigma,z)|B_{\sigma,z}((w,W),(w,W))| \\
&\le C(\sigma,z)\Big[\int_D |s(x)||\nabla u^0(x)\cdot\nabla w(x)|\,dx + \\
&\qquad + \sum_{l=1}^L |\zeta_l| \int_{e_l} |u^0(x) - U_l^0||w(x) - W_l|\,dS(x)\Big] \\
&\le C(\sigma,z)\left(\|s\|_{L^\infty(\bar{D})} + \|\zeta\|_{l^\infty}\right)\left\|(u^0,U^0)\right\|_\circ \|(w,W)\|_\circ.
\end{aligned}$$

We want to show that $T_{\sigma,z}$ is the Fréchet derivative of the mapping $\mathcal{M}$ in the point $(\sigma,z) \in \Sigma(D) \oplus \mathbb{R}_+^L$. The set $\Sigma(D) \oplus \mathbb{R}_+^L$ is an open subset of $L^\infty(\bar{D}) \oplus \mathbb{R}^L$ in the norm $\|\cdot\|_{L^\infty(\bar{D})} + \|\cdot\|_{l^\infty}$. Let

$$\|(s,\zeta)\|_{L^\infty(\bar{D})\oplus\mathbb{R}^L} < \frac{1}{2}\min\left(\left\|\frac{1}{\sigma}\right\|_{L^\infty(\bar{D})}^{-1}, z_1,\ldots,z_L\right).$$

Then $(\sigma+s, z+\zeta) \in \Sigma(D) \oplus \mathbb{R}_+^L$. Let us denote $(u,U) := \mathcal{M}(\sigma+s, z+\zeta)$. We know that for all $(v,V) \in \mathcal{H}$

$$B_{\sigma,z}((u^0,U^0),(v,V)) = \sum_{l=1}^L I_l V_l = B_{\sigma+s,z+\zeta}((u,U),(v,V)).$$

Thus for all $(v,V) \in \mathcal{H}$

$$\begin{aligned}
&B_{\sigma,z}((u-u^0, u-U^0),(v,V)) \\
&= -\int_D s(x)\nabla u(x)\cdot\nabla v(x)\,dx - \sum_{l=1}^L \zeta_l \int_{e_l} (u(x)-U_l)(v(x)-V_l)\,dS(x).
\end{aligned}$$

Hence

$$\begin{aligned}
&B_{\sigma,z}((u-u^0-w, U-U^0-W),(v,V)) \\
&= -\int_D s(x)\nabla(u-u^0)(x)\cdot\nabla v(x)\,dx + \\
&\quad - \sum_{l=1}^L \zeta_l \int_{e_l} ((u-u^0)(x) - (U_l - U_l^0))(v(x)-V_l)\,dS(x)
\end{aligned}$$

for all $(v,V) \in \mathcal{H}$. Therefore by the coercivity of the form $B_{\sigma,z}$ in $\mathcal{H}_0$,

$$\begin{aligned}
&\left\|(u-u^0-w, U-U^0-W)\right\|_\circ^2 \\
&\le C(\sigma,z)|B_{\sigma,z}((u-u^0-w,U-U^0-W),(u-u^0-w,U-U^0-W))| \\
&\le C(\sigma,z)\Big[\int_D |s(x)||\nabla(u-u^0)(x)\cdot\nabla(u-u^0-w)(x)|\,dx + \\
&\qquad + \sum_{l=1}^L |\zeta_l| \int_{e_l} |(u-u^0)(x) - (U_l-U_l^0)| \times \\
&\qquad\qquad \times |(u-u^0-w)(x) - (U_l - U_l^0 - W_l)|\,dS(x)\Big] \\
&\le C(\sigma,z)\left(\|s\|_{L^\infty(\bar{D})} + \|\zeta\|_{l^\infty}\right)\left\|(u,U)-(u^0,U^0)\right\|_\circ \left\|(u,U)-(u^0,U^0)-(w,W)\right\|_\circ.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\left\|(u-u^0, U-U^0)\right\|_\circ^2 \\
&\le C(\sigma,z)|B_{\sigma,z}((u-u^0,U-U^0),(u-u^0,U-U^0))| \\
&\le C(\sigma,z)\Bigg[\int_D |s(x)||\nabla u(x)\cdot\nabla(u-u^0)(x)|\,dx+ \\
&\qquad\qquad + \sum_{l=1}^L |\zeta_l| \int_{e_l} |u(x)-U_l||(u-u^0)(x)-(U_l-U_l^0)|\,dS(x)\Bigg] \\
&\le C(\sigma,z)\left(\|s\|_{L^\infty(\bar D)}+\|\zeta\|_{l^\infty}\right)\|(u,U)\|_\circ \left\|(u,U)-(u^0,U^0)\right\|_\circ.
\end{aligned}$$

Hence

$$\left\|(u-u^0-w, U-U^0-W)\right\|_\circ \le C(\sigma,z)\left(\|s\|_{L^\infty(\bar D)}+\|\zeta\|_{l^\infty}\right)^2\|(u,U)\|_\circ.$$

Since (u,U) depends on (s,ζ), we need to estimate its norm. By the coercivity of the form $B_{\sigma+s,z+\zeta}$ in $\mathcal{H}_0$,

$$\begin{aligned}
\|(u,U)\|_\circ^2 &\le C(\sigma,s,z,\zeta)|B_{\sigma+s,z+\zeta}((u,U),(u,U))| \\
&= C(\sigma,s,z,\zeta)|B_{\sigma,z}((u^0,U^0),(u,U))| \\
&\le C(\sigma,s,z,\zeta)\left\|(u^0,U^0)\right\|_\circ\|(u,U)\|_\circ.
\end{aligned}$$

Therefore

$$\left\|(u-u^0-w, U-U^0-W)\right\|_\circ \le C(\sigma,s,z,\zeta)\left\|(u^0,U^0)\right\|_\circ\left(\|s\|_{L^\infty(\bar D)}+\|\zeta\|_{l^\infty}\right)^2$$

where the constant $C(\sigma,s,z,\zeta)$ is of the form

$$C(\sigma,s,z,\zeta)=C(\sigma,z)\max\left(\left\|\frac{1}{\sigma+s}\right\|_{L^\infty(\bar D)},\frac{1}{z_1+\zeta_1},\dots,\frac{1}{z_L+\zeta_L}\right).$$

Thus

$$\begin{aligned}
&\frac{\|\mathcal{M}(\sigma+s,z+\zeta)-\mathcal{M}(\sigma,z)-T_{\sigma,z}(s,\zeta)\|_\circ}{\|s\|_{L^\infty(\bar D)}+\|\zeta\|_{l^\infty}} \\
&\le C(\sigma,s,z,\zeta)\|\mathcal{M}(\sigma,z)\|_\circ\left(\|s\|_{L^\infty(\bar D)}+\|\zeta\|_{l^\infty}\right)\longrightarrow 0
\end{aligned}$$

as $\|(s,\zeta)\|_{L^\infty(\bar D)\oplus\mathbb{R}^L}\to 0$. Hence $T_{\sigma,z}$ is the Fréchet derivative of $\mathcal{M}$ at the point $(\sigma,z)\in\Sigma(D)\oplus\mathbb{R}^L_+$. □

We define the projection $\pi:\mathcal{H}\to\mathbb{R}^L$ by $(u,U)\mapsto U$ for all $(u,U)\in\mathcal{H}$.

Corollary 5.8. *Let $(I_l)_{l=1}^L\in\mathbb{R}^L$ be a current pattern. The mapping*

$$U:\Sigma(D)\oplus\mathbb{R}^L_+\to\mathbb{R}^L,\quad (\sigma,z)\mapsto U(\sigma,z)$$

where $U(\sigma,z)=\pi\mathcal{M}(\sigma,z)$ is Fréchet differentiable and

$$U'(\sigma,z)=\pi\mathcal{M}'(\sigma,z)$$

for all $(\sigma,z)\in\Sigma(D)\oplus\mathbb{R}^L_+$.

Proof. By Theorem 5.7 the mapping $\mathcal{M}$ is Fréchet differentiable. Since the projection π is a bounded linear operator, the mapping U is Fréchet differentiable. The Fréchet derivative of U is obtained from the definition. □

Chapter 6

Statistical Inversion Theory

In realistic measurements we have directly observable quantities and others that cannot be observed. If some of the unobservable quantities are of our primary interest, we are dealing with an inverse problem. The interdependence of the quantities in the measurement setting is described through mathematical models. In the statistical inversion theory it is assumed that all quantities included in the model are represented by random variables. The randomness describes our degree of knowledge concerning their realizations. Our information about their values is coded into their distributions. The solution to the inverse problem is the posterior distribution of the random variables of interest after performing the measurements. We introduce the basic concepts of the statistical inversion theory. The Bayes theorem of inverse problems and Bayesian filtering method are presented. As an example of non-stationary inverse problems we study the electrical impedance process tomography problem. We view it as a state estimation problem. A discretized state estimation system is the goal of this chapter. Sections 6.1 and 6.2 are based on the book of Kaipio and Somersalo [19]. The results concerning electrical impedance process tomography (Section 6.3) are made by the author.

6.1 The Bayes Formula

In realistic measurement setting we are able to measure only a finitely many values of the directly observable quantities. For example, the measurement frame in electrical impedance tomography consists of all linearly independent injected current patterns and the corresponding set of voltage measurements. These measured values are called the data. From the data we want to compute the values of the quantities of primary interest. Usually this sort of problems are underdetermined. Hence we are able to compute only partly the quantities of primary interest. Furthermore, in numerical implementations we need to discretize our model for the measurement process. Therefore there exist only finitely many variables describing the quantities of primary interest. Thus in statistical approach to inverse problems we may assume that random variables in a model have values in $\mathbb{R}^n$ with some $n \in \mathbb{N}$. In addition, we suppose that the distributions of the random variables are absolutely continuous with respect to the Lebesgue measure. This requirement is not necessary but since we restrict ourselves to Gaussian random variables, it is acceptable. Hence the distributions of the random variables are determined by their probability densities. We denote random variables by capital letters and their realizations by lower case letters.

The statistical inversion theory is based on the *Bayes formula*. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let X and Y be random variables with values in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. We suppose that the random variable X is unobservable and of our primary interest and Y is directly observable. We call X the *unknown*, Y the *measurement* and its realization y in the actual measurement process the *data*. We assume that before performing the measurement of Y we have some information about the random variable X. This prior knowledge is coded into the probability density $x \mapsto \pi_{\mathrm{pr}}(x)$ called the *prior density*. In addition, we suppose that after analysing the measurement setting as well as all additional information available about the random variables we have found the joint probability density of X and Y denoted by $\pi(x, y)$. On the other hand, if we knew the value of the unknown, the conditional probability density of Y given this information would be

$$\pi(y \mid x) = \frac{\pi(x, y)}{\pi_{\mathrm{pr}}(x)}$$

if $\pi_{\mathrm{pr}}(x) \neq 0$. The conditional probability density of Y is called the *likelihood function* because it expresses the likehood of different measurement outcomes with given $X = x$. We assume finally that the measurement data $Y = y$ is given. The conditional probability density

$$\pi(x \mid y) = \frac{\pi(x, y)}{\pi(y)}$$

if $\pi(y) = \int_{\mathbb{R}^m} \pi(x, y)\ dx \neq 0$, is called the *posterior density* of X. This density expresses what we know about X after the observation $Y = y$. In the Bayesian framework the inverse problem can be formulated as follows: *Given the data $Y = y$, find the conditional probability density $\pi(x \mid y)$ of the variable X.* We summarize the notation and results in the following theorem, which can be referred to as the *Bayes theorem of inverse problems.*

Theorem 6.1. *Let the random variable X with values in $\mathbb{R}^n$ have a known prior probability density $\pi_{\mathrm{pr}}(x)$ and the data consists of the observed value y of the observable random variable Y with values in $\mathbb{R}^m$ such that $\pi(y) > 0$. Then the posterior probability density of X given the data y is*

$$\pi_{\mathrm{post}}(x) = \pi(x \mid y) = \frac{\pi_{\mathrm{pr}}(x)\pi(y \mid x)}{\pi(y)}. \tag{6.1}$$

The marginal density

$$\pi(y) = \int_{\mathbb{R}^m} \pi(x, y)\ dx = \int_{\mathbb{R}^m} \pi_{\mathrm{pr}}(x)\pi(y \mid x)\ dx$$

plays the role of a normalising constant and is usually of little importance. By looking at the Bayes formula (6.1) solving an inverse problem may be broken into three subtasks: (1) based on all prior information of the unknown X find a prior probability density π_{pr}, (2) find the likelihood function $\pi(y \mid x)$ that describes the interrelation between the observation and the unknown and (3) develop methods to explore the posterior probability density.

6.2 Nonstationary Inverse Problems

In several applications one encounters a situation in which measurements that constitute the data of an inverse problem are done in a nonstationary environment. More precisely, it may happen that the physical quantities that are in the focus of our primary interest are time dependent and the measured data depends on these quantities at different time instants. Inverse problems of this type are called *nonstationary inverse problems*. In some applications the time evolution model of the quantities of primary interest is given by a stochastic differential equation. The electrical impedance process tomography problem is an example of nonstationary inverse problems. The concentration distribution in a pipeline is time varying and the EIT measurement frame depends on the concentration distribution at different time instants. The time evolution of the concentration distribution is given by the stochastic convection–diffusion equation.

6.2.1 State Estimation

Often non-stationary inverse problems are viewed as a *state estimation problem*. Let $D \subset \mathbb{R}^n$ be a domain that corresponds to the object of interest. We denote by $X = X(t, x)$, $x \in D$, a distributed parameter describing the state of the object – the unknown distribution of a physical target – at time $t \geq 0$. We assume that we have a model for the time evolution of the parameter X. We suppose that instead of being a deterministic function X is a stochastic process satisfying a stochastic differential equation. This allows us to incorporate phenomena such as modelling uncertainties into the model. Let $Y = Y(t)$ denote a quantity that is directly observable at time $t \geq 0$. We assume that the dependence of Y upon the state X is known up to observation noise and modelling errors. The state estimation system consists of a pair of equations

$$dX(t) = F(t, X, R)dt + dW(t), \tag{6.2}$$

$$Y(t) = G(t, X, S). \tag{6.3}$$

Equation (6.2) is called the *state evolution equation* and is to be interpreted as a stochastic differential equation in which the function F is the evolution model function and $R = R(t)$ and $W = W(t)$ are stochastic processes. The processes R and W may represent modelling errors and uncertainties in the time evolution model. Equation (6.3) is called the *observation equation*. The function G is the observation model function and $S = S(t)$ is a stochastic process. The process S describes modelling errors and noise in the measurement process. The evolution and observation model functions are known and allowed to be nonlinear. The state estimation problem can be formulated as follows: *Estimate the state X satisfying an evolution equation of the type (6.2) based on the observed values of Y.* To be able to estimate the state X we have to solve a stochastic differential equation and represent the state evolution equation in a more useful form. Estimators of the state X are calculated by taking conditional expectation with respect to the measurements. The most commonly used estimator is the filter $\mathbb{E}(X(t)|Y(s), s \leq t)$ which is based on the current history of the measurement process.

Usually the measurements are done at discrete time instants. Hence a discrete state evolution and observation equations are needed. They may be derived from the continuous ones, especially if the evolution and observation model functions are linear. Since the computation requires space discretization, we need discretized versions of the state evolution and observation equations. Then we have two discrete stochastic processes $\{X_k\}_{k=0}^{\infty}$ and $\{Y_k\}_{k=1}^{\infty}$ with values in finite dimensional spaces and the state estimation system

$$\begin{aligned} X_{k+1} &= F_{k+1}(t_0, t_1, \ldots, t_{k+1}, X_0, X_1, \ldots, X_k, W_1, \ldots, W_{k+1}), \quad k \in \mathbb{N}_0, \\ Y_k &= G_k(t_k, X_0, X_1, \ldots, X_k, S_1, \ldots, S_k), \quad k \in \mathbb{N} \end{aligned}$$

where $\{W_k\}_{k=1}^{\infty}$ and $\{S_k\}_{k=1}^{\infty}$ are discrete stochastic processes and are called the *state* and *observation noise processes*, respectively.

6.2.2 Bayesian Filtering

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\{X_k\}_{k=0}^{\infty}$ and $\{Y_k\}_{k=1}^{\infty}$ be two discrete stochastic processes. The random variable X_k with values in $\mathbb{R}^{n_k}$ for $k \in \mathbb{N}_0$ represents the quantities we are primarily interested in and is called the *state vector*. The random variable Y_k with values in $\mathbb{R}^{m_k}$ for $k \in \mathbb{N}$ represents the measurement. We refer to it as the *observation* at the k$^{\text{th}}$ time instant. We assume that the distributions of the random variables are absolutely continuous with respect to the Lebesgue measure. We postulate the following four properties of these processes:

1. The process $\{X_k\}_{k=0}^{\infty}$ is a Markov process, i.e.,

$$\pi(x_{k+1} \mid x_0, x_1, \ldots, x_k) = \pi(x_{k+1} \mid x_k)$$

   for all $k \in \mathbb{N}_0$.

2. The process $\{Y_k\}_{k=1}^{\infty}$ is a Markov process with respect to the history of the process $\{X_k\}_{k=0}^{\infty}$, i.e.,

$$\pi(y_k \mid x_0, x_1, \ldots, x_k) = \pi(y_k \mid x_k)$$

   for all $k \in \mathbb{N}$.

3. The process $\{X_k\}_{k=0}^{\infty}$ depends on the past observations only through its own history, i.e.,

$$\pi(x_{k+1} \mid x_0, x_1, \ldots, x_k, y_1, y_2, \ldots, y_k) = \pi(x_{k+1} \mid x_k)$$

   for all $k \in \mathbb{N}$.

4. The process $\{Y_k\}_{k=1}^{\infty}$ depends on the past observations only through the history of the process $\{X_k\}_{k=0}^{\infty}$, i.e.,

$$\pi(y_{k+1} \mid x_0, x_1, \ldots, x_{k+1}, y_1, y_2, \ldots, y_k) = \pi(y_{k+1} \mid x_{k+1})$$

   for all $k \in \mathbb{N}$.

If the stochastic processes $\{X_k\}_{k=0}^{\infty}$ and $\{Y_k\}_{k=1}^{\infty}$ satisfy the condition 1–4 above, we call this pair an *evolution–observation model*. The evolution–observation model is completely specified if we know the probability density of the initial state X_0, Markov transition kernels $\pi(x_{k+1} \mid x_k)$ for all $k \in \mathbb{N}_0$ and likelihood functions $\pi(y_k \mid x_k)$ for all $k \in \mathbb{N}$. Both the Markov transition kernels $\pi(x_{k+1} \mid x_k)$ and likelihood functions $\pi(y_k \mid x_k)$ are allowed to vary in time. An example of evolution–observation models is the state estimation system

$$\begin{aligned} X_{k+1} &= F_{k+1}(X_k, W_{k+1}), \quad k \in \mathbb{N}_0, \\ Y_k &= G_k(X_k, S_k), \quad k \in \mathbb{N} \end{aligned}$$

if the state and observation noise processes satisfy the following assumptions. For all $k \neq l$ the noise vectors W_k and W_l as well as S_k and S_l are mutually independent and also mutually independent of the initial state X_0. In addition, the noise vectors W_k and S_l are mutually independent for all $k \in \mathbb{N}_0$ and $l \in \mathbb{N}$.

The inverse problem considered is to extract information of the state vectors X_k based on measurements Y_k for all $k \in \mathbb{N}$. In the Bayesian approach we try to get the posterior distribution of the state vector conditioned on the observations. Let us denote $D_k := \{y_1, y_2, \dots, y_k\}$ for all $k \in \mathbb{N}$. The conditional probability density of the state vector X_k conditioned on all the measurements $y_1, \dots, y_l$ is denoted by $\pi(x_k \mid D_l) := \pi(x_k \mid y_1, \dots, y_l)$ for all $k \in \mathbb{N}_0$ and $l \in \mathbb{N}$. Additionally, $\pi(x_k \mid D_0) := \pi(x_k)$ for all $k \in \mathbb{N}_0$.

Definition 6.2. *Let the stochastic processes $\{X_k\}_{k=0}^{\infty}$ and $\{Y_k\}_{k=1}^{\infty}$ form an evolution–observation model. The problem of determining the conditional probability density $\pi(x_{k+1} \mid D_k)$ for $k \in \mathbb{N}_0$ is called a prediction problem and $\pi(x_k \mid D_k)$ for $k \in \mathbb{N}$ a filtering problem.*

Often the prediction problem is just an intermediate step for the filtering problem. To be able to solve the state estimation problem we need to derive a recursive updating scheme where the evolution and observation updates alternate. In this type of recursive scheme the state evolution equation is used for solving the prediction problem from the filtering problem of the previous time level while the new observation is used to update the predicted probability density. Therefore we need to find formulas for the following updating steps:

1. Evolution updating: Given $\pi(x_k \mid D_k)$, find $\pi(x_{k+1} \mid D_k)$ based on the Markov transition kernel $\pi(x_{k+1} \mid x_k)$ for $k \in \mathbb{N}_0$.
2. Observation updating: Given $\pi(x_{k+1} \mid D_k)$, find $\pi(x_{k+1} \mid D_{k+1})$ based on the new observation y_{k+1} and likelihood function $\pi(y_{k+1} \mid x_{k+1})$ for $k \in \mathbb{N}_0$.

The updating formulas are given in the following theorem.

Theorem 6.3. [19, Theorem 4.2] *Let us assume that stochastic processes $\{X_k\}_{k=0}^{\infty}$ and $\{Y_k\}_{k=1}^{\infty}$ form an evolution–observation model. Then for all $k \in \mathbb{N}_0$*

(i) the evolution updating formula is

$$\pi(x_{k+1} \mid D_k) = \int_{\mathbb{R}^{n_k}} \pi(x_{k+1} \mid x_k)\pi(x_k \mid D_k)\, dx_k, \tag{6.4}$$

(ii) the observation updating formula is

$$\pi(x_{k+1} \mid D_{k+1}) = \frac{\pi(y_{k+1} \mid x_{k+1})\pi(x_{k+1} \mid D_k)}{\pi(y_{k+1} \mid D_k)} \tag{6.5}$$

where

$$\pi(y_{k+1} \mid D_k) = \int_{\mathbb{R}^{n_{k+1}}} \pi(y_{k+1} \mid x_{k+1})\pi(x_{k+1} \mid D_k)\, dx_{k+1}.$$

The integrand on the right hand side of (6.4) is simply the joint probability density of the variables X_k and X_{k+1} conditioned on the observations D_k. Hence Formula (6.4) is the conditional marginal probability density of X_{k+1}. We consider the probability density $\pi(x_{k+1} \mid D_k)$ as the prior density of X_{k+1} when the new observation y_{k+1} arrives. Then Equation (6.5) is nothing other than the Bayes formula. Therefore this method is called the *Bayesian filtering method.* If the joint probability densities of the variables X_k and X_{k+1} as well as X_{k+1} and Y_{k+1} conditioned on the observations D_k are Gaussian, the evolution and observation updating formulas (6.4) and (6.5) can be derived by the following theorem.

Theorem 6.4. [19, Theorems 3.5 and 3.6] *Let $X : \Omega \to \mathbb{R}^n$ and $Y : \Omega \to \mathbb{R}^m$ be two Gaussian random variables whose joint probability density $\pi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ is of the form*

$$\pi(x, y) \propto \exp\left(-\frac{1}{2}\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix}^{-1} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}\right)$$

where $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}^m$, $\Gamma_{11} \in \mathbb{R}^{n\times n}$, $\Gamma_{22} \in \mathbb{R}^{m\times m}$ and

$$\Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix}$$

is a positive definite symmetric $(n + m) \times (n + m)$ matrix. Then

(i) the marginal density of X is

$$\pi(x) = \int_{\mathbb{R}^m} \pi(x, y)\, dy \propto \exp\left(-\frac{1}{2}(x - x_0)^T \Gamma_{11}^{-1}(x - x_0)\right),$$

(ii) the density of X conditioned on $Y = y$ is

$$\pi(x \mid y) \propto \exp\left(-\frac{1}{2}(x - \bar{x})^T \widetilde{\Gamma}_{22}^{-1}(x - \bar{x})\right)$$

where $\bar{x} = x_0 + \Gamma_{12}\Gamma_{22}^{-1}(y - y_0)$ and $\widetilde{\Gamma}_{22} = \Gamma_{11} - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21}$.

6.3 Electrical Impedance Process Tomography

We examine a concentration distribution of a given substance in a fluid moving in a pipeline by doing electromagnetic measurements at the boundary of the pipe. In electrical impedance tomography (EIT) electric currents are applied to electrodes on the surface of an object and the resulting voltages are measured using the same electrodes. A complete set of measurements consists of all possible linearly independent injected current patterns and the corresponding set of voltage measurements. The conductivity distribution inside the object is reconstructed based on the voltage measurements. The relation between the conductivity and concentration depends on the process and is usually non-linear. In process tomography we cannot in general assume that the target remains unaltered during a full set of measurements. The time evolution of the concentration distribution needs to be modeled properly. We view the problem as a state estimation problem. The concentration distribution is treated as a stochastic process that satisfies a stochastic differential equation referred to as the state evolution equation. The measurements are described in terms of an observation equation containing the measurement noise.

Let D be an infinitely long pipe $\{x = (x_1, x') \in \mathbb{R}^d : |x'| < r\}$ with $d \geq 2$ and $r > 0$. Let $\kappa = \kappa(x)$ be the diffusion coefficient of the substance of our interest and $\mathbf{v} = \mathbf{v}(x)$ the velocity of the flow for all $x \in \bar{D}$. The diffusion coefficient and velocity distribution are assumed to be known and stationary. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We assume that the concentration distribution $C(t)$ is a stochastic process satisfying the stochastic differential equation

$$dC(t) = [LC(t) + f(t)]dt + dW(t) \tag{6.6}$$

for every $t > 0$ with the initial value $C(0) = C_0$. The operator L is the deterministic convection–diffusion operator

$$\begin{aligned} L : \mathcal{D}(L) &\to L^2(D) \\ c &\mapsto \nabla \cdot (\kappa \nabla c) - \mathbf{v} \cdot \nabla c \end{aligned} \tag{6.7}$$

with the domain

$$\mathcal{D}(L) = \left\{ c \in H^2(D) : \frac{\partial c}{\partial \nu}\bigg|_{\partial D} = 0 \right\}. \tag{6.8}$$

The boundary condition at the boundary of the pipe is included in the domain of the operator L. We assume that there is no diffusion through the pipe walls. We model with f a possible control of the system. We assume that $f(t)$, $t \geq 0$, is an $L^2(D)$-valued stochastic process. The term $dW(t)$ is a source term representing possible modelling errors where $W(t)$, $t \geq 0$, is an $L^2(D)$-valued Wiener process.

We assume that on the surface of the pipe there are L electrodes. We identify the electrode with the part of the surface it contacts. We denote these subsets of ∂D by e_l for all $1 \leq l \leq L$. At time $t > 0$ an electric current $I_l(t)$ is applied to the electrode e_l and the resulting voltage $U_l(t)$ is measured using the same electrode for all $1 \leq l \leq L$. We describe the electric potential $u(t, x)$ inside the pipe and *voltage pattern* $U(t) := (U_1(t), \ldots, U_L(t))^T$ by the *complete electrode model*

$$\nabla \cdot \sigma \nabla u = 0 \quad \text{in } D \tag{6.9}$$

$$u + z_l \sigma \frac{\partial u}{\partial \nu} = U_l \quad \text{on } e_l,\ 1 \leq l \leq L \tag{6.10}$$

$$\sigma \frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial D \setminus \cup_{l=1}^{L} e_l \tag{6.11}$$

$$\int_{e_l} \sigma \frac{\partial u}{\partial \nu}\, dS = I_l, \quad 1 \leq l \leq L \tag{6.12}$$

where $\sigma = \sigma(t, x)$ is the conductivity distribution inside the pipe and z_l is the contact impedance on the electrode e_l for all $1 \leq l \leq L$. We assume that the *current pattern*

$I(t) := (I_1(t), \ldots, I_L(t))^T$ satisfies the *conservation of charge condition* $\sum_{l=1}^L I_l(t) = 0$ and the ground potential is chosen such a way that $U_1(t) = 0$. In Chapter 5 it was shown that under appropriate regularity assumptions on the conductivity distribution, electrodes and contact impedances the complete electrode model (6.9)–(6.12) has a unique weak solution (u, U). We suppose that the contact impedances are known positive numbers. By the uniqueness of the solution to the complete electrode model the mapping $\sigma(t) \mapsto U(\sigma(t); I(t))$ from $\Sigma(D)$ to $\mathbb{R}^L$ is well defined for all current patterns $I(t)$ where $\Sigma(D)$ is a subset of $L^\infty(\bar{D})$ defined by $\Sigma(D) := \{\sigma \in L^\infty(\bar{D}) : 0 < \sigma_0 \leq \sigma(x) \leq \sigma_1 < \infty$ for almost all $x \in \bar{D}\}$. The interdependence of the conductivity and concentration distributions is usually non-linear, i.e., there exists a non-linear function $g : L^2(D) \to \Sigma(D)$ such that $\sigma(t) = g(C(t))$ for all $t > 0$. In addition, we assume that the measurement noise is additive. Then the observation equation is

$$V(t) = U(g(C(t)); I(t)) + S(t)$$

where $S(t)$, $t \geq 0$, is an $\mathbb{R}^L$-valued stochastic process independent of the process $C(t)$, $t \geq 0$. Hence the voltage pattern depends non-linearly on the concentration distribution.

We assume that the measurements are done in time instants $0 < t_1 < \ldots < t_n$. We use the notation $I^k := I(t_k)$, $C_k := C(t_k)$, $S^k := S(t_k)$ and $V^k := V(t_k)$ for all $k = 1, \ldots, n$. The state estimation system concerning the electrical impedance process tomography problem is

$$\begin{aligned} dC(t) &= [LC(t) + f(t)]dt + dW(t), \quad && t > 0, \\ V^k &= U(g(C_k); I^k) + S^k, && k = 1, \ldots, n. \end{aligned}$$

We are interested in a real-time monitoring for the flow. Therefore we should be able to solve the filtering problem $\mathbb{E}(C_k \mid V^l,\ l \leq k)$ for all $k = 1, \ldots, n$. For that reason we need to solve the stochastic convection–diffusion equation (6.6) and to present the discrete evolution equation for the concentration distribution.

6.3.1 Analytic Semigroup

According to Section 4.6 to be able to solve the stochastic convection–diffusion equation we need to show that under certain assumptions the operator L defined by (6.7) and (6.8) generates a strongly continuous analytic semigroup. We use the theory introduced in Chapters 2 and 3. Since the boundary of D is $\{x = (x_1, x') \in \mathbb{R}^d : |x'| = r\}$, it is C^∞-smooth. Hence we want to know the requirements of the coefficient functions κ and $\mathbf{v}$ such that the realization of the operator $\mathcal{A} = \nabla \cdot \kappa \nabla - \mathbf{v} \cdot \nabla$ generates an analytic semigroup if the boundary condition is defined by the operator $\mathcal{B} = \boldsymbol{\nu} \cdot \nabla$ where $\boldsymbol{\nu}$ is the exterior unit normal on the boundary ∂D. We modify the operators $\mathcal{A}$ and $\mathcal{B}$ into the form used in Chapter 3. Then

$$\mathcal{A}(x, \partial) = \sum_{i=1}^n \kappa(x)\partial_i^2 + \sum_{i=1}^n [\partial_i \kappa(x) - v_i(x)]\partial_i$$

and

$$\mathcal{B}(x, \partial) = \sum_{i=1}^n \nu_i(x)\partial_i$$

where $\mathbf{v} = (v_1, \ldots, v_n)^T$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^T$. In Chapter 3 we assumed that the coefficient functions of the operator $\mathcal{A}$ are real uniformly continuous and bounded and the coefficient functions of the operator $\mathcal{B}$ belong to the space $UC^1(\bar{D})$. Hence the functions κ and $\mathbf{v}$ have to fulfill the conditions

$$\begin{cases} \kappa : \bar{D} \to \mathbb{R}, & \kappa \in UC^1(\bar{D}), \\ \mathbf{v} : \bar{D} \to \mathbb{R}^d, & \mathbf{v} \in UC(\bar{D}). \end{cases}$$

The operator $\mathcal{A}$ has to satisfy the ellipticity condition (3.2). The principal part of $\mathcal{A}$ is $\sum_{i=1}^n \kappa(x)\partial_i^2$. Let $\xi \in \mathbb{R}^d$ and $x \in \bar{D}$. Then

$$\sum_{i=1}^n \kappa(x)\xi_i^2 = \kappa(x)|\xi|^2 \geq \inf_{x \in \bar{D}} \kappa(x)|\xi|^2.$$

Hence the function κ has to be bounded from below. We assume that there exists $\delta > 0$ such that $\kappa(x) \geq \delta$ for all $x \in \bar{D}$. The operator $\mathcal{B}$ has to fulfill the uniform nontangentiality condition (3.4). The first order terms of $\mathcal{B}$ are $\sum_{i=1}^n \nu_i(x)\partial_i$. Therefore Condition (3.4) is valid since

$$\sum_{i=1}^n \nu_i^2(x) = |\boldsymbol{\nu}(x)|^2 = 1$$

for all $x \in \partial D$. Under these assumptions the sectoriality of L follows according to Corollary 3.5. Since the domain of L is dense in $L^2(D)$, the analytic semigroup generated by L is strongly continuous.

Theorem 6.5. *The operator L is sectorial if the diffusion coefficient κ is positive and bounded from below, $\kappa(x) \geq \delta > 0$ for all $x \in \bar{D}$, and the diffusion coefficient and velocity of the flow satisfy the conditions*

$$\begin{cases} \kappa : \bar{D} \to \mathbb{R}, & \kappa \in UC^1(\bar{D}), \\ \mathbf{v} : \bar{D} \to \mathbb{R}^d, & \mathbf{v} \in UC(\bar{D}). \end{cases}$$

Hence under these assumptions the operator L generates a strongly continuous analytic semigroup $\{\mathcal{U}(t)\}_{t \geq 0}$.

6.3.2 Stochastic Convection–Diffusion Equation

We assume that the diffusion coefficient and velocity of the flow fulfill the requirements of Theorem 6.5. Let $T > 0$ and $\{\mathcal{F}_t\}_{t \in [0,T]}$ be a normal filtration in $(\Omega, \mathcal{F}, \mathbb{P})$. Let Q be a positive self-adjoint trace class operator from $L^2(D)$ to itself with $\operatorname{Ker} Q = \{0\}$ and $W(t)$, $t \in [0,T]$, a Q-Wiener process in $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $L^2(D)$ with respect to the filtration $\{\mathcal{F}_t\}_{t \in [0,T]}$. According to Section 4.6 under some assumptions of the initial value C_0 and control term f the stochastic convection–diffusion equation has the weak solution.

Theorem 6.6. *If $f \in L^1(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; L^2(D))$ and C_0 is $\mathcal{F}_0$-measurable, the stochastic convection–diffusion equation has the weak solution $C(t)$, $t \in [0,T]$, which is the predictable process given by the formula*

$$C(t) = \mathcal{U}(t)C_0 + \int_0^t \mathcal{U}(t-s)f(s)\,ds + \int_0^t \mathcal{U}(t-s)\,dW(s)$$

for all $t \in [0,T]$ almost surely.

There are some parameters in our model which we can choose rather freely and still have the weak solution to the stochastic convection–diffusion equation. The diffusion coefficient and velocity of the flow should only satisfy the requirements of Theorem 6.5. The covariance operator Q of the Wiener process can be an arbitrary positive self-adjoint trace class operator from $L^2(D)$ to itself with $\operatorname{Ker} Q = \{0\}$ by Proposition 4.17. The natural choice of the filtration is the filtration defined by the Wiener process, i.e., $\mathcal{F}_t^W = \sigma(W(s),\ s \leq t)$ for all $t \in [0,T]$. Since the filtration should be normal, by Proposition 4.35 the augmented filtration $\{\mathcal{F}_t^{W,\mathbb{P}}\}_{t\in[0,T]}$ is an appropriate choice assuming that the probability space is complete. We want that the initial value C_0 is a Gaussian $L^2(D)$-valued $\mathcal{F}_0$-measurable function with mean c_0 and covariance Γ_0. The benefits of this requirement will appear later in this section. Then the mean c_0 can be an arbitrary $L^2(D)$-function and the covariance operator Γ_0 has same requirements as Q by Proposition 4.17. The control term f should be an $L^2(D)$-valued integrable predictable process.

6.3.3 Discrete Evolution Equation Without Control

We assume that there is no control in our system, i.e., $f \equiv 0$. Then the weak solution of the stochastic convection–diffusion equation is the predictable process given by the formula

$$C(t) = \mathcal{U}(t)C_0 + \int_0^t \mathcal{U}(t-s)\, dW(s) \tag{6.13}$$

for all $t \in [0,T]$ almost surely. Since the initial value C_0 is a Gaussian random variable with mean c_0 and covariance Γ_0, the concentration distribution C has a Gaussian modification by Lemma 4.42 and Proposition 4.46. Furthermore, the mean of the Gaussian modification is $\mathcal{U}(t)c_0$ and the covariance operator is

$$\mathcal{U}(t)\Gamma_0\mathcal{U}^*(t) + \int_0^t \mathcal{U}(t-s)Q\mathcal{U}^*(t-s)\, ds \tag{6.14}$$

for all $t \in [0,T]$. We assume that the measurements are done in time instants $0 < t_1 < \ldots < t_n \leq T$. We use the notation $t_0 := 0$ and $C_k := C(t_k)$ and $\Delta_k := t_{k+1} - t_k$ for all $k = 0,\ldots,n-1$. Then the discrete evolution equation for the concentration distribution is

$$C_{k+1} = \mathcal{U}(\Delta_k)C_k + W_{k+1}$$

for all $k = 0,\ldots,n-1$ almost surely where

$$W_{k+1} := \int_{t_k}^{t_{k+1}} \mathcal{U}(t_{k+1}-s)\, dW(s)$$

by Theorem 4.39. The term W_{k+1} can be seen as a state noise for all $k = 0,\ldots,n-1$. The state noise W_{k+1} is a Gaussian random variable with mean 0 and covariance operator

$$\operatorname{Cov}(W_{k+1}) = \int_{t_k}^{t_{k+1}} \mathcal{U}(t_{k+1}-s)Q\mathcal{U}^*(t_{k+1}-s)\, ds \tag{6.15}$$

and it is independent of $\mathcal{F}_{t_k}$ for all $k = 0,\ldots,n-1$ by Lemma 4.42 and Proposition 4.46. Thus C_k and W_{k+1} are independent for all $k = 0,\ldots,n-1$. Furthermore, the state noises at different time instants are uncorrelated since

$$\begin{aligned}&\operatorname{Cor}(W_{k+1}, W_{l+1})\\&= \int_0^{t_{k+1}\wedge t_{l+1}} \chi_{[t_k,t_{k+1}]}(s)\chi_{[t_l,t_{l+1}]}(s)\mathcal{U}(t_{k+1}-s)Q\mathcal{U}^*(t_{l+1}-s)\,ds = 0\end{aligned}$$

for all $k \neq l$ by Proposition 4.41.

The discrete state estimation system for the electrical impedance process tomography problem is

$$C_{k+1} = \mathcal{U}(\Delta_k)C_k + W_{k+1}, \quad k = 0, \ldots, n-1, \tag{6.16}$$

$$V^k = U(g(C_k); I^k) + S^k, \quad k = 1, \ldots, n. \tag{6.17}$$

Since the observation model function $U \circ g$ is non-linear, the filtering problem is much more demanding than in the linear case. In the numerical implementations of this problem in the articles and proceedings papers [43, 44, 45, 40, 41, 42, 38, 39] the observation model function is linearized. In Theorem 5.8 we have shown that the mapping $\sigma \mapsto U(\sigma; I)$ is Fréchet differentiable. If the function g is Fréchet differentiable, the observation equation may be linearized.

6.3.4 Space Discretization

The realizations of the concentration distribution C are in the space $L^2(D)$. The computation requires space discretization. We need to choose a finite dimensional subspace of $L^2(D)$ and assume that the realizations of the concentration distribution are in that subspace. This causes a discretization error. Usually the discretization error is ignored in numerical implementations. The discretized state estimation system is assumed to represent the reality. In this subsection we want to analyse the stochastic nature of the discretization error in the case of electrical impedance process tomography.

Let $\{\mathcal{V}_m\}_{m=1}^{\infty}$ be a sequence of finite dimensional subspaces of $L^2(D)$ such that $\mathcal{V}_m \subset \mathcal{V}_{m+1}$ for all $m \in \mathbb{N}$ and $\overline{\cup \mathcal{V}_m} = L^2(D)$. Since $L^2(D)$ is a separable Hilbert space, there exists such a sequence, for example, $\mathcal{V}_m$ may be the subspace spanned by the m first functions in an orthonormal basis of $L^2(D)$. Let $\{\varphi_l^m\}_{l=1}^{N_m}$ be an orthonormal basis of $\mathcal{V}_m$ for all $m \in \mathbb{N}$. We denote by $(\cdot,\cdot)$ the inner product in $L^2(D)$. We define the orthogonal projection $P_m : L^2(D) \to \mathcal{V}_m$ for $m \in \mathbb{N}$ by

$$P_m f = \sum_{l=1}^{N_m} (f, \varphi_l^m)\varphi_l^m$$

for all $f \in L^2(D)$. The subspaces $\mathcal{V}_m$ are appropriate discretization spaces if $P_m f \to f$ in $L^2(D)$ as $m \to \infty$ for all $f \in L^2(D)$, i.e., the orthogonal projections P_m converge strongly to the identity operator.

Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to L^2(D)$ be a random variable. Then for all $\omega \in \Omega$

$$P_m X(\omega) = \sum_{l=1}^{N_m} (X(\omega), \varphi_l^m)\varphi_l^m = \sum_{l=1}^{N_m} (X^m(\omega))_l \varphi_l^m$$

where $X^m(\omega) := ((X(\omega), \varphi_1^m), \ldots, (X(\omega), \varphi_{N_m}^m))^T$ is an $\mathbb{R}^{N_m}$-valued random variable. We view X^m as a discretized version of the random variable X at the discretization level m. If X is a Gaussian random variable with mean $\bar{x}$ and covariance Γ, then X^m is also Gaussian [21, Theorem A.5]. Furthermore, the mean of X^m is $\mathbb{E}X^m = ((\bar{x}, \varphi_1^m), \ldots, (\bar{x}, \varphi_{N_m}^m))^T$ and the covariance matrix is defined by $(\operatorname{Cov} X^m)_{ij} := (\Gamma\varphi_i^m, \varphi_j^m)$ since

$$\mathbb{E}(X^m)_i(X^m)_j = \mathbb{E}(X, \varphi_i^m)(X, \varphi_j^m) = (\Gamma\varphi_i^m, \varphi_j^m) + (\bar{x}, \varphi_i^m)(\bar{x}, \varphi_j^m)$$

for all $i, j = 1, \ldots, N_m$.

Evolution Equation

We want to discretize the discrete evolution equation (6.16). We use the discretization level m. We form an evolution equation for the discrete $\mathbb{R}^{N_m}$-valued stochastic process $\{C_k^m\}_{k=0}^n$ where $C_k^m := ((C_k, \varphi_1^m), \ldots, (C_k, \varphi_{N_m}^m))^T$ for all $k = 0, \ldots, n$. By using the discrete evolution equation (6.16),

$$\begin{aligned}
(C_{k+1}^m)_i &= (C_{k+1}, \varphi_i^m) = (\mathcal{U}(\Delta_k)C_k + W_{k+1}, \varphi_i^m) \\
&= (\mathcal{U}(\Delta_k)P_m C_k, \varphi_i^m) + (\mathcal{U}(\Delta_k)(I - P_m)C_k, \varphi_i^m) + (W_{k+1}, \varphi_i^m) \\
&= \sum_{l=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi_l^m, \varphi_i^m)(C_k^m)_l + (E_{k+1}^m)_i + (W_{k+1}^m)_i
\end{aligned}$$

for all $i = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$ almost surely where the discrete stochastic process

$$E_{k+1}^m = ((C_k, (I - P_m)\mathcal{U}^*(\Delta_k)\varphi_1^m), \ldots, (C_k, (I - P_m)\mathcal{U}^*(\Delta_k)\varphi_{N_m}^m))^T$$

represent the discretization error and W_{k+1}^m is the state noise vector. Thus the discretized evolution equation is

$$C_{k+1}^m = A_{k+1}^m C_k^m + E_{k+1}^m + W_{k+1}^m \tag{6.18}$$

for all $k = 0, \ldots, n-1$ almost surely where the matrix A_{k+1}^m is defined by

$$(A_{k+1}^m)_{ij} := (\mathcal{U}(\Delta_k)\varphi_j^m, \varphi_i^m) \tag{6.19}$$

for all $i, j = 1, \ldots, N_m$. The discretized evolution equation (6.18) is used in the evolution updating step of the Bayesian filtering. Thereby we need to define the statistical quantities of the discrete stochastic processes $\{E_{k+1}^m\}_{k=0}^{n-1}$ and $\{W_{k+1}^m\}_{k=0}^{n-1}$.

The state noise W_{k+1} is a Gaussian random variable with mean 0 and covariance given by Formula (6.15) for all $k = 0, \ldots, n-1$. Hence the state noise vector W_{k+1}^m is Gaussian with mean 0 and covariance matrix

$$\begin{aligned}
(\operatorname{Cov}(W_{k+1}^m))_{ij} &= \left(\int_{t_k}^{t_{k+1}} \mathcal{U}(t_{k+1} - s)Q\mathcal{U}^*(t_{k+1} - s)\, ds\, \varphi_i^m, \varphi_j^m \right) \\
&= \int_{t_k}^{t_{k+1}} \left(\mathcal{U}(t_{k+1} - s)Q\mathcal{U}^*(t_{k+1} - s)\varphi_i^m, \varphi_j^m \right) ds
\end{aligned}$$

for all $i, j = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$. We define the matrix $Q_{k,l}^m(s)$ by

$$(Q_{k,l}^m(s))_{ij} := (\mathcal{U}(t_k - s)Q\mathcal{U}^*(t_l - s)\varphi_i^m, \varphi_j^m)$$

for all $i, j = 1, \dots, N_m$, $k, l = 0, \dots, n-1$ and $s \in [0, t_k \wedge t_l]$. Then

$$\mathrm{Cov}(W_{k+1}^m) = \int_{t_k}^{t_{k+1}} Q_{k+1,k+1}^m(s)\, ds \tag{6.20}$$

for all $k = 0, \dots, n-1$. Since the state noises W_k and W_l are uncorrelated for all $k \neq l$, by the Gaussianity the state noise vectors W_k^m and W_l^m are independent if $k \neq l$.

We use our knowledge of the stochastic behaviour of the continuous evolution equation (6.13) for the examination of the discretization error E_{k+1}^m for all $k = 0, \dots, n-1$. The concentration distribution C_k has a Gaussian modification with mean $\mathcal{U}(t_k)c_0$ and covariance given by Formula (6.14) where $t = t_k$ for all $k = 0, \dots, n-1$. Hence the discretization error E_{k+1}^m has a Gaussian version for all $k = 0, \dots, n-1$. The mean of the Gaussian version is given by

$$\begin{aligned}(\mathbb{E}E_{k+1}^m)_i &= \mathbb{E}(C_k, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m) = (\mathbb{E}C_k, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m) \\ &= (\mathcal{U}(t_k)c_0, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m) = (c_0, \mathcal{U}^*(t_k)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m)\end{aligned}$$

for all $i = 1, \dots, N_m$ and $k = 0, \dots, n-1$. Thus

$$(\mathbb{E}E_{k+1}^m)_i = (c_0, \mathcal{U}^*(t_{k+1})\varphi_i^m) - \sum_{l=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi_l^m, \varphi_i^m)(c_0, \mathcal{U}^*(t_k)\varphi_l^m)$$

for all $i = 1, \dots, N_m$ because

$$\begin{aligned}\mathcal{U}^*(t)(I-P_m)\mathcal{U}^*(s)f &= \mathcal{U}^*(t+s)f - \sum_{l=1}^{N_m} (\mathcal{U}^*(s)f, \varphi_l^m)\mathcal{U}^*(t)\varphi_l^m \\ &= \mathcal{U}^*(t+s)f - \sum_{l=1}^{N_m} (\mathcal{U}(s)\varphi_l^m, f)\mathcal{U}^*(t)\varphi_l^m\end{aligned}$$

for all $f \in L^2(D)$ and $s, t \in [0, T]$. Hence

$$\mathbb{E}E_{k+1}^m = \begin{bmatrix} (\mathcal{U}(t_{k+1})c_0, \varphi_1^m) \\ \vdots \\ (\mathcal{U}(t_{k+1})c_0, \varphi_{N_m}^m) \end{bmatrix} + A_{k+1}^m \begin{bmatrix} (\mathcal{U}(t_k)c_0, \varphi_1^m) \\ \vdots \\ (\mathcal{U}(t_k)c_0, \varphi_{N_m}^m) \end{bmatrix} \tag{6.21}$$

for all $k = 0, \dots, n-1$. The covariance matrix of the Gaussian version is given by

$$\begin{aligned}&(\mathrm{Cov}\, E_{k+1}^m)_{ij} \\ &= (\mathrm{Cov}(C_k)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m) \\ &= (\mathcal{U}(t_k)\Gamma_0\mathcal{U}^*(t_k)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m) + \\ &\quad + \left(\int_0^{t_k} \mathcal{U}(t_k - s)Q\mathcal{U}^*(t_k - s)\, ds\, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m\right) \\ &= (\Gamma_0\mathcal{U}^*(t_k)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m, \mathcal{U}^*(t_k)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m) + \\ &\quad + \int_0^{t_k} \big(Q\mathcal{U}^*(t_k - s)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_i^m, \mathcal{U}^*(t_k - s)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m\big)\, ds\end{aligned}$$

for all $i, j = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$. Since Γ_0 and Q are self-adjoint as the covariance operator of Gaussian random variables,

$$
\begin{aligned}
&(\mathrm{Cov}\, E^m_{k+1})_{ij} \\
&= (\Gamma_0 \mathcal{U}^*(t_{k+1})\varphi^m_i, \mathcal{U}^*(t_{k+1})\varphi^m_j) + \\
&\quad - \sum_{l=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi^m_l, \varphi^m_j)(\Gamma_0 \mathcal{U}^*(t_k)\varphi^m_l, \mathcal{U}^*(t_{k+1})\varphi^m_i) + \\
&\quad - \sum_{l=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi^m_l, \varphi^m_i)(\Gamma_0 \mathcal{U}^*(t_k)\varphi^m_l, \mathcal{U}^*(t_{k+1})\varphi^m_j) + \\
&\quad + \sum_{l,p=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi^m_l, \varphi^m_i)(\mathcal{U}(\Delta_k)\varphi^m_p, \varphi^m_j)(\Gamma_0 \mathcal{U}^*(t_k))\varphi^m_l, \mathcal{U}^*(t_k)\varphi^m_p) + \\
&\quad + \int_0^{t_k} \left(Q\mathcal{U}^*(t_{k+1}-s)\varphi^m_i, \mathcal{U}^*(t_{k+1}-s)\varphi^m_j\right) ds + \\
&\quad - \sum_{l=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi^m_l, \varphi^m_j) \int_0^{t_k} (Q\mathcal{U}^*(t_k-s)\varphi^m_l, \mathcal{U}^*(t_{k+1}-s)\varphi^m_i)\, ds + \\
&\quad - \sum_{l=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi^m_l, \varphi^m_i) \int_0^{t_k} \left(Q\mathcal{U}^*(t_k-s)\varphi^m_l, \mathcal{U}^*(t_{k+1}-s)\varphi^m_j\right) ds + \\
&\quad + \sum_{l,p=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi^m_l, \varphi^m_i)(\mathcal{U}(\Delta_k)\varphi^m_p, \varphi^m_j) \int_0^{t_k} \left(Q\mathcal{U}^*(t_k-s)\varphi^m_l, \mathcal{U}^*(t_k-s)\varphi^m_p\right) ds
\end{aligned}
$$

for all $i, j = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$. We define the matrix $\Gamma^{m,k}_{0,l}$ by

$$
(\Gamma^{m,k}_{0,l})_{ij} := (\mathcal{U}(t_k)\Gamma_0\mathcal{U}^*(t_l)\varphi^m_i, \varphi^m_j)
$$

for all $i, j = 1, \ldots, N_m$ and $k, l = 0, \ldots, n-1$. Then

$$
\begin{aligned}
\mathrm{Cov}\, E^m_{k+1} = \Gamma^{m,k+1}_{0,k+1} &- (A^m_{k+1}\Gamma^{m,k+1}_{0,k})^T - A^m_{k+1}\Gamma^{m,k+1}_{0,k} + A^m_{k+1}\Gamma^{m,k}_{0,k}(A^m_{k+1})^T + \\
&+ \int_0^{t_k} Q^m_{k+1,k+1}(s)\, ds - \int_0^{t_k} (A^m_{k+1}Q^m_{k+1,k}(s))^T\, ds + \\
&- \int_0^{t_k} A^m_{k+1}Q^m_{k+1,k}(s)\, ds + \int_0^{t_k} A^m_{k+1}Q^m_{k,k}(s)(A^m_{k+1})^T\, ds
\end{aligned}
\tag{6.22}
$$

for all $k = 0, \ldots, n-1$ where the integration is done componentwise.

Since C_k and W_{k+1} are independent, also C^m_k and W^m_{k+1} as well as E^m_{k+1} and W^m_{k+1} are mutually independent for all $k = 0, \ldots, n-1$. On the other hand, C^m_k and E^m_{k+1} are correlated for all $k = 0, \ldots, n-1$. The correlation matrix of C^m_k and E^m_{k+1} can be calculated by using the continuous evolution equation (6.13) for all $k = 0, \ldots, n-1$.

Then

$$\begin{aligned}(\mathrm{Cor}(C_k^m, E_{k+1}^m))_{ij} &= (\mathrm{Cov}(C_k)\varphi_i^m, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m) \\ &= (\mathcal{U}(t_k)\Gamma_0\mathcal{U}^*(t_k)\varphi_i^m, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m)+ \\ &\quad + \left(\int_0^{t_k} \mathcal{U}(t_k-s)Q\mathcal{U}^*(t_k-s)\, ds\; \varphi_i^m, (I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m\right) \\ &= (\Gamma_0\mathcal{U}^*(t_k)\varphi_i^m, \mathcal{U}^*(t_k)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m)+ \\ &\quad + \int_0^{t_k} \left(Q\mathcal{U}^*(t_k-s)\varphi_i^m, \mathcal{U}^*(t_k-s)(I-P_m)\mathcal{U}^*(\Delta_k)\varphi_j^m\right) ds\end{aligned}$$

for all $i, j = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$. Thus

$$\begin{aligned}(\mathrm{Cor}(C_k^m, E_{k+1}^m))_{ij} &= (\Gamma_0\mathcal{U}^*(t_k)\varphi_i^m, \mathcal{U}^*(t_{k+1})\varphi_j^m)+ \\ &\quad - \sum_{l=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi_l^m, \varphi_j^m)(\Gamma_0\mathcal{U}^*(t_k)\varphi_l^m, \mathcal{U}^*(t_k)\varphi_i^m)+ \\ &\quad + \int_0^{t_k} \left(Q\mathcal{U}^*(t_k-s)\varphi_i^m, \mathcal{U}^*(t_{k+1}-s)\varphi_j^m\right) ds+ \\ &\quad - \sum_{l=1}^{N_m} (\mathcal{U}(\Delta_k)\varphi_l^m, \varphi_j^m) \int_0^{t_k} (Q\mathcal{U}^*(t_k-s)\varphi_l^m, \mathcal{U}^*(t_k-s)\varphi_i^m)\, ds\end{aligned}$$

for all $i, j = 1, \ldots, N_m$ and hence

$$\begin{aligned}&\mathrm{Cor}(C_k^m, E_{k+1}^m) \\ &= \Gamma_{0,k}^{m,k+1} - (A_{k+1}^m \Gamma_{0,k}^{m,k})^T + \int_0^{t_k} \left(Q_{k+1,k}^m(s) - (A_{k+1}^m Q_{k,k}^m(s))^T\right) ds\end{aligned} \tag{6.23}$$

for all $k = 0, \ldots, n-1$.

According to the discretized evolution equation (6.18) the random variable C_{k+1}^m has a Gaussian version. The mean of the Gaussian version is

$$\mathbb{E}C_{k+1}^m = \mathbb{E}\left(A_{k+1}^m C_k^m + E_{k+1}^m + W_{k+1}^m\right) = A_{k+1}^m \mathbb{E}C_k^m + \mathbb{E}E_{k+1}^m \tag{6.24}$$

and the covariance matrix is

$$\begin{aligned}\mathrm{Cov}\, C_{k+1}^m &= \mathrm{Cov}\left(A_{k+1}^m C_k^m + E_{k+1}^m + W_{k+1}^m\right) \\ &= \mathrm{Cov}(A_{k+1}^m C_k^m) + \mathrm{Cor}(A_{k+1}^m C_k^m, E_{k+1}^m)+ \\ &\quad + \mathrm{Cor}(E_{k+1}^m, A_{k+1}^m C_k^m) + \mathrm{Cov}\, E_{k+1}^m + \mathrm{Cov}\, W_{k+1}^m \\ &= A_{k+1}^m \mathrm{Cov}(C_k^m)(A_{k+1}^m)^T + A_{k+1}^m \mathrm{Cor}(C_k^m, E_{k+1}^m)+ \\ &\quad + \mathrm{Cor}(C_k^m, E_{k+1}^m)^T (A_{k+1}^m)^T + \mathrm{Cov}\, E_{k+1}^m + \mathrm{Cov}\, W_{k+1}^m\end{aligned} \tag{6.25}$$

for all $k = 0, \ldots, n-1$.

Observation Equation

Since both the operator U and mapping g are non-linear, the space discretization of the observation equation (6.17) is far more difficult than the evolution equation (6.16), especially when we are interested in the discretization error. We assume that the function g is Fréchet differentiable. Then we can linearize the observation equation (6.17). For all $k = 1, \dots, n$

$$\begin{aligned} V^k &= U(g(C_k); I^k) + S^k \\ &\approx U(g(f); I^k) + U'(g(f); I^k)g'(f)(C_k - f) + S^k \\ &= U(g(f); I^k) - U'(g(f); I^k)g'(f)f + U'(g(f); I^k)g'(f)C_k + S^k \end{aligned}$$

where $f \in L^2(D)$. The point f in which the linearization is done should be chosen wisely. It may be, for example, the mean of the initial value. The linearization induces error. However, in future we ignore the linearization error. We denote $b_k := U(g(f); I^k) - U'(g(f); I^k)g'(f)f$ and $B_k := U'(g(f); I^k)g'(f)$ for all $k = 1, \dots, n$. Then $b_k \in \mathbb{R}^L$ and $B_k : L^2(D) \to \mathbb{R}^L$ is a bounded linear operator for all $k = 1, \dots, n$. The linearized observation equation is

$$V^k = B_k C_k + b_k + S^k \tag{6.26}$$

for all $k = 1, \dots, n$. Then the discretized observation equation is

$$V^k = B_k P_m C_k + B_k(I - P_m)C_k + b_k + S^k = [B_k\varphi]C_k^m + \mathcal{E}_k^m + b_k + S^k \tag{6.27}$$

for all $k = 1, \dots, n$ where $[B_k\varphi] := [B_k\varphi_1^m \dots B_k\varphi_{N_m}^m]$ is the $L \times N_m$ matrix whose l^{th} column is $B_k\varphi_l^m$ for all $l = 1, \dots, N_m$ and $\mathcal{E}_k^m := B_k(I - P_m)C_k$ represents the discretization error. The discretized observation equation (6.27) is used in the observation updating step of the Bayesian filtering. We need to define the statistical quantities of the processes $\{\mathcal{E}_k^m\}_{k=1}^n$ and $\{V^k\}_{k=1}^n$.

We use our knowledge of the stochastic behaviour of the continuous evolution equation (6.13). We assume that the process $S(t)$, $t \geq 0$, is a Gaussian process independent of the process $C(t)$, $t \geq 0$. Then S^k is independent of C_k^m and $\mathcal{E}_k^m$ for all $k = 1, \dots, n$. On the other hand, C_k^m and $\mathcal{E}_k^m$ are correlated for all $k = 1, \dots, n$. The concentration distribution C_k has a Gaussian modification with mean $\mathcal{U}(t_k)c_0$ and covariance given by Formula (6.14) where $t = t_k$ for all $k = 1, \dots, n$. Hence the discretization error $\mathcal{E}_k^m$ has a Gaussian version for all $k = 1, \dots, n$. The mean of the Gaussian version is

$$\mathbb{E}\mathcal{E}_k^m = B_k(I - P_m)\mathcal{U}(t_k)c_0$$

and the covariance matrix is

$$\begin{aligned} \operatorname{Cov}\mathcal{E}_k^m &= B_k(I - P_m)\mathcal{U}(t_k)\Gamma_0\mathcal{U}^*(t_k)(I - P_m)B_k^* + \\ &\quad + \int_0^{t_k} B_k(I - P_m)\mathcal{U}(t_k - s)Q\mathcal{U}^*(t_k - s)(I - P_m)B_k^*\, ds \end{aligned}$$

for all $k = 1, \dots, n$. The correlation matrix of C_k^m and $\mathcal{E}_k^m$ can be calculated by using the continuous evolution equation (6.13) for all $k = 1, \dots, n$. First of all,

$$\begin{aligned} \operatorname{Cor}([B_k\varphi]C_k^m, \mathcal{E}_k^m) &= \operatorname{Cor}(B_k P_m C_k, B_k(I - P_m)C_k) \\ &= B_k P_m \operatorname{Cov}(C_k)(I - P_m)B_k^* \\ &= B_k P_m \mathcal{U}(t_k)\Gamma_0\mathcal{U}^*(t_k)(I - P_m)B_k^* + \\ &\quad + \int_0^{t_k} B_k P_m \mathcal{U}(t_k - s)Q\mathcal{U}^*(t_k - s)(I - P_m)B_k^*\, ds \end{aligned}$$

for all $k = 1, \ldots, n$. If the matrix $[B_k\varphi]$ is invertable,

$$\operatorname{Cor}(C_k^m, \mathcal{E}_k^m) = [B_k\varphi]^{-1} \operatorname{Cor}([B_k\varphi]C_k^m, \mathcal{E}_k^m)$$

for all $k = 1, \ldots, n$.

In the observation updating step of the Bayesian filtering we need the joint probability distribution of C_k^m and V^k for all $k = 1, \ldots, n$. By the continuous evolution equation (6.13) the random variable C_k^m has a Gaussian version. According to the discretized observation equation (6.27) the random variable V^k has a Gaussian version and the joint probability distribution of C_k^m and V^k is Gaussian. In addition, the mean of the Gaussian version of V^k is

$$\mathbb{E}V^k = \mathbb{E}\left([B_k\varphi]C_k^m + \mathcal{E}_k^m + b_k + S^k\right) = [B_k\varphi]\mathbb{E}C_k^m + \mathbb{E}\mathcal{E}_k^m + b_k + \mathbb{E}S^k \tag{6.28}$$

and the covariance matrix is

$$\begin{aligned}
\operatorname{Cov} V^k &= \operatorname{Cov}\left([B_k\varphi]C_k^m + \mathcal{E}_k^m + b_k + S^k\right)\\
&= \operatorname{Cov}([B_k\varphi]C_k^m) + \operatorname{Cor}([B_k\varphi]C_k^m, \mathcal{E}_k^m) + \\
&\quad + \operatorname{Cor}(\mathcal{E}_k^m, [B_k\varphi]C_k^m) + \operatorname{Cov}(\mathcal{E}_k^m) + \operatorname{Cov}(S^k)\\
&= [B_k\varphi] \operatorname{Cov}(C_k^m)[B_k\varphi]^T + \operatorname{Cor}([B_k\varphi]C_k^m, \mathcal{E}_k^m) + \\
&\quad + \operatorname{Cor}([B_k\varphi]C_k^m, \mathcal{E}_k^m)^T + \operatorname{Cov}(\mathcal{E}_k^m) + \operatorname{Cov}(S^k)
\end{aligned} \tag{6.29}$$

for all $k = 1, \ldots, n$. The correlation matrix of C_k^m and V^k is

$$\begin{aligned}
\operatorname{Cor}(C_k^m, V^k) &= \operatorname{Cor}\left(C_k^m, [B_k\varphi]C_k^m + \mathcal{E}_k^m + b_k + S^k\right)\\
&= \operatorname{Cor}(C_k^m, [B_k\varphi]C_k^m) + \operatorname{Cor}(C_k^m, \mathcal{E}_k^m)\\
&= \operatorname{Cov}(C_k^m)[B_k\varphi]^T + [B_k\varphi]^{-1} \operatorname{Cor}([B_k\varphi]C_k^m, \mathcal{E}_k^m)
\end{aligned} \tag{6.30}$$

for all $k = 1, \ldots, n$.

Bayesian Filtering

The discretized state estimation system concerning the electrical impedance process tomography problem is

$$C_{k+1}^m = A_{k+1}^m C_k^m + E_{k+1}^m + W_{k+1}^m, \quad k = 0, \ldots, n-1, \tag{6.31}$$

$$V^k = [B_k\varphi]C_k^m + \mathcal{E}_k^m + b_k + S^k, \quad k = 1, \ldots, n. \tag{6.32}$$

The state noise vectors W_k^m and W_l^m are mutually independent and also independent of C_0^m for all $k \neq l$. We assume that the observation noise vectors S^k are chosen such a way that S^k and S^l are mutually independent and also independent of C_0^m for all $k \neq l$ and S^k and W_l^m are mutually independent for all $k, l = 1, \ldots, n$. Then the stochastic processes $\{C_k^m\}_{k=0}^n$ and $\{V^k\}_{k=1}^n$ form an evolution–observation model. Therefore we may use the Bayesian filtering method.

In the evolution updating step of the Bayesian filtering it is assumed that we know the conditional probability density of C_k^m with respect to some measurements $D_k := \{v^1, v^2, \ldots, v^k\}$. We need to calculate the conditional probability density of C_{k+1}^m

with respect to the data D_k. We suppose that the conditional expectation $\mathbb{E}(C_k^m|D_k)$ is a Gaussian random variable with mean $\bar{c}_k$ and covariance matrix Γ_k. According to the discretized evolution equation (6.31) we are able to present the joint distribution of C_k^m and C_{k+1}^m conditioned on the measurements D_k and know that it is Gaussian. By Theorems 6.3 and 6.4 and Formulas (6.24) and (6.25) the Gaussianity of the joint probability density implies that the conditional marginal probability density of C_{k+1}^m is Gaussian with mean

$$\bar{c}_{k+1} := A_{k+1}^m \bar{c}_k + \mathbb{E}E_{k+1}^m \tag{6.33}$$

and covariance matrix

$$\begin{aligned}\Gamma_{k+1} := A_{k+1}^m \Gamma_k (A_{k+1}^m)^T + A_{k+1}^m \operatorname{Cor}(C_k^m, E_{k+1}^m) + \\ + \operatorname{Cor}(C_k^m, E_{k+1}^m)^T (A_{k+1}^m)^T + \operatorname{Cov} E_{k+1}^m + \operatorname{Cov} W_{k+1}^m.\end{aligned} \tag{6.34}$$

Thus the evolution updating step is defined if we are able to calculate the vector $\mathbb{E}E_{k+1}^m$ and matrices A_{k+1}^m, $\operatorname{Cov} E_{k+1}^m$, $\operatorname{Cov} W_{k+1}^m$ and $\operatorname{Cor}(C_k^m, E_{k+1}^m)$ given by Formulas (6.19)–(6.23) for all $k = 0, \dots, n-1$.

In the observation updating step of the Bayesian filtering it is assumed that we know the conditional probability density of C_{k+1}^m with respect to some measured data $D_k := \{v^1, v^2, \dots, v^k\}$. A new measurement v^{k+1} is obtained. We need to calculate the conditional probability density of C_{k+1}^m with respect to measurements $D_{k+1} := \{v^1, v^2, \dots, v^{k+1}\}$. We suppose that the conditional expectation $\mathbb{E}(C_{k+1}^m|D_k)$ is a Gaussian random variable with mean $\bar{c}_{k+1}$ and covariance matrix Γ_{k+1}. By Theorems 6.3 and 6.4 the conditional probability density of C_{k+1}^m with respect to the data D_{k+1} is Gaussian with mean

$$\tilde{c}_{k+1} := \bar{c}_{k+1} + \operatorname{Cor}(C_{k+1}^m, V^{k+1}) \operatorname{Cov}(V^{k+1})^{-1} (v^{k+1} - \mathbb{E}V^{k+1}) \tag{6.35}$$

and covariance matrix

$$\widetilde{\Gamma}_{k+1} = \Gamma_{k+1} - \operatorname{Cor}(C_{k+1}^m, V^{k+1}) \operatorname{Cov}(V^{k+1})^{-1} \operatorname{Cor}(C_{k+1}^m, V^{k+1})^T. \tag{6.36}$$

Thus the observation updating step is defined if we are able to calculate the vector $\mathbb{E}V^{k+1}$ and matrices $\operatorname{Cov}(V^{k+1})$ and $\operatorname{Cor}(C_{k+1}^m, V^{k+1})$ given by Formulas (6.28)–(6.30) for all $k = 0, \dots, n-1$.

The evaluation of the matrices needed in the Bayesian filtering method depends on the discretization space $\mathcal{V}_m$, analytic semigroup $\{\mathcal{U}(t)\}_{t \geq 0}$, function c_0, operators Γ_0, Q and B_k, vector b_k and statistics of the observation noise S for all $k = 1, \dots, n$. Usually the discretization space $\mathcal{V}_m$ is chosen such a way that the projection P_m is fairly easy to calculate. For example, the basis functions φ_l^m have compact supports and they are piecewise polynomial. The function c_0 and operator Γ_0 represent our prior knowledge of the concentration distribution. The mean c_0 should illustrate the expected concentration distribution in the pipe and hence it depends heavily on the application. Since the diffusion is a smoothing operation, we may assume that the initial state is rather smooth. Thus the covariance operator Γ_0 should have some smoothness properties. Our certainty of the time evolution model is coded into the Wiener process and hence into the operator Q. The choice of Q depends on the application. The crucial factor in the evaluation of the matrices is the analytic semigroup $\{\mathcal{U}(t)\}_{t \geq 0}$. Since it is defined by Formula (2.3), only in some special cases we can present the analytic semigroup in a closed form. In Subsection 6.3.5 we study the one dimensional version of the problem. Then the analytic semigroup is a convolution operator. The operator B_k and vector b_k for all $k = 1, \dots, n$ are related to the measurement situation. We need to be able to solve the complete electrode model for a known concentration distribution and also to calculate the Fréchet derivatives of mappings U and g for that concentration distribution. The function g depends on the application. At least for strong electrolytes and multiphase mixtures relations between the conductivity and concentration distribution are studied and discussed in the literature. The observation noise S represents the accuracy of the measurement equipment.

6.3.5 One Dimensional Model Case

As an example we examine the one dimensional case. Then the pipeline is modeled by the real line. In one dimension the electrical impedance tomography is not defined, especially not in unbounded domains. We have to use some other measurement process. Since we are interested in electrical impedance process tomography, the observation equation in this model case is not specified. One possibility is to observe point values of the concentration distribution through a blurring kernel and additive noise. Then the observation equation is linear. From the point of view of the evolution equation the one dimensional example is reasonable. We present some aspects of the approach introduced in this section. Numerical implementations are not included in this thesis.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. We examine the stochastic initial value problem

$$\begin{cases} dC(t) = LC(t)dt + dW(t), & t > 0, \\ C(0) = C_0 \end{cases} \tag{6.37}$$

where the operator L is defined by

$$\begin{aligned} L : H^2(\mathbb{R}) &\to L^2(\mathbb{R}) \\ f &\mapsto \frac{d}{dx}\left(\kappa(x)\frac{d}{dx}f\right) - v(x)\frac{d}{dx}f. \end{aligned}$$

For simplicity we assume that the diffusion coefficient and velocity of the flow do not depend on the space variable, i.e., $\kappa(x) = \kappa > 0$ and $v(x) = v > 0$ for all $x \in \mathbb{R}$. Let W be an $L^2(\mathbb{R})$-valued Q-Wiener process where Q is a positive self-adjoint trace class operator from $L^2(\mathbb{R})$ to itself with $\operatorname{Ker} Q = \{0\}$. As a normal filtration we have the augmentation $\{\mathcal{F}_t^{W,\mathbb{P}}\}_{t\geq 0}$ of the filtration generated by the Wiener process.

Analytic Semigroup

By Theorem 6.5 the convection–diffusion operator

$$\begin{aligned} L : \mathcal{D}(L) \subset L^2(\mathbb{R}) &\to L^2(\mathbb{R}) \\ f &\mapsto \left(\kappa\frac{d^2}{dx^2} - v\frac{d}{dx}\right) f \end{aligned}$$

where $\mathcal{D}(L) = H^2(\mathbb{R})$ and $\kappa, v > 0$ generates an analytic semigroup. Furthermore, the semigroup is strongly continuous. The semigroup is defined by Formula (2.3). We do not want to examine the spectral properties of the convection–diffusion operator. We try to find an easier way to calculate the analytic semigroup. According to Theorem 2.8 the solution to the initial value problem

$$\begin{cases} \frac{\partial}{\partial t}c(t,x) = \kappa \frac{\partial^2}{\partial x^2}c(t,x) - v\frac{\partial}{\partial x}c(t,x), \quad t > 0, \\ c(0,x) = c_0(x) \end{cases} \tag{6.38}$$

where $c_0 \in L^2(\mathbb{R})$ is given by the analytic semigroup generated by the convection–diffusion operator L. By solving the initial value problem (6.38) using other techniques we are able to find the analytic semigroup generated by the convection–diffusion operator. We may use a Ito diffusion to solve the initial value problem (6.38) when $c_0 \in C_0^2(\mathbb{R})$ and then try to generalize the form of the solution to the initial values $c_0 \in L^2(\mathbb{R})$.

Definition 6.7. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A Ito diffusion is a stochastic process $X(t)(\omega) = X(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}^n$ satisfying a stochastic differential equation*

$$\begin{cases} dX(t) = b(X(t))dt + \sigma(X(t))dB(t), \quad t > 0, \\ X(0) = x \end{cases} \tag{6.39}$$

where $x \in \mathbb{R}$, $B(t)$ is m-dimensional Brownian motion and $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are measurable functions satisfying

$$\|b(x)\|_{\mathbb{R}^n} + \|\sigma(x)\|_{\mathbb{R}^{n\times m}} \leq C(1 + \|x\|_{\mathbb{R}^n})$$

for all $x \in \mathbb{R}^n$ with some constant $C > 0$ and

$$\|b(x) - b(y)\|_{\mathbb{R}^n} + \|\sigma(x) - \sigma(y)\|_{\mathbb{R}^{n\times m}} \leq D\|x - y\|_{\mathbb{R}^n}$$

for all $x, y \in \mathbb{R}^n$ with some constant $D > 0$.

We denote the (unique) solution of the stochastic differential equation (6.39) by $\{X^x(t)\}_{t\geq 0}$. The existence and uniqueness of a solution is proved in [21, Theorem 5.2.1].

Definition 6.8. *Let $\{X(t)\}_{t\geq 0}$ be an Ito diffusion in $\mathbb{R}^n$. The (infinitesimal) generator A of X is defined by*

$$Af(x) = \lim_{t\to 0^+} \frac{\mathbb{E}^x[f(X(t))] - f(x)}{t}$$

for $x \in \mathbb{R}^n$ where $\mathbb{E}^x$ is the expectation with respect to the law of the Ito diffusion X assuming that $X(0) = x$, i.e.,

$$\mathbb{E}^x[f(X(t))] = \mathbb{E}[f(X^x(t))] = \int_\Omega f(X^x(t))\, d\mathbb{P} = \int_{\mathbb{R}^n} f(y)\, \mathcal{L}(X^x(t))(dy).$$

The set of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that the limit exists at x is denoted by $\mathcal{D}_x(A)$ while $\mathcal{D}(A)$ denotes the set of functions for which the limit exists for all $x \in \mathbb{R}^n$.

The infinitesimal generator of an Ito diffusion has a presentation as a differential operator.

Theorem 6.9. [21, Theorem 7.3.3] *Let X be the Ito diffusion*

$$dX(t) = b(X(t))dt + \sigma(X(t))dB(t), \quad t > 0.$$

If $f \in C_0^2(\mathbb{R}^n)$, then $f \in \mathcal{D}(A)$ and

$$Af(x) = \sum_{i=1}^{n} b_i(x)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n}(\sigma\sigma^T)_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}$$

for all $x \in \mathbb{R}^n$.

Theorem 6.9 indicates that Ito diffusions may be used for solving initial value problems.

Theorem 6.10. [21, Theorem 8.1.1] *Let X be an Ito diffusion in $\mathbb{R}^n$ with generator A. Let $f \in C_0^2(\mathbb{R}^n)$.*

(i) We define

$$u(t,x) = \mathbb{E}^x[f(X(t))] \tag{6.40}$$

for all $t > 0$ and $x \in \mathbb{R}^n$. Then $u(t,\cdot) \in \mathcal{D}(A)$ for each $t > 0$ and

$$\frac{\partial}{\partial t}u(t,x) = Au(t,x), \quad t > 0,\ x \in \mathbb{R}^n, \tag{6.41}$$

$$u(0,x) = f(x), \qquad x \in \mathbb{R}^n \tag{6.42}$$

where the right hand side of (6.41) is to be interpreted as A applied to the function $x \mapsto u(t,x)$ for each $t > 0$.

(ii) Moreover, if $w(t,x) \in C^{1,2}(\mathbb{R}\times\mathbb{R}^n)$ is a bounded function satisfying (6.41) and (6.42), then $w(t,x) = u(t,x)$ given by (6.40) for all $t > 0$ and $x \in \mathbb{R}^n$.

By Theorem 6.9 the generator of the Ito diffusion

$$\begin{cases} dX(t) = -vdt + \sqrt{2\kappa}dB(t), \\ X(0) = x \end{cases}$$

is the convection–diffusion operator

$$A = \kappa\frac{d^2}{dx^2} - v\frac{d}{dx}$$

and $C_0^2(\mathbb{R}) \subset \mathcal{D}(A)$. Thus according to Theorem 6.10 the solution to the initial value problem (6.38) where $c_0 \in C_0^2(\mathbb{R})$ is

$$c(t,x) = \mathbb{E}^x[c_0(X(t))]$$

for all $t > 0$ and $x \in \mathbb{R}$. But

$$X^x(t) = x - vt + \sqrt{2\kappa}B(t)$$

for all $t > 0$. Thus for all $t > 0$

$$X^x(t) \sim \mathcal{N}(x - vt, 2\kappa t)$$

and the density function of $X^x(t)$ is

$$\pi(y) = \frac{1}{2\sqrt{\pi\kappa t}} \exp\left(-\frac{(x-y-vt)^2}{4\kappa t}\right)$$

for all $y \in \mathbb{R}$. Hence

$$\begin{aligned} c(t,x) &= \mathbb{E}[c_0(X^x(t))] = \mathbb{E}[c_0(x - vt + \sqrt{2\kappa}B(t))] \\ &= \frac{1}{2\sqrt{\pi\kappa t}} \int_{-\infty}^{\infty} c_0(y) e^{-\frac{(x-y-vt)^2}{4\kappa t}}\, dy \end{aligned}$$

for all $t > 0$ and $x \in \mathbb{R}$. Let us denote

$$\Phi(t,x) = \frac{1}{2\sqrt{\pi\kappa t}} \exp\left(-\frac{(x-vt)^2}{4\kappa t}\right)$$

for all $t > 0$ and $x \in \mathbb{R}$. Then

$$c(t,x) = (\Phi(t,\cdot) * c_0)(x) \tag{6.43}$$

for all $t > 0$ and $x \in \mathbb{R}$ where

$$(\Phi(t,\cdot) * f)(x) := \int_{-\infty}^{\infty} \Phi(t, x-y) f(y)\, dy$$

for all $f \in L^2(\mathbb{R})$. Thus the solution to the initial value problem (6.38) is the convolution of the intial value c_0 with the probability density Φ if $c_0 \in C_0^2(\mathbb{R})$. We want to generalize this result to L^2-initial values.

We define an operator family $\{\mathcal{U}(t)\}_{t\geq 0}$ by

$$\begin{cases} \mathcal{U}(0)f = f, \\ (\mathcal{U}(t)f)(x) = (\Phi(t,\cdot) * f)(x), \quad t > 0, \end{cases}$$

for all $f \in L^2(\mathbb{R})$. Then $\mathcal{U}(t)$ is clearly linear for all $t \geq 0$. Furthermore, $\mathcal{U}(t)$ is bounded for all $t \geq 0$ since

$$\begin{aligned} |\Phi(t,\cdot) * f(x)| &\leq \int_{-\infty}^{\infty} \Phi(t, x-y) |f(y)|\, dy \\ &\leq \left(\int_{-\infty}^{\infty} \Phi(t,x-y)\, dy\right)^{\frac{1}{2}} \left(\int_{-\infty}^{\infty} \Phi(t,x-y)|f(y)|^2\, dy\right)^{\frac{1}{2}} \\ &= \left(\int_{-\infty}^{\infty} \Phi(t,x-y)|f(y)|^2\, dy\right)^{\frac{1}{2}} \end{aligned}$$

and hence

$$\begin{aligned} \|\mathcal{U}(t)f\|_{L^2(\mathbb{R})}^2 &= \int_{-\infty}^{\infty} |\Phi(t,\cdot) * f(x)|^2\, dx \leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi(t,x-y)|f(y)|^2\, dy\, dx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Phi(t,x-y)\, dx\, |f(y)|^2\, dy = \|f\|_{L^2(\mathbb{R})}^2 \end{aligned}$$

for all $f \in L^2(\mathbb{R})$. Thus $\mathcal{U}(t)$ is a bounded linear operator from $L^2(\mathbb{R}))$ to itself for all $t \geq 0$. Additionally,

$$\begin{aligned}
\mathcal{U}(t)\mathcal{U}(s)f(x) &= \frac{1}{4\pi\kappa\sqrt{st}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{((x-y)-vt)^2}{4\kappa t}} e^{-\frac{((y-z)-vs)^2}{4\kappa s}} f(z)\, dzdy \\
&= \frac{1}{4\pi\kappa\sqrt{st}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{((t+s)y-sx-tz)^2}{4\kappa st(t+s)}}\, dy\, e^{-\frac{((x-z)-v(t+s))^2}{4\kappa(t+s)}} f(z)\, dz \\
&= \frac{1}{2\sqrt{\pi\kappa(t+s)}} \int_{-\infty}^{\infty} e^{-\frac{((x-z)-v(t+s))^2}{4\kappa(t+s)}} f(z)\, dz \\
&= \mathcal{U}(t+s)f(x)
\end{aligned}$$

for all $f \in L^2(\mathbb{R})$, $s, t > 0$ and $x \in \mathbb{R}$. Therefore $\{\mathcal{U}(t)\}_{t\geq 0}$ is a semigroup since $\mathcal{U}(t)\mathcal{U}(0) = \mathcal{U}(t) = \mathcal{U}(0)\mathcal{U}(t)$ for all $t > 0$.

Let $c_0 \in L^2(\mathbb{R})$. The solution to the initial value problem (6.38) is $c(t,x) = \mathcal{U}(t)c_0(x)$ for all $t \geq 0$ and $x \in \mathbb{R}$ because $c(0,x) = \mathcal{U}(0)c_0(x) = c_0(x)$ and

$$\left(\frac{\partial}{\partial t} - \kappa\frac{\partial^2}{\partial x^2} + v\frac{\partial}{\partial x}\right) c(t,x) = \left(\left(\frac{\partial}{\partial t} - \kappa\frac{\partial^2}{\partial x^2} + v\frac{\partial}{\partial x}\right)\Phi(t,\cdot)\right) * c_0(x) = 0.$$

Hence according to Theorem 2.8 the semigroup $\{\mathcal{U}(t)\}_{t\geq 0}$ is the strongly continuous analytic semigroup generated by the convection–diffusion operator.

Wiener Process and the Initial Value

Our prior knowledge of the application we are interested in is coded into the choice of the initial value and covariance operator of the Wiener process. The initial value C_0 is a Gaussian random variable measurable with respect to the σ-algebra $\mathcal{F}_0^{W,\mathbb{P}}$. Hence we need to choose the mean c_0 and covariance operator Γ_0. In this model case our prior assumption is that the concentration distribution is almost uniform because in some real life applications that may be expected. Hence the mean could be a constant function. Since it should belong to $L^2(\mathbb{R})$, we have to do a cutting. In electrical impedance process tomography only finite number of electrodes are set on the surface of the pipe. Therefore we get information only from a part of the pipe. Our knowledge of the concentration distribution outside the so called measurement region is slight. Hence we may assume that the mean is a constant in the measurement region $|x| \leq M$ for some $M > 0$ and decays exponentially outside of it, for instance

$$c_0(x) = \begin{cases} c_0 & \text{if } |x| \leq M, \\ c_0 e^{-(|x|-M)} & \text{if } |x| > M, \end{cases} \tag{6.44}$$

for all $x \in \mathbb{R}$ where c_0 is a positive constant.

We need to choose an appropriate covariance operator for the initial value C_0. If the stochastic initial value problem (6.37) has the strong solution, by Definition 4.44 for almost all $(t,\omega) \in \Omega_T$ the solution $C(t,\omega)$ belongs to the domain of the convection–diffusion operator, i.e., $C(t,\omega) \in H^2(\mathbb{R})$ for almost all $(t,\omega) \in \Omega_T$. Thus we may expect that the initial value has some sort of smoothness properties. We assume that

$$\left(1 - \frac{d^2}{dx^2}\right) C_0 = \eta$$

where η is the *Gaussian white noise* in $L^2(\mathbb{R})$. Then $\mathbb{E}[(f,\eta)(g,\eta)] = (f,g)$ for all $f,g \in L^2(\mathbb{R})$. Thus for all $f,g \in C_0^\infty(\mathbb{R})$

$$
\begin{aligned}
(f,g) &= \mathbb{E}\left(\left(\left(1-\frac{d^2}{dx^2}\right)f, C_0\right)\left(\left(1-\frac{d^2}{dx^2}\right)g, C_0\right)\right) \\
&= \left(\Gamma_0\left(1-\frac{d^2}{dx^2}\right)f, \left(1-\frac{d^2}{dx^2}\right)g\right).
\end{aligned}
$$

We assume that Γ_0 is a convolution operator, i.e., $\Gamma_0 f = \gamma_0 * f$ for some $\gamma_0 \in L^2(\mathbb{R})$. Then by the Parseval formula,

$$
\begin{aligned}
(f,g) &= \left(\mathcal{F}\left((\gamma_0 * \left(1-\frac{d^2}{dx^2}\right)f\right), \mathcal{F}\left(\left(1-\frac{d^2}{dx^2}\right)g\right)\right) \\
&= (\hat{\gamma}_0(1+\xi^2)\hat{f}, (1+\xi^2)\hat{g}) = (\hat{\gamma}_0(1+\xi^2)^2\hat{f}, \hat{g})
\end{aligned}
$$

for all $f,g \in C_0^\infty(\mathbb{R})$. Hence we have $\hat{\gamma}_0(\xi) = (1+\xi^2)^{-2}$ for all $\xi \in \mathbb{R}$. Thus by the calculus of residues,

$$
\gamma_0(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{e^{ix\xi}}{(1+\xi^2)^2}\, d\xi = \sqrt{\frac{\pi}{8}}(1+|x|)e^{-|x|}
$$

for all $x \in \mathbb{R}$. Unfortunately, Γ_0 defined as an integral operator having the integral kernel $\gamma_0(x-y)$ is not a trace class operator. We have to do some sort of modification. We define an integral operator $\widetilde{\Gamma}_0$ with the integral kernel $\tilde{\gamma}_0(x,y) = w(x)\gamma_0(x-y)w(y)$ where the function w is exponentially decaying at infinity, $w(x) = 1$ when $|x| < N$ with some $N > 0$ and $0 < w(x) \leq 1$ for all $x \in \mathbb{R}$. Then $\widetilde{\Gamma}_0$ is self-adjoint since $\gamma_0(x-y) = \gamma_0(y-x)$ for all $x,y \in \mathbb{R}$. By the Parseval formula,

$$
(\widetilde{\Gamma}_0 f, f) = (\gamma_0 * (wf), wf) = (\hat{\gamma}_0\widehat{wf}, \widehat{wf}) = \int_{-\infty}^{\infty} \hat{\gamma}_0(\xi)|\widehat{wf}(\xi)|^2\, d\xi
$$

for all $f \in L^2(\mathbb{R})$. Since $\hat{\gamma}_0(\xi) > 0$ for all $\xi \in \mathbb{R}$, the operator $\widetilde{\Gamma}_0$ is positive. If $\widetilde{\Gamma}_0 f = 0$, then $(\widetilde{\Gamma}_0 f, f) = 0$. Thus $\widehat{wf} = 0$ almost everywhere. Hence $f = 0$ almost everywhere because $w > 0$. Therefore the kernel of $\widetilde{\Gamma}_0$ is trivial. The operator $\widetilde{\Gamma}_0$ is a composition of three operators, $\widetilde{\Gamma}_0 = M_w m_{\hat{\gamma}_0} M_w$ where

$$
M_w : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad f \mapsto wf
$$

is a multiplier and

$$
m_{\hat{\gamma}_0} : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad f \mapsto \mathcal{F}^{-1}(\hat{\gamma}_0\hat{f})
$$

is a Fourier multiplier. Furthermore, $m_{\hat{\gamma}_0} = m^2_{\hat{\gamma}_0^{1/2}}$. So

$$
\widetilde{\Gamma}_0 = M_w m^2_{\hat{\gamma}_0^{1/2}} M_w = \left(M_w m_{\hat{\gamma}_0^{1/2}}\right)\left(m_{\hat{\gamma}_0^{1/2}} M_w\right) = B^*B
$$

where

$$
Bf := m_{\hat{\gamma}_0^{1/2}} M_w f = \mathcal{F}^{-1}\left(\hat{\gamma}_0^{1/2}\widehat{wf}\right) = \mathcal{F}^{-1}\left(\hat{\gamma}_0^{1/2}\right) * (wf)
$$

for all $f \in L^2(\mathbb{R})$. Thus B is an integral operator with the integral kernel

$$
b(x,y) = \mathcal{F}^{-1}\left(\hat{\gamma}_0^{1/2}\right)(x-y)w(y) = \frac{w(y)}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{e^{i(x-y)\xi}}{1+\xi^2}\, d\xi = \sqrt{\frac{\pi}{2}}e^{-|x-y|}w(y)
$$

for all $x, y \in \mathbb{R}$. Since b is square integrable in $\mathbb{R}^2$, by Example D.7 the operator B is a Hilbert-Schmidt operator. Hence according to Proposition D.12 the operator $\widetilde{\Gamma}_0$ is nuclear. Therefore $\widetilde{\Gamma}_0$ is an appropriate covariance operator for a Gaussian random variable and it is a smoothing operator. In future we shall mark it without the tilde.

In this model case we assume that our model for the flow is rather accurate. Hence we use the same covariance operator for the Wiener process than for the initial value.

Discretization Space

We need a family $\{\mathcal{V}_m\}_{m=1}^{\infty}$ of finite dimensional subspaces of $L^2(\mathbb{R})$ satisfying the following conditions

(i) $\mathcal{V}_m \subseteq \mathcal{V}_{m+1}$ for all $m \in \mathbb{N}$,

(ii) $\overline{\cup_m \mathcal{V}_m} = L^2(\mathbb{R})$ and

(iii) $P_m f \to f$ in $L^2(\mathbb{R})$ as $m \to \infty$ for all $f \in L^2(\mathbb{R})$ where P_m is the orthogonal projection from $L^2(\mathbb{R})$ to $\mathcal{V}_m$.

Let us choose

$$\mathcal{V}_m := \operatorname{span}\left\{\sqrt{m}\chi_{[\frac{l-1}{m}-m, \frac{l}{m}-m]},\ l = 1, \ldots, 2m^2\right\}$$

for all $m \in \mathbb{N}$. Then $\mathcal{V}_m \subseteq \mathcal{V}_{m+1}$ and $\dim \mathcal{V}_m \leq 2m^2$ for all $m \in \mathbb{N}$.

Lemma 6.11. *$\overline{\cup_m \mathcal{V}_m} = L^2(\mathbb{R})$.*

Proof. Since $\mathcal{V}_m \subset L^2(\mathbb{R})$ for all $m \in \mathbb{N}$, then $\overline{\cup_m \mathcal{V}_m}$ is a closed subspace of $L^2(\mathbb{R})$. We want to show that the orthocomplement of $\overline{\cup_m \mathcal{V}_m}$ is trivial. Let $f \in L^2(\mathbb{R})$ be such that $(f, \psi) = 0$ for all $\psi \in \overline{\cup_m \mathcal{V}_m}$. Specially, for all intervals $I \subset \mathbb{R}$ such that $m(I) < \infty$ we have $(f, \chi_I) = 0$ because $\chi_I \in \overline{\cup_m \mathcal{V}_m}$. Thus

$$\frac{1}{m(I)} \int_I f(x)\, dx = 0$$

for all intervals $I \subset \mathbb{R}$ such that $m(I) < \infty$. Since $f \in L^2(\mathbb{R})$, then $f \in L^1(I)$ for all intervals $I \subset \mathbb{R}$ such that $m(I) < \infty$. Since for L^1-functions almost all points are the Lebesgue points,

$$f(x) = \lim_{n \to \infty} \frac{1}{m(I_n)} \int_{I_n} f(x)\, dx = 0$$

for almost all $x \in \mathbb{R}$ where $I_n := (x - r_n, x + r_n)$ and $\lim_{n\to\infty} r_n = 0$. Hence $f \equiv 0$ and the orthocomplement of $\overline{\cup_m \mathcal{V}_m}$ is trivial. Thus $\overline{\cup_m \mathcal{V}_m} = L^2(\mathbb{R})$. $\square$

We denote

$$\psi_l^m := \sqrt{m}\chi_{[\frac{l-1}{m}-m, \frac{l}{m}-m]}$$

for all $l = 1, \ldots, 2m^2$ and $m \in \mathbb{N}$. Since $(\psi_i^m, \psi_j^m) = \delta_{ij}$ for all $i, j = 1, \ldots, 2m^2$, the family $\{\psi_l^m\}_{l=1}^{2m^2}$ is an orthonormal basis of $\mathcal{V}_m$ for all $m \in \mathbb{N}$. Thus $\dim \mathcal{V}_m = 2m^2$. We can define the orthogonal projections $P_m : L^2(\mathbb{R}) \to \mathcal{V}_m$ by

$$P_m f = \sum_{l=1}^{2m^2} (f, \psi_l^m)\psi_l^m = \sum_{l=-m^2+1}^{m^2} m \int_{\frac{l-1}{m}}^{\frac{l}{m}} f \, dx \chi_{[\frac{l-1}{m}, \frac{l}{m}]}$$

for all $f \in L^2(\mathbb{R})$.

Lemma 6.12. *$P_m f \to f$ in $L^2(\mathbb{R})$ as $m \to \infty$ for all $f \in L^2(\mathbb{R})$.*

Proof. We can change the basis of $\mathcal{V}_m$ such a way that the new basis $\{\varphi_l^m\}_{l=1}^{2m^2}$ is an orthonormal basis of $\mathcal{V}_m$ and

$$\{\varphi_l^{m-1}\}_{l=1}^{2(m-1)^2} \subset \{\varphi_l^m\}_{l=1}^{2m^2}$$

for all $m \in \mathbb{N}$. We start with the basis of $\mathcal{V}_1$. The new basis of $\mathcal{V}_2$ is made by adding linearly independent members of the old basis to the basis of $\mathcal{V}_1$ and by using the Gram-Schmidt orthogonalization procedure. So, the new basis at level m is obtained by adding linearly independent members of the old basis to the basis at level $m-1$ and by using the Gram-Schmidt orthogonalization procedure. Thus the basis is a growing family of functions and we can index them by the appearance. In this way we get an orthonormal basis $\{\varphi_l\}_{l=1}^{\infty}$ of $\cup_m \mathcal{V}_m$. The change of the basis does not change the projection, since

$$P_m f = \sum_{l=1}^{2m^2} (f, \psi_l^m)\psi_l^m = \sum_{l=1}^{2m^2} (f, \psi_l^m) \sum_{j=1}^{2m^2} (\psi_l^m, \varphi_j)\varphi_j = \sum_{j=1}^{2m^2} (f, \varphi_j)\varphi_j$$

for all $f \in L^2(\mathbb{R})$.

Let $\varepsilon > 0$ and $f \in L^2(\mathbb{R})$. Then by Lemma 6.11 there exists $f_\varepsilon \in \cup_m \mathcal{V}_m$ such that $\|f - f_\varepsilon\|_{L^2(\mathbb{R})} < \varepsilon/3$. Since $\{\varphi_l\}_{l=1}^{\infty}$ is an orthonormal basis of $\cup_m \mathcal{V}_m$, there exists $M_\varepsilon \in \mathbb{N}$ such that $\|f_\varepsilon - P_{M_\varepsilon} f_\varepsilon\|_{L^2(\mathbb{R})} < \varepsilon/3$. Thus

$$\begin{aligned}\|P_{M_\varepsilon} f - f\|_{L^2(\mathbb{R})} &\leq \|P_{M_\varepsilon} f - P_{M_\varepsilon} f_\varepsilon\|_{L^2(\mathbb{R})} + \|P_{M_\varepsilon} f_\varepsilon - f_\varepsilon\|_{L^2(\mathbb{R})} + \|f_\varepsilon - f\|_{L^2(\mathbb{R})} \\ &\leq \|P_{M_\varepsilon}\| \|f_\varepsilon - f\|_{L^2(\mathbb{R})} + \|P_{M_\varepsilon} f_\varepsilon - f_\varepsilon\|_{L^2(\mathbb{R})} + \|f_\varepsilon - f\|_{L^2(\mathbb{R})} < \varepsilon.\end{aligned}$$

Hence $P_m f \to f$ in $L^2(\mathbb{R})$ as $n \to \infty$ for all $f \in L^2(\mathbb{R})$. Furthermore, $\{\varphi_l\}_{l=1}^{\infty}$ is a basis of $L^2(\mathbb{R})$. □

According to Lemmas 6.11 and 6.12 the family $\{\mathcal{V}_m\}_{m=1}^{\infty}$ form a family of appropriate discretization spaces in $L^2(\mathbb{R})$. The basis functions of $\mathcal{V}_m$ are the simplest one, constant functions with finite supports.

Discretized Evolution Equation

The choice of the discretization level m depends on how accurate and how fast computation we want to have. The support of a function in $\mathcal{V}_m$ belongs to the interval $[-m, m]$. Since we know that the measurements give information only from a part of the pipe, the discretization level need not to be bigger than half of the width of the measurement region. By the calculation in Subsection 6.3.4 the discretized evolution equation is

$$C_{k+1}^m = A_{k+1}^m C_k^m + E_{k+1}^m + W_{k+1}^m$$

for all $k = 0, \dots, n-1$. The matrix A_{k+1}^m is defined by $(A_{k+1}^m)_{ij} := (\mathcal{U}(\Delta_k)\psi_j^m, \psi_i^m)$ for all $i, j = 1, \dots, 2m^2$. We are able to calculate the elements of the matrix A_{k+1}^m for all $k = 0, \dots, n-1$. First of all, for all $l = 1, \dots, 2m^2$, $t > 0$ and $x \in \mathbb{R}$

$$\begin{aligned}
\mathcal{U}(t)\psi_l^m(x) &= \frac{\sqrt{m}}{2\sqrt{\pi\kappa t}} \int_{\frac{l-1}{m}-m}^{\frac{l}{m}-m} e^{-\frac{(x-y-vt)^2}{4\kappa t}}\, dy = \sqrt{\frac{m}{\pi}} \int_{\frac{mx-l+m^2-mvt}{2m\sqrt{\kappa t}}}^{\frac{mx-l+1+m^2-mvt}{2m\sqrt{\kappa t}}} e^{-z^2}\, dz \\
&= \sqrt{m}\left[I\left(\frac{mx-l+1+m^2-mvt}{2m\sqrt{\kappa t}}\right) - I\left(\frac{mx-l+m^2-mvt}{2m\sqrt{\kappa t}}\right)\right]
\end{aligned}$$

where

$$I(x) := \frac{1}{\sqrt{\pi}} \int_{-\infty}^{x} e^{-z^2}\, dz = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \int_{0}^{x} e^{-z^2}\, dz = \frac{1}{2} + 2\,\mathrm{erf}(x)$$

for all $x \in \mathbb{R}$ and erf is the so called *error function*. Thus

$$\begin{aligned}
&(\mathcal{U}(\Delta_k)\psi_j^m, \psi_i^m) \\
&= m \int_{\frac{i-1}{m}-m}^{\frac{i}{m}-m} \left[I\left(\frac{mx-j+1+m^2-mv\Delta_k}{2m\sqrt{\kappa\Delta_k}}\right) - I\left(\frac{mx-j+m^2-mv\Delta_k}{2m\sqrt{\kappa\Delta_k}}\right)\right] dx \\
&= 2m\sqrt{\kappa\Delta_k} \int_{\frac{i-j-mv\Delta_k}{2m\sqrt{\kappa\Delta_k}}}^{\frac{i-j+1-mv\Delta_k}{2m\sqrt{\kappa\Delta_k}}} I(y)\, dy - 2m\sqrt{\kappa\Delta_k} \int_{\frac{i-j-1-mv\Delta_k}{2m\sqrt{\kappa\Delta_k}}}^{\frac{i-j-mv\Delta_k}{2m\sqrt{\kappa\Delta_k}}} I(y)\, dy
\end{aligned}$$

for all $i, j = 1, \dots, 2m^2$. Since

$$\int_{-\infty}^{x} I(y)\, dy = \frac{e^{-x^2}}{2\sqrt{\pi}} + xI(x) = \frac{e^{-x^2}}{2\sqrt{\pi}} + \frac{x}{2} + 2x\,\mathrm{erf}(x)$$

for all $x \in \mathbb{R}$, the elements of the matrix A_{k+1}^m are given by functions known by mathematical softwares.

Since both E_{k+1}^m and W_{k+1}^m are Gaussian random variables, the knowledge of the means and covariance and correlation operators is sufficient to be able to present the distribution of C_{k+1}^m for all $k = 0, \dots, n-1$. By the calculation in the previous subsection only the vector $\mathbb{E}E_{k+1}^m$ and matrices $\mathrm{Cov}\, E_{k+1}^m$, $\mathrm{Cov}\, W_{k+1}^m$ and $\mathrm{Cor}(C_k^m, E_{k+1}^m)$ for all $k = 0, \dots, n-1$ are required. According to Formulas (6.21)–(6.23) it is enough to know how to calculate the inner products $(\mathcal{U}(t)c_0, \psi_i^m)$ and $(\mathcal{U}(t)\Gamma_0\mathcal{U}^*(s)\psi_i^m, \psi_j^m)$ and the integral

$$\int_r^s (\mathcal{U}(u-\tau)Q\mathcal{U}^*(t-\tau)\psi_i^m, \psi_j^m)\, d\tau \tag{6.45}$$

for all $i, j = 1, \dots, 2m^2$ and $0 \le r \le s \le t \le u \le T$. We shall need the adjoint operator of $\mathcal{U}(t)$ for all $t > 0$. Let $f, g \in L^2(\mathbb{R})$. Then

$$\begin{aligned}
(\mathcal{U}(t)f, g) &= \int_{-\infty}^{\infty} (\Phi(t,\cdot) * f)(x) g(x)\, dx \\
&= \int_{-\infty}^{\infty} f(y) \int_{-\infty}^{\infty} \frac{1}{2\sqrt{\pi\kappa t}} e^{-\frac{(x-y-vt)^2}{4\kappa t}} g(x)\, dx dy \\
&= \int_{-\infty}^{\infty} f(y)(\Phi^*(t,\cdot) * g)(y)\, dy
\end{aligned}$$

where

$$\Phi^*(t,x) := \frac{1}{2\sqrt{\pi\kappa t}} \exp\left(-\frac{(x+vt)^2}{4\kappa t}\right)$$

for all $t > 0$ and $x \in \mathbb{R}$. Thus

$$\begin{cases} \mathcal{U}^*(0)f = f, \\ (\mathcal{U}^*(t)f)(x) = (\Phi^*(t,\cdot) * f)(x), \quad t > 0, \end{cases}$$

for all $f \in L^2(\mathbb{R})$. Hence similarly as above

$$\mathcal{U}^*(t)\psi_l^m(x) = \sqrt{m}\left[I\left(\frac{mx - l + 1 + m^2 + mvt}{2m\sqrt{\kappa t}}\right) - I\left(\frac{mx - l + m^2 + mvt}{2m\sqrt{\kappa t}}\right)\right]$$

for all $l = 1, \ldots, 2m^2$, $t > 0$ and $x \in \mathbb{R}$. The function c_0 is given by Formula (6.44). Since the mean of the initial value has exponentially decaying tails, we need to know how to calculate integrals of $e^{-\beta|x|}I(x)$ for some $\beta > 0$. Let $\alpha \in \mathbb{R}$. Then

$$\begin{aligned}\int_{-\infty}^{\alpha} e^{\beta x} I(x)\, dx &= \frac{1}{\beta}e^{\alpha\beta}I(\alpha) - \frac{1}{\beta}e^{\frac{\beta^2}{4}} I\left(\alpha - \frac{\beta}{2}\right) \\ &= \frac{1}{2\beta}\left(e^{\alpha\beta} - e^{\frac{\beta^2}{4}}\right) + \frac{2}{\beta}\left(e^{\alpha\beta}\operatorname{erf}(\alpha) - e^{\frac{\beta^2}{4}}\operatorname{erf}\left(\alpha - \frac{\beta}{2}\right)\right)\end{aligned}$$

and

$$\begin{aligned}\int_{\alpha}^{\infty} e^{-\beta x} I(x)\, dx &= \frac{1}{\beta}e^{-\alpha\beta}I(\alpha) + \frac{1}{\beta}e^{\frac{\beta^2}{4}}\left(1 - I\left(\alpha + \frac{\beta}{2}\right)\right) \\ &= \frac{1}{2\beta}\left(e^{-\alpha\beta} + e^{\frac{\beta^2}{4}}\right) + \frac{2}{\beta}\left(e^{-\alpha\beta}\operatorname{erf}(\alpha) - e^{\frac{\beta^2}{4}}\operatorname{erf}\left(\alpha + \frac{\beta}{2}\right)\right).\end{aligned}$$

Therefore $(\mathcal{U}(t)c_0, \psi_l^m) = (c_0, \mathcal{U}^*(t)\psi_l^m)$ for all $l = 1, \ldots, 2m^2$ is given by functions known by the mathematical softwares because

$$\begin{aligned}&(c_0, \mathcal{U}^*(t)\psi_l^m) \\ &= c_0\sqrt{m}\int_{-\infty}^{-M} e^{x-M}\left[I\left(\frac{x - \frac{l-1}{m} + m + vt}{2\sqrt{\kappa t}}\right) - I\left(\frac{x - \frac{l}{m} + m + vt}{2\sqrt{\kappa t}}\right)\right] dx + \\ &\quad + c_0\sqrt{m}\int_{-M}^{M}\left[I\left(\frac{x - \frac{l-1}{m} + m + vt}{2\sqrt{\kappa t}}\right) - I\left(\frac{x - \frac{l}{m} + m + vt}{2\sqrt{\kappa t}}\right)\right] dx + \\ &\quad + c_0\sqrt{m}\int_{M}^{\infty} e^{-x+M}\left[I\left(\frac{x - \frac{l-1}{m} + m + vt}{2\sqrt{\kappa t}}\right) - I\left(\frac{x - \frac{l}{m} + m + vt}{2\sqrt{\kappa t}}\right)\right] dx \\ &= 2c_0\sqrt{m\kappa t}\, e^{-M + \frac{l-1}{m} - m - vt}\int_{-\infty}^{\frac{-mM - l + 1 + m^2 + mvt}{2m\sqrt{\kappa t}}} e^{2\sqrt{\kappa t}y} I(y)\, dy + \\ &\quad - 2c_0\sqrt{m\kappa t}\, e^{-M + \frac{l}{m} - m - vt}\int_{-\infty}^{\frac{-mM - l + m^2 + mvt}{2m\sqrt{\kappa t}}} e^{2\sqrt{\kappa t}y} I(y)\, dy + \\ &\quad + 2c_0\sqrt{m\kappa t}\int_{\frac{-mM - l + 1 + m^2 + mvt}{2m\sqrt{\kappa t}}}^{\frac{mM - l + 1 + m^2 + mvt}{2m\sqrt{\kappa t}}} I(z)\, dz - 2c_0\sqrt{m\kappa t}\int_{\frac{-mM - l + m^2 + mvt}{2m\sqrt{\kappa t}}}^{\frac{mM - l + m^2 + mvt}{2m\sqrt{\kappa t}}} I(z)\, dz \\ &\quad + 2c_0\sqrt{m\kappa t}\, e^{M - \frac{l-1}{m} + m + vt}\int_{\frac{mM - l + 1 + m^2 + mvt}{2m\sqrt{\kappa t}}}^{\infty} e^{-2\sqrt{\kappa t}y} I(y)\, dy + \\ &\quad + 2c_0\sqrt{m\kappa t}\, e^{M - \frac{l}{m} + m + vt}\int_{\frac{mM - l + m^2 + mvt}{2m\sqrt{\kappa t}}}^{\infty} e^{-2\sqrt{\kappa t}y} I(y)\, dy.\end{aligned}$$

We have chosen that the covariance operators Γ_0 and Q are the integral operator with the integral kernel $w(x)\gamma_0(x-y)w(y)$ where $\hat{\gamma}_0(\xi) = (1+\xi^2)^{-2}$ for all $\xi \in \mathbb{R}$ and w is exponentially decaying at infinity, $w(x) = 1$ when $|x| < N$ with some $N > 0$ and $0 < w(x) \leq 1$ for all $x \in \mathbb{R}$. We are not able to calculate the inner product $(\mathcal{U}(t)\Gamma_0\mathcal{U}^*(s)\psi_i^m, \psi_j^m)$ for $i, j = 1, \ldots, 2m^2$ and $0 \leq s \leq t \leq T$ in a closed form. By using the Parseval formula we notice that

$$\begin{aligned}(\mathcal{U}(t)\Gamma_0\mathcal{U}^*(s)\psi_i^m, \psi_j^m) &= (\gamma_0 * (w\mathcal{U}^*(s)\psi_i^m), w\mathcal{U}^*(t)\psi_j^m) \\ &= (\hat{\gamma}_0\mathcal{F}(w\mathcal{U}^*(s)\psi_i^m), \mathcal{F}(w\mathcal{U}^*(t)\psi_j^m))\end{aligned}$$

for all $i, j = 1, \ldots, 2m^2$ and $0 \leq s \leq t \leq T$. The Fourier transform of γ_0 is known. We can use the fast Fourier transform (FFT) algorithm to compute the Fourier transform of $w\mathcal{U}^*(t)\psi_i^m$ for all $i = 1, \ldots, 2m^2$ and $0 \leq t \leq T$. We need use some numerical quadrature to calculate an approximation of the integral

$$\int_{-\infty}^{\infty} \hat{\gamma}_0(\xi)\mathcal{F}(w\mathcal{U}^*(s)\psi_i^m)(\xi)\mathcal{F}(w\mathcal{U}^*(t)\psi_j^m)(\xi)\, d\xi.$$

In addition, Integral (6.45) has to be computed numerically. Consequently, we have all information needed to perform the evolution updating step of the Bayesian filtering.

6.4 Conclusions

In this thesis we have examined the non-stationary inverse problem concerning electrical impedance process tomography. We have viewed it as a state estimation problem. We have presented the continuous infinite dimensional state estimation system corresponding to the problem. By studying the infinite dimensional evolution equation and linearizing the observation equation we have been able to introduce the discretized state estimation system relating to the problem. The finite dimensional state estimation problem has been solved in the Gaussian context by using the Bayesian filtering method. However, the method introduced in Section 6.3 can be applied to all non-stationary inverse problems in which the time evolution is modeled by a linear stochastic differential equation with a sectorial operator and the observation equation is linear or linearizable.

The solution (6.33)–(6.36) to the Bayesian filtering is valid only in the Gaussian case. The assumption of Gaussianity seemed to be natural since the solution to the infinite dimensional state evolution equation is a Gaussian process if the initial value is assumed be a Gaussian random variable. Despite of the initial value the state noise is always a Gaussian process. In some application non-Gaussian initial values may be reasonable. Nonetheless, the non-Gaussian case is beyond the scope of this thesis.

The main weakness of the method introduced in this thesis is the use of analytic semigroups in solving the infinite dimensional state evolution equation. Since the analytic semigroup is defined by using the spectral properties of the infinitesimal generator, only in some special cases we can present the analytic semigroup in a closed form. Some other ways of solving infinite dimensional linear stochastic differential equations should be researched.

The examination of the continuous infinite dimensional state evolution equation is beneficial for solving non-stationary inverse problems by the state estimation method. The numerical computation requires space discretization. The discretization error can be considered only by knowing the stochastic nature of the time evolution of the object of interest. As was seen in the electrical impedance process tomography problem in Section 6.3 the knowledge of the continuous infinite dimensional state evolution equation of the concentration distribution allows us to calculate the probability distribution of the discretization errors both in the evolution and observation equation. When the discretization error is taken into account, the state estimation system is discretization invariant and hence the solution to the non-stationary inverse problem does not depend on the discretization. Then we need not choose the discretization level as high as possible for ensuring the accuracy of the computation. If the aim in the Bayesian filtering is to have a real time monitoring of the object of interest, we may use such a discretization level that the computation is fast enough. Other discretization invariant estimation methods have been developed in the PhD thesis of Lasanen [25].

In Subsection 6.3.5 we have studied a one dimensional version of the process tomography problem. Since electrical impedance tomography is not defined in the one dimension, the model case only illustrates the time evolution model used in the electrical impedance process tomography problem. By numerical implementation of the one dimensional version we would be able to visualize the discretization invariance of the method. Unfortunately, we were not able to include numerical results to this thesis. They will be presented in further publications of the author.

Appendix A

Resolvent

In this appendix we introduce basic properties of the resolvent set and operator of a linear operator. Let $(E, \|\cdot\|_E)$ be a Banach space. We denote by $B(E)$ the space of bounded linear operators from E to E equipped with the operator norm

$$\|A\|_{B(E)} := \sup\{\|Ax\|_E : x \in E,\ \|x\|_E \leq 1\}$$

for all $A \in B(E)$.

Definition A.1. *Let $A : \mathcal{D}(A) \subset E \to E$ be linear. The resolvent set $\rho(A)$ and the spectrum $\sigma(A)$ of the operator A are*

$$\rho(A) := \{\lambda \in \mathbb{C} : \exists\ (\lambda I - A)^{-1} \in B(E)\} \quad \textit{and} \quad \sigma(A) := \mathbb{C} \setminus \rho(A).$$

If $\lambda \in \rho(A)$, we denote $R(\lambda, A) := (\lambda I - A)^{-1}$. The operator $R(\lambda, A)$ is said to be the *resolvent operator* or simply the resolvent of the operator A. The so called *resolvent identity*

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda)R(\lambda, A)R(\mu, A)$$

is valid for all $\lambda, \mu \in \rho(A)$.

Let $D \subset \mathbb{C}$ be open. The function $\lambda \mapsto T(\lambda)$ from D to $B(E)$ is said to be *analytic* (or *holomorphic*) if for every disc $B(a, r)$ in D there exists a series

$$\sum_{n=0}^{\infty} A_n(\lambda - a)^n$$

where $A_n \in B(E)$ which converges in $B(E)$ to $T(\lambda)$ for all $\lambda \in B(a, r)$.

Proposition A.2. *Let $\lambda_0 \in \rho(A)$. Then the disc*

$$B\left(\lambda_0, \|R(\lambda_0, A)\|_{B(E)}^{-1}\right) := \left\{\lambda \in \mathbb{C} : |\lambda - \lambda_0| < \|R(\lambda_0, A)\|_{B(E)}^{-1}\right\}$$

is contained in $\rho(A)$ and for all λ in that disc

$$\begin{aligned} R(\lambda, A) &= R(\lambda_0, A)[I + (\lambda - \lambda_0)R(\lambda_0, A)]^{-1} \\ &= \sum_{n=0}^{\infty} (-1)^n (\lambda - \lambda_0)^n R^{n+1}(\lambda_0, A). \end{aligned} \tag{A.1}$$

Therefore the resolvent set $\rho(A)$ is open in $\mathbb{C}$ and the mapping $\lambda \mapsto R(\lambda, A)$ is analytic in $\rho(A)$.

Proof. Let $\lambda_0 \in \rho(A)$. For every $y \in E$ the equation $\lambda x - Ax = y$ is equivalent to the equation $z + (\lambda - \lambda_0)R(\lambda_0, A)z = y$ where $z := (\lambda_0 - A)x$. If $\|(\lambda - \lambda_0)R(\lambda_0, A)\|_{B(E)} < 1$, then $I + (\lambda - \lambda_0)R(\lambda_0, A)$ is invertable with a bounded inverse. Hence

$$x = R(\lambda_0, A)\left[I + (\lambda - \lambda_0)R(\lambda_0, A)\right]^{-1} y = \sum_{n=0}^{\infty} (-1)^n (\lambda - \lambda_0)^n R^{n+1}(\lambda_0, A)y.$$

Thus

$$R(\lambda, A) = R(\lambda_0, A)\left[I + (\lambda - \lambda_0)R(\lambda_0, A)\right]^{-1} = \sum_{n=0}^{\infty} (-1)^n (\lambda - \lambda_0)^n R^{n+1}(\lambda_0, A)$$

if $|\lambda - \lambda_0| < \|R(\lambda_0, A)\|_{B(E)}^{-1}$. So the disc

$$\left\{\lambda \in \mathbb{C} : |\lambda - \lambda_0| < \|R(\lambda_0, A)\|_{B(E)}^{-1}\right\}$$

belongs to $\rho(A)$ and Expansion (A.1) is valid in that disc. Therefore $\rho(A)$ is open and the mapping $\lambda \mapsto R(\lambda, A)$ is holomorphic in $\rho(A)$. $\square$

Appendix B

Vector Valued Functions

In this appendix we introduce the Bochner integration theory for Banach space valued functions. In Section B.1 we have gathered measure theoretical notation used in this appendix. The Bochner integration theory for functions with values in an arbitrary Banach space is presented in Sections B.2 and B.4. The special case of operator valued functions is considered in Sections B.3 and B.5. The main references of this appendix are the books of Hille and Phillips [16] and Kuttler [24]. The Bochner integral can also be found among others in the books of Diestel and Uhl [9] and Yosida [54], in the master's thesis of Hytönen [17] and in the PhD thesis of Mikkola [29].

B.1 Basic Definitions of Measure Theory

In this section we recall the basic notation of the measure theory on account of consistence. Nevertheless, we assume the Lebesgue integration theory for scalar valued functions to be known. The books of Kuttler [24] or Rudin [37] can be used as a reference.

Let Ω be a set. A collection $\mathcal{F}$ of subsets of Ω is said to be a *σ-algebra* in Ω if $\mathcal{F}$ has the following properties

(i) $\Omega \in \mathcal{F}$,

(ii) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$,

(iii) if $A = \cup_{n=1}^{\infty} A_n$ and $A_n \in \mathcal{F}$ for all $n \in \mathbb{N}$, then $A \in \mathcal{F}$.

If $\mathcal{F}$ is a σ-algebra in Ω, then $(\Omega, \mathcal{F})$ is called a *measurable space* and the members of $\mathcal{F}$ are called the *measurable sets* in Ω. Let $(\Omega, \mathcal{F})$ and $(E, \mathcal{G})$ be measurable spaces. A function $x : \Omega \to E$ is said to be *measurable* if $x^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{G}$, i.e., the inverse image of a measurable set is measurable.

Let E be a topological space. Then the *Borel σ-algebra* of E is the smallest σ-algebra containing all open subsets of E. It is denoted by $\mathcal{B}(E)$ and the elements of

$\mathcal{B}(E)$ are called the *Borel sets* of E. An E-valued measurable function is a mapping $x : \Omega \to E$ which is measurable from $(\Omega, \mathcal{F})$ to $(E, \mathcal{B}(E))$.

Let $(\Omega, \mathcal{F})$ be a measurable space. A function $\mu : \mathcal{F} \to [0, \infty]$ is a *positive measure* if $\mu(A) < \infty$ at least for one $A \in \mathcal{F}$ and μ is σ-additive, i.e., if $\{A_i\}$ is a disjoint countable collection of measurable sets,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

The triplet $(\Omega, \mathcal{F}, \mu)$ is called a *measure space*. The measure space $(\Omega, \mathcal{F}, \mu)$ is σ-*finite* if Ω is a countable union of sets Ω_i with $\mu(\Omega_i) < \infty$.

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $(E, \mathcal{G})$ a measurable vector space. A function $x : \Omega \to E$ is called *simple* if it is of the form

$$x(\omega) = \sum_{k=1}^{n} a_k \chi_{A_k}(\omega)$$

for all $\omega \in \Omega$ where $n \in \mathbb{N}$, $a_k \in E$, $A_k \in \mathcal{F}$ such that $\mu(A_k) < \infty$ for all $k = 1, \ldots, n$ and $A_i \cap A_j = \emptyset$ if $i \neq j$, and

$$\chi_{A_k}(\omega) := \begin{cases} 1 & \text{if } \omega \in A_k, \\ 0 & \text{if } \omega \notin A_k. \end{cases}$$

A simple function has only a finite number of values and the measure of the set in which a simple function is nonzero is finite.

B.2 Strong and Weak Measurability

Let $(\Omega, \mathcal{F}, \mu)$ be a σ-finite measure space and $(E, \|\cdot\|_E)$ a Banach space. We define two different kinds of measurabilities for functions from Ω to E. They will be used in the definition of the Bochner integral.

Definition B.1. *(i) A function $x : \Omega \to E$ is said to be strongly measurable if there exists a sequence $\{x_n\}_{n=1}^{\infty}$ of simple functions converging pointwise to x.*

(ii) A function $x : \Omega \to E$ is said to be weakly measurable if for each $f \in E'$ the scalar function $\omega \mapsto \langle x(\omega), f \rangle$ is measurable.

Clearly, if x is measurable from $(\Omega, \mathcal{F})$ to $(E, \mathcal{B}(E))$, it is weakly measurable. Our task it to verify that strongly and weakly measurable functions are often measurable and vice versa. The separability of the range of a function is the necessary and sufficient condition.

Theorem B.2. *A function x is strongly measurable if and only if it is measurable and $x(\Omega)$ is separable.*

Proof. "$\Leftarrow$" Let us assume that x is measurable and $x(\Omega)$ is separable. Let $\{a_k\}_{k=1}^{\infty}$ be dense in $x(\Omega)$ and $n \in \mathbb{N}$. We set for all $k = 1, \ldots, n$

$$U_k^n := \left\{ z \in E : \|z - a_k\|_E \leq \min_{1 \leq l \leq n} \|z - a_l\|_E \right\}$$

and $B_k^n := x^{-1}(U_k^n)$. Then U_k^n is a Borel set in E and hence B_k^n is measurable for all $k = 1, \ldots, n$. We form disjoint sets D_k^n by

$$D_k^n := B_k^n \setminus \left(\bigcup_{l<k} B_l^n \right)$$

for all $k = 1, \ldots, n$. We define a sequence $\{x_n\}_{n=1}^{\infty}$ of measurable functions by

$$x_n(\omega) := \sum_{k=1}^{n} a_k \chi_{D_k^n}(\omega)$$

for all $\omega \in \Omega$. Thus x_n is the nearest approximation of x in the set $\{a_k\}_{k=1}^n$. Since $\{a_k\}_{k=1}^{\infty}$ is dense in $x(\Omega)$, the functions x_n converge pointwise to x. Since $(\Omega, \mathcal{F}, \mu)$ is σ-finite, there exist sets $\Omega_n \uparrow \Omega$ such that $\mu(\Omega_n) < \infty$. We define $y_n(\omega) := \chi_{\Omega_n}(\omega) x_n(\omega)$ for all $n \in \mathbb{N}$. Then y_n converges pointwise to x since for all $\omega \in \Omega$ and n large enough $\omega \in \Omega_n$. Functions y_n are simple because $\mu(\Omega_n) < \infty$. Hence x is strongly measurable.

"$\Rightarrow$" Let us assume that x is strongly measurable. Then there exists a sequence $\{x_n\}_{n=1}^{\infty}$ of simple functions converging pointwise to x. A set $x_n^{-1}(U)$ is measurable for all open $U \subset E$ and $n \in \mathbb{N}$ since it is a finite union of measurable sets. Let $U \subset E$ be open and $\{V_m\}_{m=1}^{\infty}$ a sequence of open sets satisfying the conditions

$$\overline{V}_m \subseteq U, \quad \overline{V}_m \subseteq V_{m+1} \quad \text{and} \quad U = \bigcup_{m=1}^{\infty} V_m.$$

Then

$$x^{-1}(V_m) \subseteq \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} x_k^{-1}(V_m) \subseteq x^{-1}(\overline{V}_m)$$

since x_n converges pointwise to x. Hence

$$x^{-1}(U) = \bigcup_{m=1}^{\infty} x^{-1}(V_m) \subseteq \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} x_k^{-1}(V_m) \subseteq \bigcup_{m=1}^{\infty} x^{-1}(\overline{V}_m) \subseteq x^{-1}(U).$$

Thus

$$x^{-1}(U) = \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} x_k^{-1}(V_m).$$

Since $x_k^{-1}(V_m)$ is measurable for all $k, m \in \mathbb{N}$, then $x^{-1}(U)$ is measurable for all open $U \subset E$ as a countable intersection and unions of measurable sets. Since $\{C : x^{-1}(C) \in \mathcal{F}\}$ is a σ-algebra containing all open sets, $x^{-1}(B)$ is measurable for all Borel sets B in E. Hence x is measurable.

We still have to show that $x(\Omega)$ is separable. Let

$$D := \{a \in E : \exists n \in \mathbb{N} \text{ such that } x_n(\omega) = a \text{ for some } \omega \in \Omega\}.$$

Then $x_n(\Omega) \subset D$ for all $n \in \mathbb{N}$ and D is a countable set and dense in $\bar{D}$. Since x_n converges pointwise to x, then $x(\Omega) \subset \bar{D}$. As a subset of a separable set $x(\Omega)$ is separable. $\square$

To be able to prove that the measurabilities coincide if the ranges of functions are separable, we need the following lemma. The lemma is interesting itself.

Lemma B.3. *If E is separable and x is weakly measurable, the scalar function $\|x(\cdot)\|_E$ is measurable.*

Proof. Let x be a weakly measurable function. Let us set

$$A := \{\omega \in \Omega : \|x(\omega)\|_E \leq a\} \quad \text{and} \quad A_f := \{\omega \in \Omega : |\langle x(\omega), f\rangle| \leq a\}$$

where $a \in \mathbb{R}$ and $f \in E'$. It is enough to prove that A is measurable. By the definition,

$$A \subseteq \bigcap_{\|f\|_{E'} \leq 1} A_f.$$

If $\omega \in \cap_{\|f\|_{E'} \leq 1} A_f$, then $|\langle x(\omega), f\rangle| \leq a$ for all $f \in B'$ where B' is the unit ball of E'. But by the Hahn-Banach theorem there exists $f_\omega \in E'$ such that $\|f_\omega\|_{E'} = 1$ and $\langle x(\omega), f_\omega\rangle = \|x(\omega)\|_E$. Hence $\|x(\omega)\|_E \leq a$ and

$$A = \bigcap_{\|f\|_{E'} \leq 1} A_f.$$

We need the functional analytic fact that for separable Banach spaces the unit ball of the dual space is weak*-separable.

Lemma B.4. [54, pp. 131–132] *Let E be a separable Banach space and B' the unit ball in E'. Then there exists a sequence $\{f_n\}_{n=1}^\infty \subseteq B'$ with the property that for every $f_0 \in B'$ there exists a subsequence $\{f_{n_k}\}_{k=1}^\infty$ of $\{f_n\}_{n=1}^\infty$ such that $\langle x, f_0\rangle = \lim_{k\to\infty}\langle x, f_{n_k}\rangle$ for every $x \in E$.*

Let $\{f_n\}_{n=1}^\infty$ be the sequence in Lemma B.4. If $\omega \in \cap_{n=1}^\infty A_{f_n}$, then $|\langle x(\omega), f_n\rangle| \leq a$ for all n. Therefore $|\langle x(\omega), f\rangle| = \lim_{k\to\infty} |\langle x(\omega), f_{n_k}\rangle| \leq a$ for all $f \in B'$. Thus

$$A = \bigcap_{\|f\|_{E'} \leq 1} A_f = \bigcap_{n=1}^\infty A_{f_n}.$$

Since x is weakly measurable, A_{f_n} is measurable for all $n \in \mathbb{N}$. As a countable intersection of measurable sets A is measurable. Hence the scalar function $\|x(\cdot)\|_E$ is measurable. $\square$

The following theorem combines the strong and weak measurability.

Theorem B.5. *A function x is strongly measurable if and only if it is weakly measurable and $x(\Omega)$ is separable.*

Proof. "$\Rightarrow$" Let x be strongly measurable. Then $x(\Omega)$ is separable by Theorem B.2. Since x is strongly measurable, there exists a sequence $\{x_n\}_{n=1}^\infty$ of simple functions such that x_n converges pointwise to x. Then by the continuity of the dual operation $\langle x_n(\omega), f\rangle \to \langle x(\omega), f\rangle$ as $n \to \infty$ for each $\omega \in \Omega$ and for all $f \in E'$. Since $\langle x(\cdot), f\rangle$ is a limit of measurable complex functions, it is measurable. Hence x is weakly measurable.

"$\Leftarrow$" Let x be weakly measurable and $x(\Omega)$ separable. Let $\{a_k\}_{k=1}^{\infty}$ be dense in $x(\Omega)$. Then $x - a_k$ is weakly measurable for all $k \in \mathbb{N}$. Since $x(\Omega)$ separable, by Lemma B.3 the norm $\|x(\cdot) - a_k\|_E$ is measurable for all $k \in \mathbb{N}$. Let $\varepsilon > 0$. We define for all $k \in \mathbb{N}$

$$A_k^{\varepsilon} := \{\omega \in \Omega : \|x(\omega) - a_k\|_E < \varepsilon\}.$$

Then $A_k^{\varepsilon} \in \mathcal{F}$ and $\cup_k A_k^{\varepsilon} = \Omega$ since $\{a_k\}_{k=1}^{\infty}$ is dense in $x(\Omega)$. We set

$$D_k^{\varepsilon} := A_k^{\varepsilon} \setminus \bigcup_{l<k} A_l^{\varepsilon}$$

for all $k \in \mathbb{N}$. Then the sets D_k^{ε} are measurable, disjoint and $\cup_k D_k^{\varepsilon} = \Omega$. We define

$$x_{\varepsilon}(\omega) := \sum_{k=1}^{\infty} a_k \chi_{D_k^{\varepsilon}}(\omega)$$

for all $\omega \in \Omega$. Clearly, x_{ε} is a countable valued function and $\|x(\omega) - x_{\varepsilon}(\omega)\|_E < \varepsilon$ for all $\omega \in \Omega$. Since $(\Omega, \mathcal{F}, \mu)$ is σ-finite, there exist sets $\Omega_n \uparrow \Omega$ such that $\mu(\Omega_n) < \infty$. We define for all $n \in \mathbb{N}$ and $\omega \in \Omega$

$$y_n(\omega) := \sum_{k=1}^{n} a_k \chi_{D_k^{\frac{1}{n}} \cap \Omega_n}(\omega) = \chi_{\Omega_n}(\omega) \chi_{\bigcup_{k \le n} D_k^{\frac{1}{n}}}(\omega) x_{\frac{1}{n}}(\omega).$$

Then $\{y_n\}_{n=1}^{\infty}$ is a sequence of simple function converging pointwise to x. Thus x is strongly measurable. $\square$

We have actually proved a somewhat stronger result that the statement of the theorem would indicate.

Corollary B.6. *A function x is strongly measurable if and only if it is the uniform limit of a sequence of countable valued functions.*

The following corollary is the summary of this section.

Corollary B.7. *Let x be a function from Ω to E. Then the following three statements are equivalent:*

(i) x is measurable and $x(\Omega)$ is separable,

(ii) x is strongly measurable,

(iii) x is weakly measurable and $x(\Omega)$ is separable.

In a separable Banach space all three measurabilities are equivalent.

B.3 Operator Valued Functions

Let $(\Omega, \mathcal{F}, \mu)$ be a σ-finite measure space and $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ Banach spaces. We denote by $B(E, F)$ the space of bounded linear operator from E to F with the operator norm

$$\|U\|_{B(E,F)} := \sup\{\|Ux\|_F : x \in E,\ \|x\|_E \le 1\}$$

for all $U \in B(E,F)$. Since $B(E,F)$ is a Banach space, the theory introduced in the previous section can be applied to operator valued functions. Nevertheless, it is convenient to define the strong and weak measurability for operator valued functions differently than in Definition B.1.

Definition B.8. *(i) The operator valued function $U : \Omega \to B(E,F)$ is said to be uniformly measurable if there exists a sequence of simple operator valued functions converging pointwise to U in the uniform operator topology.*

(ii) The operator valued function $U : \Omega \to B(E,F)$ is said to be strongly measurable if the vector valued function $U(\cdot)x$ is strongly measurable in the sense of Definition B.1 for all $x \in E$.

(iii) The operator valued function $U : \Omega \to B(E,F)$ is said to be weakly measurable if the vector valued function $U(\cdot)x$ is weakly measurable in the sense of Definition B.1 for all $x \in E$.

It is clear that the uniform measurability of an operator valued function U is the same as the strong measurability of U considered as a vector valued function in the Banach space $B(E,F)$.

The connection between the three different types of measurability for operator valued functions is given by the following theorem.

Theorem B.9. *(i) The operator valued function U is strongly measurable if and only if it is weakly measurable and $U(\Omega)x$ is separable for each $x \in E$.*

(ii) The operator valued function U is uniformly measurable if and only if it is weakly measurable and $U(\Omega)$ is separable.

Proof. The statement (i) is an immediate consequence of Theorem B.5. The statement (ii) is not as obvious. The proof is similar to the one of Theorem B.5.

"$\Rightarrow$" Let us assume that U is uniformly measurable. Then there exists a sequence $\{U_n\}_{n=1}^{\infty}$ of simple operator valued functions converging pointwise to U in the uniform operator topology, i.e., $\|U_n(\omega) - U(\omega)\|_{B(E,F)} \to 0$ as $n \to \infty$ for all $\omega \in \Omega$. Thus for each $x \in E$ and $f \in F'$

$$\begin{aligned} |\langle U_n(\omega)x, f\rangle - \langle U(\omega)x, f\rangle| &\le \|U_n(\omega)x - U(\omega)x\|_F \|f\|_{F'} \\ &\le \|U_n(\omega) - U(\omega)\|_{B(E,F)} \|x\|_E \|f\|_{F'} \longrightarrow 0 \end{aligned}$$

as $n \to \infty$ for all $\omega \in \Omega$. Finite valued scalar functions $\langle U_n(\cdot)x, f\rangle$ are measurable for each $x \in E$ and $f \in F'$. Therefore $\langle U(\cdot)x, f\rangle$ is measurable for each $x \in E$ and $f \in F'$ as a limit of measurable scalar functions. Hence U is weakly measurable.

We still have to show that $U(\Omega)$ is separable. Let

$$D := \{A \in B(E,F) : \exists n \in \mathbb{N} \text{ such that } U_n(\omega) = A \text{ for some } \omega \in \Omega\}.$$

Then $U_n(\Omega) \subset D$ for all $n \in \mathbb{N}$. In addition, D is a countable set and dense in $\bar{D}$. Since U_n converges pointwise to U, then $U(\Omega) \subset \bar{D}$. As a subset of a separable set $U(\Omega)$ is separable.

"$\Leftarrow$" Let U be weakly measurable and $U(\Omega)$ separable. Let $\{U_n\}_{n=1}^{\infty} \subset B(E,F)$ be dense in $U(\Omega)$. For every $n \in \mathbb{N}$ we can find a sequence $\{x_m^n\}_{m=1}^{\infty} \subset E$ such that $\|x_m^n\|_E = 1$ and

$$\|U_n x_m^n\|_F \geq \|U_n\|_{B(E,F)} - \frac{1}{m}$$

for all $m \in \mathbb{N}$. Since $U(\Omega)$ is separable, also $U(\Omega)x$ is separable for all $x \in E$. Since U is weakly measurable, U is strongly measurable by Theorem B.5. Since U is strongly measurable, for every $x \in E$ there exists a sequence $\{f_n^x\}_{n=1}^{\infty}$ of simple F-valued functions such that f_n^x converges pointwise to $U(\cdot)x$. Thus

$$|\|U(\omega)x\|_F - \|f_n^x(\omega)\|_F| \leq \|U(\omega)x - f_n^x(\omega)\|_F \longrightarrow 0$$

as $n \to \infty$ for all $\omega \in \Omega$. Hence $\|U(\cdot)x\|_F$ is measurable for each $x \in E$ as a limit of measurable scalar functions. Thus $\|U(\cdot)x_m^n\|_F$ is measurable for all $m, n \in \mathbb{N}$. Also the function

$$F(\omega) := \sup_{m,n\in\mathbb{N}} \|U(\omega)x_m^n\|_F$$

for all $\omega \in \Omega$ is measurable. Clearly, $F(\omega) \leq \|U(\omega)\|_{B(E,F)}$ for all $\omega \in \Omega$. Actually, an equality holds. For given $\omega \in \Omega$ and $m \in \mathbb{N}$ there exists $n \in \mathbb{N}$ depending on ω and m such that

$$\|U(\omega) - U_n\|_{B(E,F)} \leq \frac{1}{m}.$$

Hence for all $m \in \mathbb{N}$ and $\omega \in \Omega$

$$\begin{aligned} F(\omega) &\geq \|U(\omega)x_m^n\|_F \geq \|U_n x_m^n\|_F - \|U(\omega)x_m^n - U_n x_m^n\|_F \\ &\geq \|U_n\|_{B(E,F)} - \|U(\omega) - U_n\|_{B(E,F)} - \frac{1}{m} \\ &\geq \|U_n\|_{B(E,F)} - \frac{2}{m} \geq \|U\|_{B(E,F)} - \frac{3}{m}. \end{aligned}$$

Thus $F(\omega) = \|U(\omega)\|_{B(E,F)}$ for all $\omega \in \Omega$ and $\|U(\cdot)\|_{B(E,F)}$ is measurable.

Since U is weakly measurable, $U - U_n$ is weakly measurable for all $n \in \mathbb{N}$. Hence $\|U(\cdot) - U_n\|_{B(E,F)}$ is measurable for all $n \in \mathbb{N}$. Let $\varepsilon > 0$. We define for all $n \in \mathbb{N}$

$$A_n^{\varepsilon} := \{\omega \in \Omega : \|U(\omega) - U_n\|_{B(E,F)} < \varepsilon\}.$$

Then $A_n^{\varepsilon} \in \mathcal{F}$ and $\cup_n A_n^{\varepsilon} = \Omega$ since $\{U_n\}_{n=1}^{\infty}$ is dense in $U(\Omega)$. We set

$$D_n^{\varepsilon} := A_n^{\varepsilon} \setminus \bigcup_{k<n} A_k^{\varepsilon}$$

for all $n \in \mathbb{N}$. Then the sets D_n^{ε} are measurable, disjoint and $\cup_n D_n^{\varepsilon} = \Omega$. We define

$$U_{\varepsilon}(\omega) := \sum_{n=1}^{\infty} U_n \chi_{D_n^{\varepsilon}}(\omega).$$

Clearly, U_{ε} is a countable valued function and $\|U(\omega) - U_{\varepsilon}(\omega)\|_{B(E,F)} < \varepsilon$ for all $\omega \in \Omega$. Since $(\Omega, \mathcal{F}, \mu)$ is σ-finite, there exist sets $\Omega_n \uparrow \Omega$ such that $\mu(\Omega_n) < \infty$. We define for all $m \in \mathbb{N}$ and $\omega \in \Omega$

$$V_m(\omega) := \sum_{n=1}^{m} U_n \chi_{D_n^{\frac{1}{m}} \cap \Omega_m}(\omega) = \chi_{\Omega_m}(\omega) \chi_{\bigcup_{k\leq m} D_k^{\frac{1}{m}}}(\omega) U_{\frac{1}{m}}(\omega).$$

Then $\{V_m\}_{m=1}^{\infty}$ is a sequence of simple operator valued function converging pointwise to U in the uniform operator topology. Thus U is uniformly measurable. $\square$

B.4 The Bochner Integral

Let $(\Omega, \mathcal{F}, \mu)$ be a σ-finite measure space and $(E, \|\cdot\|_E)$ a Banach space. Let x be a simple function

$$x = \sum_{k=1}^{n} a_k \chi_{A_k}$$

from Ω to E. We define the *Bochner integral* of x to be

$$\int_\Omega x(\omega)\, d\mu := \sum_{k=1}^{n} a_k \mu(A_k).$$

Then the Bochner integral is well defined and linear on the set of simple functions.

Definition B.10. *A strongly measurable function x is Bochner integrable if there exists a sequence $\{x_n\}_{n=1}^\infty$ of simple functions converging pointwise to x and satisfying*

$$\int_\Omega \|x_n(\omega) - x_m(\omega)\|_E \, d\mu \longrightarrow 0 \tag{B.1}$$

as $m, n \to \infty$. If x is Bochner integrable, we define the Bochner integral of x to be

$$\int_\Omega x(\omega)\, d\mu := \lim_{n\to\infty} \int_\Omega x_n(\omega)\, d\mu.$$

We need to prove that the previous definition is appropriate.

Lemma B.11. *The Bochner integral of a Bochner integrable function is well defined.*

Proof. Let x be a simple function. Then

$$\begin{aligned}\left\| \int_\Omega x(\omega)\, d\mu \right\|_E &= \left\| \sum_{k=1}^{n} a_k \mu(A_k) \right\|_E \leq \sum_{k=1}^{n} \|a_k\|_E \mu(A_k) \\ &= \int_\Omega \sum_{k=1}^{n} \|a_k\|_E \chi_{A_k}(\omega)\, d\mu = \int_\Omega \|x(\omega)\|_E \, d\mu.\end{aligned}$$

Hence

$$\left\| \int_\Omega x(\omega)\, d\mu \right\|_E \leq \int_\Omega \|x(\omega)\|_E \, d\mu$$

for each simple function x.

Let x be a Bochner integrable function and $\{x_n\}_{n=1}^\infty$ a sequence of simple functions converging pointwise to x and satisfying Condition (B.1). Then $\left\{\int_\Omega x_n(\omega)\, d\mu\right\}_{n=1}^\infty$ is a Cauchy sequence in E since

$$\left\| \int_\Omega x_n(\omega)\, d\mu - \int_\Omega x_m(\omega)\, d\mu \right\|_E \leq \int_\Omega \|x_n(\omega) - x_m(\omega)\|_E \, d\mu \longrightarrow 0$$

as $m, n \to \infty$. Since E is complete,

$$\lim_{n\to\infty} \int_\Omega x_n(\omega)\, d\mu$$

exists in E.

We need to show that the Bochner integral does not depend on the choice of the sequence satisfying Condition (B.1). Let $\{x_n\}_{n=1}^{\infty}$ and $\{y_m\}_{m=1}^{\infty}$ be two sequences satisfying Condition (B.1) and converging pointwise to x. Let $\varepsilon > 0$. Then by Fatou's lemma,

$$\begin{aligned}
&\left\| \int_\Omega x_n(\omega)\, d\mu - \int_\Omega y_m(\omega)\, d\mu \right\|_E \\
&\leq \int_\Omega \|x_n(\omega) - y_m(\omega)\|_E \, d\mu \\
&\leq \int_\Omega \|x_n(\omega) - x(\omega)\|_E \, d\mu + \int_\Omega \|y_m(\omega) - x(\omega)\|_E \, d\mu \\
&\leq \liminf_{k\to\infty} \int_\Omega \|x_n(\omega) - x_k(\omega)\|_E \, d\mu + \liminf_{k\to\infty} \int_\Omega \|y_m(\omega) - y_k(\omega)\|_E \, d\mu \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

if $m, n \in \mathbb{N}$ are large enough. Since $\varepsilon > 0$ is arbitrary, this shows that the integrals converge to the same limit. Hence the Bochner integral is well defined. □

There is an equivalent way to define Bochner integrable functions.

Theorem B.12. *A function x is Bochner integrable if and only if it is strongly measurable and*

$$\int_\Omega \|x(\omega)\|_E \, d\mu < \infty.$$

If x is Bochner integrable, there exists a sequence $\{y_n\}_{n=1}^{\infty}$ of simple functions such that y_n converges pointwise to x and satisfies $\|y_n(\omega)\|_E \leq 2\|x(\omega)\|_E$ for all $n \in \mathbb{N}$ and $\omega \in \Omega$ and

$$\int_\Omega \|y_n(\omega) - y_m(\omega)\|_E \, d\mu \longrightarrow 0$$

as $m, n \to \infty$. In addition,

$$\lim_{n\to\infty} \int_\Omega \|x(\omega) - y_n(\omega)\|_E \, d\mu = 0.$$

Proof. "$\Rightarrow$" Let x be Bochner integrable. By the definition it is strongly measurable. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of simple functions satisfying Condition (B.1) and converging pointwise to x. By Fatou's lemma,

$$\int_\Omega \|x(\omega)\|_E \, d\mu \leq \liminf_{n\to\infty} \int_\Omega \|x_n(\omega)\|_E \, d\mu.$$

The right hand side is finite since the sequence $\{x_n\}_{n=1}^{\infty}$ satisfies Condition (B.1) and thus $\left\{\int_\Omega \|x_n(\omega)\|_E \, d\mu\right\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$. So

$$\int_\Omega \|x(\omega)\|_E \, d\mu < \infty$$

for each Bochner integrable x.

"$\Leftarrow$" Let x be a strongly measurable function and

$$\int_\Omega \|x(\omega)\|_E \, d\mu < \infty.$$

Then there exists a sequence $\{x_n\}_{n=1}^\infty$ of simple functions converging pointwise to x. Let us define

$$y_n(\omega) := \begin{cases} x_n(\omega) & \text{if } \|x_n(\omega)\|_E \leq 2\|x(\omega)\|_E, \\ 0 & \text{if } \|x_n(\omega)\|_E > 2\|x(\omega)\|_E, \end{cases}$$

for all $n \in \mathbb{N}$. If $x(\omega) = 0$, then $y_n(\omega) = 0$ for all $n \in \mathbb{N}$. If $\|x(\omega)\|_E > 0$, for all $n \in \mathbb{N}$ large enough $y_n(\omega) = x_n(\omega)$. Thus y_n converges pointwise to x and $\|y_n(\omega)\|_E \leq 2\|x(\omega)\|_E$ for all $n \in \mathbb{N}$ and $\omega \in \Omega$. Since $\|y_n(\omega) - y_m(\omega)\|_E \leq 4\|x(\omega)\|_E$ for all $m, n \in \mathbb{N}$ and $\omega \in \Omega$ and $\|x(\cdot)\|_E \in L^1(\Omega)$, by Lebesgue's dominated convergence theorem,

$$\lim_{m,n\to\infty} \int_\Omega \|y_n(\omega) - y_m(\omega)\|_E \, d\mu = 0.$$

Hence the sequence $\{y_n\}_{n=1}^\infty$ satisfies Condition (B.1). Therefore x is Bochner integrable.

Since y_n converges pointwise to x and $\|y_n(\omega) - y_m(\omega)\|_E \leq 4\|x(\omega)\|_E$ for all $m, n \in \mathbb{N}$ and $\omega \in \Omega$ and $\|x(\cdot)\|_E \in L^1(\Omega)$,

$$0 = \lim_{n\to\infty} \lim_{m\to\infty} \int_\Omega \|y_n(\omega) - y_m(\omega)\|_E \, d\mu = \lim_{n\to\infty} \int_\Omega \|y_n(\omega) - x(\omega)\|_E \, d\mu$$

by Lebesgue's dominated convergence theorem. □

Main properties of the Bochner integral are presented in the following theorem.

Theorem B.13. *(i) If x and $\tilde{x}$ are Bochner integrable and $\alpha, \beta \in \mathbb{C}$,*

$$\alpha \int_\Omega x(\omega) \, d\mu + \beta \int_\Omega \tilde{x}(\omega) \, d\mu = \int_\Omega (\alpha x(\omega) + \beta \tilde{x}(\omega)) \, d\mu,$$

i.e., the Bochner integral is a linear operator from the set of Bochner integrable functions to E.

(ii) If x is Bochner integrable,

$$\left\| \int_\Omega x(\omega) \, d\mu \right\|_E \leq \int_\Omega \|x(\omega)\|_E \, d\mu.$$

(iii) Let F be a Banach space. If x is Bochner integrable and A is a bounded linear operator from E to F,

$$A \int_\Omega x(\omega) \, d\mu = \int_\Omega Ax(\omega) \, d\mu.$$

Proof. (i) Let $\{y_n\}_{n=1}^\infty$ and $\{\tilde{y}_n\}_{n=1}^\infty$ be sequences of simple functions stated in Theorem B.12 corresponding to the Bochner integrable functions x and $\tilde{x}$, respectively.

If $\alpha, \beta \in \mathbb{C}$, then $\alpha x + \beta \tilde{x}$ is Bochner integrable since the simple function $\alpha y_n + \beta \tilde{y}_n$ converges pointwise to $\alpha x + \beta \tilde{x}$ and

$$\begin{aligned}&\int_\Omega \|\alpha y_n(\omega) + \beta \tilde{y}_n(\omega) - (\alpha y_m(\omega) + \beta \tilde{y}_m(\omega))\|_E \, d\mu \\ &\le |\alpha| \int_\Omega \|y_n(\omega) - y_m(\omega)\|_E \, d\mu + |\beta| \int_\Omega \|\tilde{y}_n(\omega) - \tilde{y}_m(\omega)\|_E \, d\mu \longrightarrow 0\end{aligned}$$

as $m, n \to \infty$. The Bochner integral is linear on the set of simple functions. Hence

$$\alpha \int_\Omega y_n(\omega) \, d\mu + \beta \int_\Omega \tilde{y}_n(\omega) \, d\mu = \int_\Omega (\alpha y_n(\omega) + \beta \tilde{y}_n(\omega)) \, d\mu$$

for all $\alpha, \beta \in \mathbb{C}$. Therefore

$$\begin{aligned}\alpha \int_\Omega x(\omega) \, d\mu + \beta \int_\Omega \tilde{x}(\omega) \, d\mu &= \lim_{n\to\infty} \left(\alpha \int_\Omega y_n(\omega) \, d\mu + \beta \int_\Omega \tilde{y}_n(\omega) \, d\mu \right) \\ &= \lim_{n\to\infty} \int_\Omega (\alpha y_n(\omega) + \beta \tilde{y}_n(\omega)) \, d\mu \\ &= \int_\Omega (\alpha x(\omega) + \beta \tilde{x}(\omega)) \, d\mu\end{aligned}$$

for all $\alpha, \beta \in \mathbb{C}$.

(ii) Let $\{y_n\}_{n=1}^\infty$ be a sequence of simple functions stated in Theorem B.12 corresponding to the Bochner integrable function x. In the proof of Lemma B.11 we showed that

$$\left\| \int_\Omega x(\omega) \, d\mu \right\|_E \le \int_\Omega \|x(\omega)\|_E \, d\mu$$

for each simple function x. By the continuity of a norm,

$$\left\| \int_\Omega x(\omega) \, d\mu \right\|_E = \lim_{n\to\infty} \left\| \int_\Omega y_n(\omega) \, d\mu \right\|_E \le \lim_{n\to\infty} \int_\Omega \|y_n(\omega)\|_E \, d\mu.$$

Since y_n converges pointwise to x, $\|y_n(\omega)\|_E \le 2\|x(\omega)\|_E$ for all $n \in \mathbb{N}$ and $\omega \in \Omega$ and $\|x(\cdot)\|_E \in L^1(\Omega)$, according to Lebesgue's dominated convergence theorem,

$$\left\| \int_\Omega x(\omega) \, d\mu \right\|_E \le \int_\Omega \|x(\omega)\|_E \, d\mu$$

for each Bochner integrable x.

(iii) Let $A \in B(E, F)$ and x be a simple function. Then

$$\begin{aligned}A \int_\Omega x(\omega) \, d\mu &= A \sum_{k=1}^n a_k \mu(A_k) = \sum_{k=1}^n A a_k \mu(A_k) \\ &= \int_\Omega \sum_{k=1}^n A a_k \chi_{A_k}(\omega) \, d\mu = \int_\Omega A x(\omega) \, d\mu.\end{aligned}$$

Let $\{y_n\}_{n=1}^\infty$ be a sequence of simple functions stated in Theorem B.12 corresponding to the Bochner integrable function x. Then Ay_n converges pointwise to Ax and

$$\int_\Omega \|Ay_n(\omega) - Ay_m(\omega)\|_F \, d\mu \le \|A\|_{B(E,F)} \int_\Omega \|y_n(\omega) - y_m(\omega)\|_E \, d\mu \longrightarrow 0$$

as $m, n \to \infty$ Hence Ax is Bochner integrable. By the continuity of the operator A,

$$A \int_{\Omega} x(\omega)\, d\mu = \lim_{n\to\infty} A \int_{\Omega} y_n(\omega)\, d\mu = \lim_{n\to\infty} \int_{\Omega} A y_n(\omega)\, d\mu = \int_{\Omega} Ax(\omega)\, d\mu$$

for each Bochner integrable x. □

Theorem B.12 allows us to define the space of Bochner integrable functions.

Definition B.14. *A function x belongs to the space $L^p(\Omega, \mathcal{F}, \mu; E)$ for $1 \leq p < \infty$ if x is strongly measurable and*

$$\int_{\Omega} \|x(\omega)\|_E^p \, d\mu < \infty.$$

We identify two functions in $L^p(\Omega, \mathcal{F}, \mu; E)$ if they are equal almost everywhere, i.e., $x(\omega) = y(\omega)$ for all $\omega \in \Omega \setminus A$ with $\mu(A) = 0$.

We denote

$$\|x\|_p := \|x\|_{L^p(\Omega,\mathcal{F},\mu;E)} := \left(\int_{\Omega} \|x(\omega)\|_E^p \, d\mu \right)^{\frac{1}{p}}. \tag{B.2}$$

It is clear that $L^p(\Omega, \mathcal{F}, \mu; E)$ is a norm space with the norm given by Formula (B.2). In fact, $L^p(\Omega, \mathcal{F}, \mu; E)$ is a Banach space. In the proof of the completeness we use the following lemma.

Lemma B.15. *If $\{x_n\}_{n=1}^{\infty}$ is a sequence in $L^p(\Omega, \mathcal{F}, \mu; E)$ satisfying*

$$\sum_{n=1}^{\infty} \|x_{n+1} - x_n\|_p < \infty,$$

there exists $x \in L^p(\Omega, \mathcal{F}, \mu; E)$ such that x_n converges pointwise almost everywhere and in $L^p(\Omega, \mathcal{F}, \mu; E)$ to x.

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in $L^p(\Omega, \mathcal{F}, \mu; E)$ satisfying the assumption. We denote

$$g_n(\omega) := \sum_{k=1}^{n} \|x_{k+1}(\omega) - x_k(\omega)\|_E$$

for all $n \in \mathbb{N}$ and $\omega \in \Omega$. Then g_n belongs to $L^p(\Omega)$ since it is measurable and

$$\|g_n\|_{L^p(\Omega)} \leq \sum_{k=1}^{n} \|x_{k+1} - x_k\|_p < \infty$$

for all $n \in \mathbb{N}$. We set

$$g(\omega) := \lim_{n\to\infty} g_n(\omega) = \sum_{k=1}^{\infty} \|x_{k+1}(\omega) - x_k(\omega)\|_E$$

for all $\omega \in \Omega$. By Lebesgue's monotone convergence theorem,

$$\|g\|_{L^p(\Omega)} = \lim_{n\to\infty} \|g_n\|_{L^p(\Omega)} \leq \sum_{k=1}^{\infty} \|x_{k+1} - x_k\|_p < \infty.$$

Hence $g \in L^p(\Omega)$ and thus $g(\omega) < \infty$ for almost all $\omega \in \Omega$. We mark the set in which $g(\omega) = \infty$ with A. Then $\mu(A) = 0$.

Let $m > n$. Then

$$x_m(\omega) - x_n(\omega) = \sum_{k=n}^{m-1} (x_{k+1}(\omega) - x_k(\omega))$$

for all $\omega \in \Omega$. Thus

$$\|x_m(\omega) - x_n(\omega)\|_E \leq \sum_{k=n}^{m-1} \|x_{k+1}(\omega) - x_k(\omega)\|_E \leq \sum_{k=n}^{\infty} \|x_{k+1}(\omega) - x_k(\omega)\|_E$$

for all $\omega \in \Omega$. If $\omega \notin A$,

$$\sum_{k=n}^{\infty} \|x_{k+1}(\omega) - x_k(\omega)\|_E \longrightarrow 0$$

as $n \to \infty$. Hence if $\omega \notin A$, then $\{x_n(\omega)\}_{n=1}^{\infty}$ is a Cauchy sequence in E and $\lim_{n\to\infty} x_n(\omega)$ exists. We denote

$$x(\omega) := \begin{cases} \lim_{n\to\infty} x_n(\omega) & \text{if } \omega \notin A, \\ 0 & \text{if } \omega \in A. \end{cases}$$

Since x_n is strongly measurable, $x_n(\Omega)$ is separable for all $n \in \mathbb{N}$ by Theorem B.2. Let $\{a_k^n\}_{k=1}^{\infty}$ be a dense subset of $x_n(\Omega)$ for all $n \in \mathbb{N}$. Let us set $D := \{a_k^n\}_{k,n=1}^{\infty}$. Then $D \cap x_n(\Omega)$ is dense in $x_n(\Omega)$ and $\bar{D}$ is separable. Additionally, by the definition $x(\Omega) \subset \bar{D}$. Hence $x(\Omega)$ is separable. If $f \in E'$, according to continuity of the dual operation $\langle x(\omega), f\rangle = \lim_{n\to\infty}\langle x_n(\omega), f\rangle$ if $\omega \notin A$, and $\langle x(\omega), f\rangle = 0$ if $\omega \in A$. Hence $\langle x(\cdot), f\rangle$ is measurable as a limit of measurable scalar functions $\langle x_n(\cdot)\chi_{A^c}(\cdot), f\rangle$. So x is weakly measurable. Since x is weakly measurable and $x(\Omega)$ is separable, x is strongly measurable by Theorem B.5.

Let $\varepsilon > 0$. By Fatou's lemma,

$$\int_\Omega \|x(\omega) - x_n(\omega)\|_E^p \, d\mu \leq \liminf_{m\to\infty} \int_\Omega \|x_m(\omega) - x_n(\omega)\|_E^p \, d\mu.$$

For n large enough

$$\|x_m - x_n\|_p \leq \left[\int_\Omega \left(\sum_{k=n}^{m-1} \|x_{k+1}(\omega) - x_k(\omega)\|_E\right)^p\right]^{\frac{1}{p}} \leq \sum_{k=n}^{\infty} \|x_{k+1} - x_k\|_p < \varepsilon$$

if $m > n$. Since $\varepsilon > 0$ is arbitrary, $\lim_{n\to\infty} \|x - x_n\|_p = 0$.

We still have to prove that $x \in L^p(\Omega, \mathcal{F}, \mu; E)$. Let $\varepsilon > 0$. For $n \in \mathbb{N}$ large enough

$$\|x\|_p \leq \|x - x_n\|_p + \|x_n\|_p \leq \|x_n\|_p + \varepsilon < \infty.$$

Hence $x \in L^p(\Omega, \mathcal{F}, \mu; E)$. □

Theorem B.16. *$L^p(\Omega, \mathcal{F}, \mu; E)$ is complete. In addition, every Cauchy sequence has a subsequence which converges pointwise almost everywhere.*

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence in $L^p(\Omega, \mathcal{F}, \mu; E)$. We choose such a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ that $\|x_{n_{k+1}} - x_{n_k}\|_p \leq 2^{-k}$. Then the subsequence satisfies the assumptions of Lemma B.15. By the lemma the subsequence converges pointwise almost everywhere and in $L^p(\Omega, \mathcal{F}, \mu; E)$ to $x \in L^p(\Omega, \mathcal{F}, \mu; E)$. Since a limit is unique, the theorem is proved. □

By Theorem B.12 we know that the set of simple functions is dense in $L^1(\Omega, \mathcal{F}, \mu; E)$. The next theorem states that simple functions are also dense in $L^p(\Omega, \mathcal{F}, \mu; E)$ for $p > 1$ if the measure μ is finite.

Theorem B.17. *Assume that $\mu(\Omega) < \infty$. If $x \in L^p(\Omega, \mathcal{F}, \mu; E)$ for $p \geq 1$, there exists a sequence $\{x_n\}_{n=1}^{\infty}$ of simple functions such that x_n converges pointwise almost everywhere and in $L^p(\Omega, \mathcal{F}, \mu; E)$ to x and satisfies $\|x_n(\omega)\|_E \leq 2\|x(\omega)\|_E$ for all $n \in \mathbb{N}$ and almost all $\omega \in \Omega$.*

Proof. Simple functions belong to $L^p(\Omega, \mathcal{F}, \mu; E)$ for $p \geq 1$ since they are strongly measurable and

$$\|x\|_p^p = \sum_{k=1}^{n} \|a_k\|_E^p \mu(A_k) < \infty$$

for a simple function x. Let $x \in L^p(\Omega, \mathcal{F}, \mu; E)$. By Hölder's inequality,

$$\|x\|_{L^1(\Omega;E)} \leq \|1\|_{L^{\frac{p}{p-1}}(\Omega)} \|x\|_{L^p(\Omega;E)} = \mu(\Omega)^{1-\frac{1}{p}} \|x\|_{L^p(\Omega;E)} < \infty.$$

So x is Bochner integrable. By Theorem B.12 there exists a sequence $\{x_n\}_{n=1}^{\infty}$ of simple functions such that x_n converges pointwise almost everywhere and in $L^1(\Omega, \mathcal{F}, \mu; E)$ to x and satisfies $\|x_n(\omega)\|_E \leq 2\|x(\omega)\|_E$ for all $n \in \mathbb{N}$ and almost all $\omega \in \Omega$. Since $\|x(\omega) - x_n(\omega)\|_E^p \leq 3^p \|x(\omega)\|_E^p$ for all $n \in \mathbb{N}$ and almost all $\omega \in \Omega$, by Lebesgue's dominated convergence theorem,

$$\|x - x_n\|_p^p = \int_{\Omega} \|x(\omega) - x_n(\omega)\|_E^p \, d\mu \longrightarrow 0$$

as $n \to \infty$. □

Bochner integrable functions can be approximated by simple functions in L^1-norm. The following theorem states that the approximating functions can be chosen in such a way that the values of approximating functions are values of the original function if we allow the approximating functions to have countably many values.

Theorem B.18. *Let x be Bochner integrable and $\varepsilon > 0$. Then there exists a subdivision of Ω into disjoint sets $\{A_k\}_{k=1}^{\infty} \subset \mathcal{F}$ such that for an arbitrary $\omega_k \in A_k$ the function*

$$x_\varepsilon = \sum_{k=1}^{\infty} x(\omega_k) \chi_{A_k}$$

is countable valued, Bochner integrable and satisfies the relation

$$\int_{\Omega} \|x(\omega) - x_\varepsilon(\omega)\|_E \, d\mu < \varepsilon.$$

Furthermore, this remains valid for all refinements of the above subdivision.

Proof. Since $(\Omega, \mathcal{F}, \mu)$ is σ-finite, there exist $\Omega_n \uparrow \Omega$ such that $\mu(\Omega_n) < \infty$. We denote $\tilde{S}_n := \Omega_n \setminus \cup_{k<n}\Omega_k$. Then $\tilde{S}_n \in \mathcal{F}$ are disjoint and $\cup_n \tilde{S}_n = \Omega$. We may redefine a subdivision $\{S_n\}$ of Ω such that S_n are disjoint and $0 < \mu(S_n) < \infty$ for all $n \in \mathbb{N}$ by adding all $\tilde{S}_k$ such that $\mu(\tilde{S}_k) = 0$ to some $\tilde{S}_l$ with $\mu(\tilde{S}_l) > 0$.

Let x be Bochner integrable and $\varepsilon > 0$. Since x is strongly measurable, by Corollary B.6 for every $n \in \mathbb{N}$ there exists a countable valued function $x_{\varepsilon,n}$ such that

$$\|x_{\varepsilon,n}(\omega) - x(\omega)\|_E < \frac{2^{-n-1}\varepsilon}{\mu(S_n)}$$

for all $\omega \in \Omega$. For every $n \in \mathbb{N}$ let $\{a_l^n\}_{l=1}^\infty$ be the set of all values of the function $x_{\varepsilon,n}$ on S_n. We denote $A_l^n := x_{\varepsilon,n}^{-1}(a_l^n) \cap S_n$ for all $l, n \in \mathbb{N}$. Then A_l^n are disjoint for all $l, n \in \mathbb{N}$ and $\cup_{l=1}^\infty A_l^n = S_n$ for all $n \in \mathbb{N}$.

Let $\omega_l^n \in A_l^n$ be arbitrary for all $l, n \in \mathbb{N}$. We define the function

$$x_\varepsilon(\omega) := \sum_{l,n=1}^\infty x(\omega_l^n)\chi_{A_l^n}(\omega)$$

for all $\omega \in \Omega$. Then x_ε is countable valued and hence strongly measurable. Let $\omega \in A_l^n$ for some $l, n \in \mathbb{N}$. Then $x_{\varepsilon,n}(\omega) = x_{\varepsilon,n}(\omega_l^n) = a_l^n$. Hence

$$\|x_\varepsilon(\omega) - x(\omega)\|_E \le \|x(\omega_l^n) - x_{\varepsilon,n}(\omega_l^n)\|_E + \|x_{\varepsilon,n}(\omega) - x(\omega)\|_E < \frac{2^{-n}\varepsilon}{\mu(S_n)}$$

for all $\omega_l^n \in A_l^n$. Thus

$$\begin{aligned}\int_\Omega \|x(\omega) - x_\varepsilon(\omega)\|_E \, d\mu &\le \sum_{l,n=1}^\infty \int_\Omega \|x(\omega_l^n) - x(\omega)\|_E \chi_{A_l^n}(\omega) \, d\mu \\ &< \sum_{l,n=1}^\infty \frac{2^{-n}\varepsilon}{\mu(S_n)}\mu(A_l^n) = \sum_{n=1}^\infty 2^{-n}\varepsilon = \varepsilon.\end{aligned}$$

Furthermore,

$$\begin{aligned}\int_\Omega \|x_\varepsilon(\omega)\|_E \, d\mu &\le \int_\Omega \|x(\omega) - x_\varepsilon(\omega)\|_E \, d\mu + \int_\Omega \|x(\omega)\|_E \, d\mu \\ &< \int_\Omega \|x(\omega)\|_E \, d\mu + \varepsilon < \infty.\end{aligned}$$

Hence x_ε is Bochner integrable. The construction of the function x_ε allows all refinements. □

Bounded operators commute with the Bochner integral by Theorem B.13. The boundedness is not a necessary condition.

Theorem B.19. *Let F be a Banach space and $T : \mathcal{D}(T) \subset E \to F$ a closed linear operator. If $x \in L^1(\Omega, \mathcal{F}, \mu; E)$, $x(\omega) \in \mathcal{D}(T)$ for almost all $\omega \in \Omega$ and $Tx \in L^1(\Omega, \mathcal{F}, \mu; F)$,*

$$T \int_\Omega x(\omega) \, d\mu = \int_\Omega Tx(\omega) \, d\mu.$$

Proof. Let T be a closed linear operator from $\mathcal{D}(T) \subset E$ to F and x a Bochner integrable function such that $x(\omega) \in \mathcal{D}(T)$ for all $\omega \in \Omega$ and Tx is Bochner integrable. Let $\varepsilon > 0$. By Theorem B.18 there exists two subdivisions of Ω, one corresponding to an ε-approximation of x and the other an ε-approximation of Tx. Let $\{A_n\}_{n=1}^{\infty}$ be a common refinement of these two subdivisions and $\omega_n \in A_n$ for all $n \in \mathbb{N}$. We define

$$x_\varepsilon(\omega) := \sum_{n=1}^{\infty} x(\omega_n)\chi_{A_n}(\omega)$$

for all $\omega \in \Omega$. Then x_ε and Tx_ε are Bochner integrable and

$$\int_\Omega \|x(\omega) - x_\varepsilon(\omega)\|_E \, d\mu < \varepsilon \quad \text{and} \quad \int_\Omega \|Tx(\omega) - Tx_\varepsilon(\omega)\|_F \, d\mu < \varepsilon.$$

Thus

$$\int_\Omega x_\varepsilon(\omega) \, d\mu = \sum_{n=1}^{\infty} x(\omega_n)\mu(A_n) = \lim_{N\to\infty} \sum_{n=1}^{N} x(\omega_n)\mu(A_n)$$

and

$$\int_\Omega Tx_\varepsilon(\omega) \, d\mu = \sum_{n=1}^{\infty} Tx(\omega_n)\mu(A_n) = \lim_{N\to\infty} T\left(\sum_{n=1}^{N} x(\omega_n)\mu(A_n)\right).$$

Since T is a closed linear operator,

$$\int_\Omega x_\varepsilon(\omega) \, d\mu \in \mathcal{D}(T) \quad \text{and} \quad T\int_\Omega x_\varepsilon(\omega) \, d\mu = \int_\Omega Tx_\varepsilon(\omega) \, d\mu.$$

Let $\{\varepsilon_n\}_{n=1}^{\infty}$ be a sequence of positive numbers converging to zero. Since x_{ε_n} converges to x in $L^1(\Omega, \mathcal{F}, \mu; E)$ and Tx_{ε_n} to Tx in $L^1(\Omega, \mathcal{F}, \mu; F)$,

$$\lim_{n\to\infty} \int_\Omega x_{\varepsilon_n}(\omega) \, d\mu = \int_\Omega x(\omega) \, d\mu$$

and

$$\lim_{n\to\infty} T\int_\Omega x_{\varepsilon_n}(\omega) \, d\mu = \lim_{n\to\infty} \int_\Omega Tx_{\varepsilon_n}(\omega) \, d\mu = \int_\Omega Tx(\omega) \, d\mu.$$

Since T is a closed linear operator,

$$\int_\Omega x(\omega) \, d\mu \in \mathcal{D}(T) \quad \text{and} \quad T\int_\Omega x(\omega) \, d\mu = \int_\Omega Tx(\omega) \, d\mu.$$

If $x(\omega) \in \mathcal{D}(T)$ only for almost all $\omega \in \Omega$, the definitions of x_ε and Tx_ε have to be changed in a set of measure zero. Hence the statement is proved. $\square$

Since there is not an order relation in an arbitrary Banach space, there do not exist versions of Lebesgue's monotone convergence theorem and Fatou's lemma. But a modification of Lebesgue's dominated convergence theorem holds also in this setting.

Theorem B.20. *Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of strongly measurable functions. If x is strongly measurable and x_n converges pointwise almost everywhere to x with $\|x_n(\omega)\|_E \le g(\omega)$ for all $n \in \mathbb{N}$ and almost all $\omega \in \Omega$ where $g \in L^1(\Omega)$, then x is Bochner integrable and*

$$\int_\Omega x(\omega) \, d\mu = \lim_{n\to\infty} \int_\Omega x_n(\omega) \, d\mu.$$

Proof. Since $\|x_n(\omega)\|_E \le g(\omega)$ for all $n \in \mathbb{N}$ and almost all $\omega \in \Omega$ and $g \in L^1(\Omega)$, the functions x_n are Bochner integrable. Since $\|x_n(\omega) - x_m(\omega)\|_E \le 2g(\omega)$ for all $m, n \in \mathbb{N}$ and almost all $\omega \in \Omega$, by Lebesgue's dominated convergence theorem,

$$\lim_{m,n\to\infty} \int_\Omega \|x_n(\omega) - x_m(\omega)\|_E \, d\mu = 0.$$

So $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence in $L^1(\Omega, \mathcal{F}, \mu; E)$. According to Theorem B.16 there exists a subsequence $\{x_{n_k}\}_{k=1}^\infty$ and $y \in L^1(\Omega, \mathcal{F}, \mu; E)$ such that x_{n_k} converges pointwise almost everywhere and in $L^1(\Omega, \mathcal{F}, \mu; E)$ to y. But $\lim_{k\to\infty} x_{n_k}(\omega) = x(\omega)$ for almost all $\omega \in \Omega$. So $x = y$ almost everywhere. Hence x is Bochner integrable and x_n converges in $L^1(\Omega, \mathcal{F}, \mu; E)$ to x. Thus

$$\left\| \int_\Omega x(\omega)\, d\mu - \int_\Omega x_n(\omega)\, d\mu \right\|_E \le \int_\Omega \|x_n(\omega) - x(\omega)\|_E \, d\mu \longrightarrow 0$$

as $n \to \infty$. Therefore

$$\int_\Omega x(\omega)\, d\mu = \lim_{n\to\infty} \int_\Omega x_n(\omega)\, d\mu$$

and the theorem is proved. □

B.5 The Bochner Integral of Operator Valued Functions

We must distinguish between the uniform Bochner integral and strong Bochner integral of operator valued functions. If $U : \Omega \to B(E,F)$ is uniformly measurable and

$$\int_\Omega \|U(\omega)\|_{B(E,F)}\, d\mu < \infty,$$

$U \in L^1(\Omega, \mathcal{F}, \mu; B(E,F))$ and the theory of Section B.4 applies directly. In that case,

$$\int_\Omega U(\omega)\, d\mu \in B(E,F)$$

and is the limit in the uniform operator topology of the approximating integrals. The operator $\int_\Omega U(\omega)\, d\mu$ is called the *uniform Bochner integral* of U. On the other hand if U is strongly measurable and

$$\int_\Omega \|U(\omega)x\|_F\, d\mu < \infty$$

for all $x \in E$, i.e., $U(\cdot)x \in L^1(\Omega, \mathcal{F}, \mu; F)$ for each $x \in E$, the theory of Section B.4 merely asserts that

$$\int_\Omega U(\omega)x\, d\mu = V(x)$$

is an element of F. It requires additional argument to show that $V \in B(E,F)$.

Theorem B.21. *If $U(\cdot)x \in L^1(\Omega, \mathcal{F}, \mu; F)$ for each $x \in E$,*

$$Vx = \int_\Omega U(\omega)x\, d\mu$$

defines a bounded linear operator from E to F.

Proof. Let $U(\cdot)x \in L^1(\Omega, \mathcal{F}, \mu; F)$ for all $x \in E$. Then V is well defined and linear on E. In order to show that V is bounded we consider a transformation W from E to $L^1(\Omega, \mathcal{F}, \mu; F)$ defined by $(Wx)(\omega) := U(\omega)x$ for all $\omega \in \Omega$. One sees directly that W is linear. If $x_n \to x$ in E and $Wx_n \to y$ in $L^1(\Omega, \mathcal{F}, \mu; F)$ as $n \to \infty$, then $(Wx_n)(\omega) \to (Wx)(\omega)$ as $n \to \infty$ for all $\omega \in \Omega$ and there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ such that $(Wx_{n_k})(\omega) \to y(\omega)$ as $k \to \infty$ for almost all $\omega \in \Omega$. Since the limit is unique, $y(\omega) = (Wx)(\omega)$ for almost all $\omega \in \Omega$, i.e., $y = Wx$ in $L^1(\Omega, \mathcal{F}, \mu; F)$. Thus W is closed. By the closed graph theorem W is bounded. Therefore for all $x \in E$

$$\|Vx\|_F \leq \int_\Omega \|U(\omega)x\|_F \, d\mu = \|Wx\|_{L^1(\Omega;F)} \leq \|W\|_{B(E,L^1(\Omega,\mathcal{F},\mu;F))}\|x\|_E.$$

Hence V is bounded. □

The operator V is called the *strong Bochner integral* of U and denoted by

$$V := \int_\Omega U(\omega) \, d\mu.$$

Since uniformly measurable operator valued functions are strongly measurable, we have two different integrals for functions in $L^1(\Omega, \mathcal{F}, \mu; B(E, F))$. The following theorem shows that the integrals coincide.

Theorem B.22. *If $U \in L^1(\Omega, \mathcal{F}, \mu; B(E, F))$, the uniform and strong Bochner integral are equal.*

Proof. Let U be a simple operator valued function. Then

$$\left(\int_\Omega U(\omega) \, d\mu\right) x = \sum_{k=1}^{n} U_k x \mu(A_k) = \int_\Omega U(\omega)x \, d\mu$$

for all $x \in E$.

Let $U \in L^1(\Omega, \mathcal{F}, \mu; B(E, F))$. Then U is uniformly measurable and hence strongly measurable. Furthermore,

$$\int_\Omega \|U(\omega)\|_{B(E,F)} \, d\mu < \infty$$

and thus for each $x \in E$

$$\int_\Omega \|U(\omega)x\|_F \, d\mu \leq \int_\Omega \|U(\omega)\|_{B(E,F)} \, d\mu \, \|x\|_E < \infty.$$

Therefore both the uniform and strong Bochner integrals are defined. Since $U \in L^1(\Omega, \mathcal{F}, \mu; B(E, F))$, there exists a sequence $\{U_n\}_{n=1}^{\infty}$ of simple operator valued functions converging pointwise almost everywhere to U in the uniform operator topology and satisfying

$$\int_\Omega \|U_n(\omega) - U_m(\omega)\|_{B(E,F)} \, d\mu \longrightarrow 0$$

as $m, n \to \infty$. Thus for each $x \in E$ the sequence $\{U_n(\cdot)x\}_{n=1}^{\infty}$ of simple F-valued functions converges pointwise almost everywhere to $U(\cdot)x$ and satisfies

$$\int_{\Omega} \|U_n(\omega)x - U_m(\omega)x\|_F \, d\mu \longrightarrow 0$$

as $m, n \to \infty$. Hence for all $x \in E$

$$\left(\int_{\Omega} U(\omega) \, d\mu \right) x = \lim_{n \to \infty} \left(\int_{\Omega} U_n(\omega) \, d\mu \right) x = \lim_{n \to \infty} \int_{\Omega} U_n(\omega)x \, d\mu = \int_{\Omega} U(\omega)x \, d\mu$$

and therefore the uniform and strong Bochner integrals have the same value. □

Appendix C

Integration Along a Curve

In this appendix the Bochner integration theory introduced in Appendix B is used to define the integral of a vector valued function along a curve in the complex plane. This sort of integrals are needed in the definition of the analytic semigroup generated by a sectorial operator in Chapter 2.

Let $(E, \|\cdot\|_E)$ be a Banach space and γ a curve in $\mathbb{C}$, i.e., there exists such a parametrization

$$\gamma = \{\lambda \in \mathbb{C} : \lambda = \gamma(\varphi) := \gamma_1(\varphi) + i\gamma_2(\varphi),\ \varphi \in [a,b] \subset \mathbb{R}\}$$

where $a < b$ that γ_i, $i = 1, 2$, are piecewise continuously differentiable functions from $[a,b]$ to $\mathbb{R}$. We say that γ is a curve in a set $D \subset \mathbb{C}$ if $\gamma \subset D$. Let $x : \mathbb{C} \to E$ be a vector valued function. We define the integral of x along the curve γ to be the Bochner integral

$$\int_\gamma x(\lambda)\, d\lambda := \int_a^b x(\gamma(\varphi))\gamma'(\varphi)\, d\varphi$$

if $x(\gamma(\cdot))$ is strongly measurable from $[a,b]$ to E and

$$\int_a^b \|x(\gamma(\varphi))\|_E |\gamma'(\varphi)|\, d\varphi < \infty.$$

If $(F, \|\cdot\|_F)$ is a Banach space and $U : \mathbb{C} \to B(E,F)$ is an operator valued function, the integral of U along the curve γ can be defined as a uniform Bochner integral

$$\int_\gamma U(\lambda)\, d\lambda := \int_a^b U(\gamma(\varphi))\gamma'(\varphi)\, d\varphi$$

if $U(\gamma(\cdot))$ is uniformly measurable from $[a,b]$ to $B(E,F)$ and

$$\int_a^b \|U(\gamma(\varphi))\|_{B(E,F)} |\gamma'(\varphi)|\, d\varphi < \infty.$$

Then the integral

$$\int_\gamma U(\lambda)\, d\lambda$$

is a bounded linear operator from E to F.

C.1 Analytic Functions

We want to show that some of the results of the complex analysis are valid for operator valued functions.

Definition C.1. *Let $(E, \|\cdot\|_E)$ be a Banach space and $D \subset \mathbb{C}$ open. The function $x : D \to E$ is said to be holomorphic (or analytic) in D if for every disc $B(a,r) \subset D$ there exists a series*

$$\sum_{n=0}^{\infty} c_n(\lambda - a)^n$$

where $c_n \in E$ which converges in E to $x(\lambda)$ for all $\lambda \in B(a,r)$.

If $U : \mathbb{C} \to B(E,F)$ is analytic in an open set $D \subset \mathbb{C}$, the function $\langle U(\lambda)x, f\rangle$ is an analytic scalar function in D for all $x \in E$ and $f \in F'$. Let γ be a curve in D. Since the function $\varphi \mapsto \gamma(\varphi)$ is continuous, the function $\varphi \mapsto \langle U(\gamma(\varphi))x, f\rangle$ is measurable as a composite function of a continuous and analytic function for all $x \in E$ and $f \in F'$. Hence $U(\gamma(\cdot))$ is weakly measurable from $[a,b]$ to $B(E,F)$. Since $[a,b]$ is separable and $U(\gamma(\cdot))$ is continuous, $U(\gamma([a,b]))$ is separable. Therefore $U(\gamma(\cdot))$ is uniformly measurable by Theorem B.9. If the length $|\gamma| := \int_a^b |\gamma'(\varphi)|\, d\varphi$ of γ is finite,

$$\int_a^b \|U(\gamma(\varphi))\|_{B(E,F)} |\gamma'(\varphi)|\, d\varphi \le |\gamma| \max_{\varphi \in [a,b]} \|U(\gamma(\varphi))\|_{B(E,F)} < \infty.$$

Hence the integral

$$\int_\gamma U(\lambda)\, d\lambda$$

is defined for analytic functions U if the length of γ is finite. If there exists information about the behaviour of the norm of an analytic operator valued function, the integral along a curve with infinite length may be defined.

Let γ be a closed curve, i.e., $\gamma(a) = \gamma(b)$. By the Cauchy integral theorem,

$$\begin{aligned} 0 &= \oint_\gamma \langle U(\lambda)x, f\rangle\, d\lambda = \int_a^b \langle U(\gamma(\varphi))x, f\rangle \gamma'(\varphi)\, d\varphi \\ &= \left\langle \int_a^b U(\gamma(\varphi))x\gamma'(\varphi)\, d\varphi, f \right\rangle = \left\langle \oint_\gamma U(\lambda)x\, d\lambda, f \right\rangle \end{aligned}$$

for all $x \in E$ and $f \in F'$. Thus

$$\oint_\gamma U(\lambda)\, d\lambda = 0.$$

Therefore the Cauchy integral theorem is valid for holomorphic operator valued functions.

On the other hand, let γ be a closed curve and $\xi \notin \gamma$. By the Cauchy integral formula,

$$\begin{aligned} \langle U(\xi)x, f\rangle \operatorname{Ind}_\gamma(\xi) &= \frac{1}{2\pi i} \oint_\gamma \frac{\langle U(\lambda)x, f\rangle}{\xi - \lambda}\, d\lambda = \frac{1}{2\pi i} \int_a^b \frac{\langle U(\gamma(\varphi))x, f\rangle}{\xi - \gamma(\varphi)} \gamma'(\varphi)\, d\varphi \\ &= \left\langle \frac{1}{2\pi i} \int_a^b \frac{U(\gamma(\varphi))x}{\xi - \gamma(\varphi)} \gamma'(\varphi)\, d\varphi, f \right\rangle = \left\langle \frac{1}{2\pi i} \oint_\gamma \frac{U(\lambda)x}{\xi - \lambda}\, d\lambda, f \right\rangle \end{aligned}$$

for all $x \in E$ and $f \in F'$ where

$$\operatorname{Ind}_\gamma(\xi) := \frac{1}{2\pi i} \oint_\gamma \frac{d\lambda}{\xi - \lambda}.$$

Thus

$$U(\xi) \operatorname{Ind}_\gamma(\xi) = \frac{1}{2\pi i} \oint_\gamma \frac{U(\lambda)}{\xi - \lambda}\, d\lambda.$$

Therefore the Cauchy integral formula is valid for holomorphic operator valued functions.

The following theorem is the summary of this section.

Theorem C.2. *Let $U : \mathbb{C} \to B(E, F)$ be analytic in an open set $D \subset \mathbb{C}$ and γ a closed curve in D. Then*

$$\oint_\gamma U(\lambda)\, d\lambda = 0.$$

If $\xi \notin \gamma$,

$$U(\xi) \operatorname{Ind}_\gamma(\xi) = \frac{1}{2\pi i} \oint_\gamma \frac{U(\lambda)}{\xi - \lambda}\, d\lambda.$$

Appendix D

Special Operators

In this chapter we present some special bounded linear operators in Banach and Hilbert spaces. We consider the spaces of nuclear and Hilbert-Schmidt operators. References of this chapter are the books of Da Prato and Zabczyk [35], Kuo [23], Pietsch [32] and Treves [52]. Nuclear operators are also treaded among others in the book of Köthe [22] and Hilbert-Schmidt operators in the books of Dunford and Schwartz [10] and Köthe [22].

D.1 Hilbert-Schmidt Operators

Let $(U, (\cdot,\cdot)_U)$ and $(H, (\cdot,\cdot)_H)$ be separable Hilbert spaces.

Lemma D.1. *Let $\{e_k\}_{k=1}^\infty$ and $\{d_k\}_{k=1}^\infty$ be two orthonormal bases in U. Then*

$$\sum_{k=1}^\infty \|Te_k\|_H^2 = \sum_{k=1}^\infty \|Td_k\|_H^2$$

for a linear operator T from U to H.

Proof. Let $\{f_k\}_{k=1}^\infty$ be an orthonormal basis in H. Then for a linear operator T

$$\begin{aligned}\sum_{k=1}^\infty \|Te_k\|_H^2 &= \sum_{k=1}^\infty\sum_{j=1}^\infty |(Te_k, f_j)_H|^2 = \sum_{j=1}^\infty\sum_{k=1}^\infty |(e_k, T^* f_j)_U|^2 = \sum_{j=1}^\infty \|T^* f_j\|_U^2 \\ &= \sum_{j=1}^\infty\sum_{k=1}^\infty |(T^* f_j, d_k)_U|^2 = \sum_{k=1}^\infty\sum_{j=1}^\infty |(f_j, Td_k)_H|^2 = \sum_{k=1}^\infty \|Td_k\|_H^2.\end{aligned}$$

Thus if the series $\sum_{k=1}^\infty \|Te_k\|_H^2$ converges for some $\{e_k\}_{k=1}^\infty$, it converges for any other $\{d_k\}_{k=1}^\infty$ and if the series $\sum_{k=1}^\infty \|Te_k\|_H^2$ is divergent for some $\{e_k\}_{k=1}^\infty$, it is for any other $\{d_k\}_{k=1}^\infty$. □

Definition D.2. *A linear operator $T : U \to H$ is said to be a Hilbert-Schmidt operator if*

$$\sum_{k=1}^\infty \|Te_k\|_H^2 < \infty$$

for an orthonormal basis $\{e_k\}_{k=1}^\infty$ in U.

By Lemma D.1 the definition of Hilbert-Schmidt operators is independent of the choice of the basis $\{e_k\}_{k=1}^{\infty}$. We denote by $B(U,H)$ the Banach space of all bounded linear operators from U into H endowed with the operator norm

$$\|T\|_{B(U,H)} := \sup\{\|Tx\|_H : x \in U,\ \|x\|_U \leq 1\}$$

for all $T \in B(U,H)$ and by $B_2(U,H)$ the collection of Hilbert-Schmidt operators from U to H. We define the *Hilbert-Schmidt norm* by

$$\|T\|_{B_2(U,H)} := \left(\sum_{k=1}^{\infty} \|Te_k\|_H^2\right)^{\frac{1}{2}}$$

for all $T \in B_2(U,H)$. If $U = H$, we use the notation $B(H) := B(H,H)$ and $B_2(H) := B_2(H,H)$.

Theorem D.3. *Let $(U,(\cdot,\cdot)_U)$, $(H,(\cdot,\cdot)_H)$ and $(E,(\cdot,\cdot)_E)$ be separable Hilbert spaces and $Q \in B(E,U)$, $R \in B(H,E)$ and $S,T \in B_2(U,H)$. Then*

(i) $\|\alpha T\|_{B_2(U,H)} = |\alpha|\|T\|_{B_2(U,H)}$ *for all* $\alpha \in \mathbb{C}$,

(ii) $\|S+T\|_{B_2(U,H)} \leq \|S\|_{B_2(U,H)} + \|T\|_{B_2(U,H)}$,

(iii) $\|T^*\|_{B_2(H,U)} = \|T\|_{B_2(U,H)}$,

(iv) $\|T\|_{B(U,H)} \leq \|T\|_{B_2(U,H)}$,

(v) RT is a Hilbert-Schmidt operator from U to E and

$$\|RT\|_{B_2(U,E)} \leq \|R\|_{B(H,E)}\|T\|_{B_2(U,H)},$$

(vi) TQ is a Hilbert-Schmidt operator from E to H and

$$\|TQ\|_{B_2(E,H)} \leq \|Q\|_{B(E,U)}\|T\|_{B_2(U,H)}.$$

Proof. The statement (i) is obvious.

(ii) Let $S,T \in B_2(U,H)$ and $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis in U. Then by the Minkowski inequality,

$$\left(\sum_{k=1}^{\infty} \|(S+T)e_k\|_H^2\right)^{\frac{1}{2}} \leq \left(\sum_{k=1}^{\infty} \|Se_k\|_H^2\right)^{\frac{1}{2}} + \left(\sum_{k=1}^{\infty} \|Te_k\|_H^2\right)^{\frac{1}{2}}.$$

Thus $\|S+T\|_{B_2(U,H)} \leq \|S\|_{B_2(U,H)} + \|T\|_{B_2(U,H)}$.

The statement (iii) is a consequence of the proof of Lemma D.1.

(iv) Let T be a Hilbert-Schmidt operator from U to H and $x \in U$. Let $\{f_k\}_{k=1}^{\infty}$ be an orthonormal basis in H. Then

$$\begin{aligned}\|Tx\|_H^2 &= \sum_{k=1}^{\infty} |(Tx,f_k)|^2 = \sum_{k=1}^{\infty} |(x,T^*f_k)|^2 \\ &\leq \|x\|_U^2 \sum_{k=1}^{\infty} \|T^*f_k\|_U^2 = \|x\|_U^2 \|T^*\|_{B_2(H,U)}^2 \\ &= \|T\|_{B_2(U,H)}^2 \|x\|_U^2.\end{aligned}$$

Thus $\|T\|_{B(U,H)} \leq \|T\|_{B_2(U,H)}$.

(v) Let $R \in B(H,E)$ and $T \in B_2(U,H)$. Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal basis in U. Then

$$\|RT\|_{B_2(U,E)}^2 = \sum_{k=1}^{\infty} \|RTe_k\|_E^2 \leq \|R\|_{B(H,E)}^2 \sum_{k=1}^{\infty} \|Te_k\|_H^2 = \|R\|_{B(H,E)}^2 \|T\|_{B_2(U,H)}^2.$$

Hence RT is a Hilbert-Schmidt operator from U to E and the claimed inequality is valid.

(vi) Let $Q \in B(E,U)$ and $T \in B_2(U,H)$. Then by the statements (iii) and (v),

$$\|TQ\|_{B_2(E,H)} = \|Q^*T^*\|_{B_2(H,E)} \leq \|Q^*\|_{B(U,E)} \|T^*\|_{B_2(H,U)} = \|Q\|_{B(E,U)} \|T\|_{B_2(U,H)}$$

and hence TQ is a Hilbert-Schmidt operator from E to H. □

Corollary D.4. *The Hilbert-Schmidt norm is a norm in $B_2(U,H)$.*

By Theorem D.3 Hilbert-Schmidt operators are bounded, i.e., $B_2(U,H) \subseteq B(U,H)$. If U is finite dimensional, $B_2(U,H) = B(U,H)$. But if U is infinite dimensional, $B_2(U,H) \subset B(U,H)$, e.g. the identity operator is bounded but not a Hilbert-Schmidt operator.

Proposition D.5. *A Hilbert-Schmidt operator from U to H is compact.*

Proof. Let T be a Hilbert-Schmidt operator from U to H and $\{e_k\}_{k=1}^{\infty}$ an orthonormal basis in U. Then $\sum_{k=1}^{\infty} \|Te_k\|_H^2 < \infty$. Let $\{f_k\}_{k=1}^{\infty}$ be an orthonormal basis in H. Then $Tx = \sum_{k=1}^{\infty} (Tx, f_k)_H f_k$ for all $x \in U$. Thus

$$\begin{aligned} \left\| Tx - \sum_{k=1}^{n} (Tx, f_k)_H f_k \right\|_H^2 &= \sum_{k=n+1}^{\infty} |(Tx, f_k)_H|^2 = \sum_{k=n+1}^{\infty} |(x, T^* f_k)_U|^2 \\ &\leq \|x\|_U^2 \sum_{k=n+1}^{\infty} \|T^* f_k\|_U^2 \longrightarrow 0 \end{aligned}$$

as $n \to \infty$ for all $x \in U$ since $\sum_{j=1}^{\infty} \|T^* f_j\|_U^2 = \sum_{k=1}^{\infty} \|Te_k\|_H^2$ by the proof of Lemma D.1. Hence T is the limit of finite rank operators in the operator norm. Therefore T is compact. □

We equip the norm space $B_2(U,H)$ with the *Hilbert-Schmidt inner product*

$$(S,T)_{B_2(U,H)} := \sum_{k=1}^{\infty} (Se_k, Te_k)_H \tag{D.1}$$

for all $S, T \in B_2(U,H)$ where $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis in U. Then $B_2(U,H)$ is a Hilbert space.

Proposition D.6. *The space of Hilbert-Schmidt operators from U into H is a separable Hilbert space.*

Proof. Let $\{e_k\}_{k=1}^\infty$ and $\{f_k\}_{k=1}^\infty$ be orthonormal bases in U and H, respectively. The series on the right hand side of (D.1) converges absolutely since $2|(Se_k, Te_k)_H| \le \|Se_k\|_H^2 + \|Te_k\|_H^2$ for all $S, T \in B_2(U,H)$. The Hilbert-Schmidt inner product is independent of the basis $\{e_k\}_{k=1}^\infty$ because $\sum_{k=1}^\infty (Se_k, Te_k)_H = \sum_{k=1}^\infty (T^* f_k, S^* f_k)_H$ for all $S, T \in B_2(U,H)$. Since $(T,T)_{B_2(U,H)} = \|T\|_{B_2(U,H)}^2$ and $(\cdot,\cdot)_H$ is an inner product in H, the Hilbert-Schmidt inner product is an inner product in $B_2(U,H)$. Hence $B_2(U,H)$ is an inner product space.

To prove the completeness of $B_2(U,H)$ let us assume that $\{T_n\}_{n=1}^\infty$ is a Cauchy sequence in $B_2(U,H)$. By Theorem D.3 the sequence $\{T_n\}_{n=1}^\infty$ is also a Cauchy sequence in $B(U,H)$. Since $B(U,H)$ is a Banach space, there exists $T \in B(U,H)$ such that $\|T - T_n\|_{B(U,H)} \to 0$ as $n \to \infty$. We need to prove that $T \in B_2(U,H)$ and $\|T - T_n\|_{B_2(U,H)} \to 0$ as $n \to \infty$. Since $\{T_n\}_{n=1}^\infty$ is a Cauchy sequence in $B_2(U,H)$, it is bounded, i.e., there exists $C > 0$ such that $\|T_n\|_{B_2(U,H)} \le C$ for all $n \in \mathbb{N}$. Let $\{e_k\}_{k=1}^\infty$ be an orthonormal basis in U. Then for each $N \in \mathbb{N}$

$$\sum_{k=1}^N \|Te_k\|_H^2 = \lim_{n\to\infty} \sum_{k=1}^N \|T_n e_k\|_H^2 \le \lim_{n\to\infty} \|T_n\|_{B_2(U,H)}^2 \le C^2.$$

Hence $\|T\|_{B_2(U,H)} \le C$ and $T \in B_2(U,H)$. Let $\varepsilon > 0$. Then there exists $M > 0$ such that $\|T_m - T_n\|_{B_2(U,H)} < \varepsilon$ for all $m, n \ge M$. Then for $m \ge M$ and each $N \in \mathbb{N}$

$$\begin{aligned}\sum_{k=1}^N \|(T - T_m)e_k\|_H^2 &= \lim_{n\to\infty} \sum_{k=1}^N \|(T_n - T_m)e_k\|_H^2 \le \lim_{n\to\infty} \|T_n - T_m\|_{B_2(U,H)}^2 \\ &\le \limsup_{n\to\infty} \|T_n - T_m\|_{B_2(U,H)}^2 \le \varepsilon^2.\end{aligned}$$

Hence $\|T - T_m\|_{B_2(U,H)} \le \varepsilon$ for all $m \ge M$. Therefore $B_2(U,H)$ is complete.

Let $T \in B_2(U,H)$ and $\{f_k\}_{k=1}^\infty$ be an orthonormal basis in H. Then for all $x \in U$

$$\begin{aligned}Tx &= \sum_{k=1}^\infty (Tx, f_k)_H f_k = \sum_{k,l=1}^\infty (x, e_l)_U (Te_l, f_k)_H f_k \\ &= \sum_{k,l=1}^\infty (T, f_k \otimes e_l)_{B_2(U,H)} (f_k \otimes e_l)(x)\end{aligned}$$

where $(f_k \otimes e_l)(x) = (x, e_l)_U f_k$ for all $x \in U$. The set $\{f_k \otimes e_l\}_{k,l=1}^\infty$ is orthonormal in $B_2(U,H)$. Furthermore,

$$\begin{aligned}&\left\| T - \sum_{k=1}^n \sum_{l=1}^m (Te_l, f_k)_H (f_k \otimes e_l) \right\|_{B_2(U,H)}^2 \\ &= \sum_{j=1}^\infty \left\| \sum_{k=n+1}^\infty \sum_{l=m+1}^\infty (Te_l, f_k)_H (f_k \otimes e_l)(e_j) \right\|_H^2 \\ &= \sum_{j=m+1}^\infty \left\| \sum_{k=n+1}^\infty (Te_j, f_k)_H f_k \right\|_H^2 \\ &= \sum_{j=m+1}^\infty \sum_{k=n+1}^\infty |(Te_j, f_k)_H|^2 \longrightarrow 0\end{aligned}$$

as $m, n \to \infty$ since $\sum_{j,k=1}^{\infty} |(Te_j, f_k)_H|^2 = \sum_{j=1}^{\infty} \|Te_j\|_H^2 < \infty$. Hence $\{f_k \otimes e_l\}_{k,l=1}^{\infty}$ is an orthonormal basis in $B_2(U, H)$. $\square$

As an example of a Hilbert-Schmidt operator we present the Hilbert-Schmidt integral operator in $L^2(\mathbb{R})$.

Example D.7 (The Hilbert-Schmidt integral operator). *Let $k \in L^2(\mathbb{R}^2)$. We define the operator $K : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ by*

$$Kf(t) = \int_{-\infty}^{\infty} k(t, s) f(s) \, ds$$

for all $t \in \mathbb{R}$. Then K is a Hilbert-Schmidt operator and $\|K\|_{B_2(L^2(\mathbb{R}))} = \|k\|_{L^2(\mathbb{R}^2)}$.

Proof. Let $f \in L^2(\mathbb{R})$. Then

$$\begin{aligned}
\|Kf\|_{L^2(\mathbb{R})}^2 &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} k(t, s) f(s) \, ds \right|^2 \, dt \le \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |k(t, s) f(s)| \, ds \right)^2 \, dt \\
&\le \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |k(t, s)|^2 \, ds \int_{-\infty}^{\infty} |f(s)|^2 \, ds \, dt = \|k\|_{L^2(\mathbb{R}^2)}^2 \|f\|_{L^2(\mathbb{R})}^2.
\end{aligned}$$

Thus K is a bounded linear operator from $L^2(\mathbb{R})$ to itself and $\|K\|_{B(L^2(\mathbb{R}))} \le \|k\|_{L^2(\mathbb{R}^2)}$.

To show that K is actually a Hilbert-Schmidt operator we use the Fubini theorem. Since

$$\int_{\mathbb{R}^2} \left| \overline{k(t, s)} \right|^2 \, dsdt = \int_{\mathbb{R}^2} |k(t, s)|^2 \, dsdt < \infty,$$

by the Fubini theorem,

$$\int_{-\infty}^{\infty} \left| \overline{k(t, s)} \right|^2 \, ds < \infty$$

for almost all $t \in \mathbb{R}$, i.e., $\overline{k(t, \cdot)} \in L^2(\mathbb{R})$ for almost all $t \in \mathbb{R}$. Let $\{e_n\}_{n=1}^{\infty}$ be an orthonormal basis in $L^2(\mathbb{R})$. Then for almost all $t \in \mathbb{R}$

$$\left\| \overline{k(t, \cdot)} \right\|_{L^2(\mathbb{R})}^2 = \sum_{n=1}^{\infty} \left| \left(\overline{k(t, \cdot)}, e_n(\cdot) \right)_{L^2(\mathbb{R})} \right|^2 = \sum_{n=1}^{\infty} \left| \int_{-\infty}^{\infty} \overline{k(t, s) e_n(s)} \, ds \right|^2.$$

Hence by Lebesgue's monotone convergence theorem,

$$\begin{aligned}
\|k\|_{L^2(\mathbb{R}^2)}^2 = \left\| \bar{k} \right\|_{L^2(\mathbb{R}^2)}^2 &= \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \left| \int_{-\infty}^{\infty} \overline{k(t, s) e_n(s)} \, ds \right|^2 \, dt \\
&= \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} k(t, s) e_n(s) \, ds \right|^2 \, dt.
\end{aligned}$$

Therefore

$$\sum_{n=1}^{\infty} \|Ke_n\|_{L^2(\mathbb{R})}^2 = \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} k(t, s) e_n(s) \, ds \right|^2 \, dt = \|k\|_{L^2(\mathbb{R}^2)}^2.$$

Hence K is a Hilbert-Schmidt operator and $\|K\|_{B_2(L^2(\mathbb{R}))} = \|k\|_{L^2(\mathbb{R}^2)}$. $\square$

D.2 Nuclear Operators

Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be Banach spaces.

Definition D.8. *A bounded linear operator T is said to be nuclear if there exist sequences $\{a_j\}_{j=1}^\infty \subset F$ and $\{\varphi_j\}_{j=1}^\infty \subset E'$ such that T has the representation*

$$Tx = \sum_{j=1}^\infty a_j \langle x, \varphi_j \rangle$$

for all $x \in E$ and

$$\sum_{j=1}^\infty \|a_j\|_F \|\varphi_j\|_{E'} < \infty.$$

Proposition D.9. *A nuclear operator from E to F is compact.*

Proof. Let T be a nuclear operator from E to F. Then there exist sequences $\{a_j\}_{j=1}^\infty \subset F$ and $\{\varphi_j\}_{j=1}^\infty \subset E'$ such that T has for all $x \in E$ the representation $Tx = \sum_{j=1}^\infty a_j \langle x, \varphi_j \rangle$ with $\sum_{j=1}^\infty \|a_j\|_F \|\varphi_j\|_{E'} < \infty$. Then

$$\left\| Tx - \sum_{j=1}^n a_j \langle x, \varphi_j \rangle \right\|_F \leq \sum_{j=n+1}^\infty \|a_j\|_F |\langle x, \varphi_j \rangle| \leq \|x\|_E \sum_{j=n+1}^\infty \|a_j\|_F \|\varphi_j\|_{E'} \longrightarrow 0$$

as $n \to \infty$ for all $x \in E$. Hence T is the limit of finite rank operators in the operator norm. Therefore T is compact. $\square$

Let $B_1(E,F)$ be the collection of nuclear operators from E into F. We use the notation $B_1(E) := B_1(E,E)$. We endow $B_1(E,F)$ with the norm

$$\|T\|_{B_1(E,F)} := \inf \left\{ \sum_{j=1}^\infty \|a_j\|_F \|\varphi_j\|_{E'} : Tx = \sum_{j=1}^\infty a_j \langle x, \varphi_j \rangle \text{ for all } x \in E \right\}$$

for all $T \in B_1(E,F)$.

Theorem D.10. *Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ and $(G, \|\cdot\|_G)$ be Banach spaces and $Q \in B(G,E)$, $R \in B(F,G)$ and $S, T \in B_1(E,F)$. Then*

(i) $\|\alpha T\|_{B_1(E,F)} = |\alpha| \|T\|_{B_1(E,F)}$ *for all* $\alpha \in \mathbb{C}$,

(ii) $\|S + T\|_{B_1(E,F)} \leq \|S\|_{B_1(E,F)} + \|T\|_{B_1(E,F)}$,

(iii) $\|T\|_{B(E,F)} \leq \|T\|_{B_1(E,F)}$,

(iv) $\|T'\|_{B_1(F',E')} \leq \|T\|_{B_1(E,F)}$,

(v) *RT is a nuclear operator from E to G and*

$$\|RT\|_{B_1(E,G)} \leq \|R\|_{B(F,G)} \|T\|_{B_1(E,F)},$$

(vi) TQ is a nuclear operator from G to F and

$$\|TQ\|_{B_1(G,F)} \le \|Q\|_{B(G,E)}\|T\|_{B_1(E,F)}.$$

Proof. The statement (i) is obvious.

(ii) Let $S, T \in B_1(E,F)$ and $\varepsilon > 0$. Then there exist sequences $\{a_j\}_{j=1}^{\infty}, \{b_j\}_{j=1}^{\infty} \subset F$ and $\{\varphi_j\}_{j=1}^{\infty}, \{\phi_j\}_{j=1}^{\infty} \subset E'$ such that S and T have the representations $Sx = \sum_{j=1}^{\infty} a_j\langle x, \varphi_j\rangle$ and $Tx = \sum_{j=1}^{\infty} b_j\langle x, \phi_j\rangle$ for all $x \in E$ with $\sum_{j=1}^{\infty} \|a_j\|_F\|\varphi_j\|_{E'} < \|S\|_{B_1(E,F)} + \varepsilon/2$ and $\sum_{j=1}^{\infty} \|b_j\|_F\|\phi_j\|_{E'} < \|T\|_{B_1(E,F)} + \varepsilon/2$. We define the sequences $\{c_j\}_{j=1}^{\infty} \subset F$ and $\{\psi_j\}_{j=1}^{\infty} \subset E'$ by $c_{2j+1} := a_j$ and $c_{2j} := b_j$ and $\psi_{2j+1} := \varphi_j$ and $\psi_{2j} := \phi_j$ for all $j \in \mathbb{N}$. Then $(S+T)x = \sum_{j=1}^{\infty} c_j\langle x, \psi_j\rangle$ for all $x \in E$ and

$$\|S+T\|_{B_1(E,F)} \le \sum_{j=1}^{\infty} \|c_j\|_F\|\psi_j\|_{E'} < \|S\|_{B_1(E,F)} + \|T\|_{B_1(E,F)} + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $\|S+T\|_{B_1(E,F)} \le \|S\|_{B_1(E,F)} + \|T\|_{B_1(E,F)}$.

(iii) If T is a nuclear operator from E to F,

$$\|Tx\|_F \le \sum_{j=1}^{\infty} \|a_j\|_F|\langle x, \varphi_j\rangle| \le \|x\|_E \sum_{j=1}^{\infty} \|a_j\|_F\|\varphi_j\|_{E'}$$

for all representations $Tx = \sum_{j=1}^{\infty} a_j\langle x, \varphi_j\rangle$ and $x \in E$. By taking the infimum over all representations we get $\|T\|_{B(E,F)} \le \|T\|_{B_1(E,F)}$.

(iv) Let T be a nuclear operator from E to F and $\varepsilon > 0$. Then there exist sequences $\{a_j\}_{j=1}^{\infty} \subset F$ and $\{\varphi_j\}_{j=1}^{\infty} \subset E'$ such that T has the representation $Tx = \sum_{j=1}^{\infty} a_j\langle x, \varphi_j\rangle$ for all $x \in E$ and $\sum_{j=1}^{\infty} \|a_j\|_F\|\varphi_j\|_{E'} < \|T\|_{B_1(E,F)} + \varepsilon$. Hence for all $\phi \in F'$

$$\langle Tx, \phi\rangle = \left\langle \sum_{j=1}^{\infty} a_j\langle x, \varphi_j\rangle, \phi \right\rangle = \sum_{j=1}^{\infty}\langle x, \varphi_j\rangle\langle a_j, \phi\rangle = \left\langle x, \sum_{j=1}^{\infty}\langle a_j, \phi\rangle\varphi_j \right\rangle = \langle x, T'\phi\rangle.$$

Thus the Banach adjoint $T' \in B(F', E')$ has the representation $T'\phi = \sum_{j=1}^{\infty}\langle a_j, \phi\rangle\varphi_j$ for all $\phi \in F'$. Since $F \subset F''$ and $\|a_j\|_{F''} = \|a_j\|_F$ and hence $\sum_{j=1}^{\infty} \|a_j\|_{F''}\|\varphi_j\|_{E'} < \|T\|_{B_1(E,F)} + \varepsilon$, the Banach adjoint T' is nuclear and $\|T'\|_{B_1(F',E')} < \|T\|_{B_1(E,F)} + \varepsilon$. Since $\varepsilon > 0$ is arbitrary, $\|T'\|_{B_1(F',E')} \le \|T\|_{B_1(E,F)}$.

(v) Let $R \in B(F,G)$ and $T \in B_1(E,F)$. Then $RTx = \sum_{j=1}^{\infty} Ra_j\langle x, \varphi_j\rangle$ for all representation $Tx = \sum_{j=1}^{\infty} a_j\langle x, \varphi_j\rangle$ and $x \in E$. Thus

$$\|RT\|_{B_1(E,G)} \le \sum_{j=1}^{\infty} \|Ra_j\|_G\|\varphi_j\|_{E'} \le \|R\|_{B(F,G)} \sum_{j=1}^{\infty} \|a_j\|_F\|\varphi_j\|_{E'} < \infty.$$

Hence $RT \in B_1(E,G)$. By taking the infimum over all representations of T we get the claimed inequality.

(vi) Let $Q \in B(G,E)$ and $T \in B_1(E,F)$. Then

$$TQy = \sum_{j=1}^{\infty} a_j\langle Qy, \varphi_j\rangle = \sum_{j=1}^{\infty} a_j\langle y, Q'\varphi_j\rangle$$

for each representation $Tx = \sum_{j=1}^{\infty} a_j \langle x, \varphi_j \rangle$ and $y \in G$. Thus

$$\|TQ\|_{B_1(G,F)} \le \sum_{j=1}^{\infty} \|a_j\|_F \|Q'\varphi_j\|_{G'} \le \|Q'\|_{B(E',G')} \sum_{j=1}^{\infty} \|a_j\|_F \|\varphi_j\|_{E'} < \infty.$$

Hence $TQ \in B_1(G,F)$. By taking the infimum over all representations of T we get the claimed inequality since $\|Q'\|_{B(E',G')} = \|Q\|_{B(G,E)}$. □

As a corollary of Theorem D.10 $B_1(E,F)$ is a norm space with the norm $\|\cdot\|_{B_1(E,F)}$. Actually, $B_1(E,F)$ is complete.

Theorem D.11. *The space of nuclear operator from E to F is a Banach space.*

Proof. To prove the completeness let us assume that $\{T_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $B_1(E,F)$. By Theorem D.10 the sequence $\{T_n\}_{n=1}^{\infty}$ is also a Cauchy sequence in $B(E,F)$. Since $B(E,F)$ is a Banach space, there exists $T \in B(E,F)$ such that $\|T - T_n\|_{B(E,F)} \to 0$ as $n \to \infty$. We need to prove that $T \in B_1(E,F)$ and $\|T - T_n\|_{B_1(E,F)} \to 0$ as $n \to \infty$. We determine a monotonically increasing sequence $\{n_k\}_{k=1}^{\infty}$ of indices such that $\|T_l - T_m\|_{B_1(E,F)} < 1/2^{k+2}$ for all $l, m \ge n_k$. Then for all $k \in \mathbb{N}$ there exist sequences $\{a_j^k\}_{j=1}^{\infty} \subset F$ and $\{\varphi_j^k\}_{j=1}^{\infty} \subset E'$ such that the nuclear operator $T_{n_{k+1}} - T_{n_k}$ has the representation $(T_{n_{k+1}} - T_{n_k})x = \sum_{j=1}^{\infty} a_j^k \langle x, \varphi_j^k \rangle$ for all $x \in E$ and $\sum_{j=1}^{\infty} \|a_j^k\|_F \|\varphi_j^k\|_{E'} < 1/2^{k+2}$. Let $k \in \mathbb{N}$. Consequently, for all $p \in \mathbb{N}$

$$(T_{n_{k+p}} - T_{n_k})x = \sum_{l=k}^{k+p-1} (T_{n_{l+1}} - T_{n_l})x = \sum_{l=k}^{k+p-1} \sum_{j=1}^{\infty} a_j^l \langle x, \varphi_j^l \rangle$$

for all $x \in E$. By taking the limit $p \to \infty$ we obtain the identity $(T - T_{n_k})x = \sum_{l=k}^{\infty} \sum_{j=1}^{\infty} a_j^l \langle x, \varphi_j^l \rangle$ for all $x \in E$ because the series on the right hand side converges absolutely. Since

$$\|T - T_{n_k}\|_{B_1(E,F)} \le \sum_{l=k}^{\infty} \sum_{j=1}^{\infty} \|a_j^l\|_F \|\varphi_j^l\|_{E'} \le \frac{1}{2^{k+1}},$$

the operator $T - T_{n_k}$ is nuclear and hence is also T. Finally,

$$\|T - T_n\|_{B_1(E,F)} \le \|T - T_{n_k}\|_{B_1(E,F)} + \|T_{n_k} - T_n\|_{B_1(E,F)} < \frac{1}{2^k}$$

for all $n \ge n_k$ and hence $\|T - T_n\|_{B_1(E,F)} \to 0$ as $n \to \infty$. □

In separable Hilbert spaces the product of two Hilbert-Schmidt operators is nuclear.

Proposition D.12. *Let $(U, (\cdot,\cdot)_U)$, $(H, (\cdot,\cdot)_H)$ and $(E, (\cdot,\cdot)_E)$ be separable Hilbert spaces. If $T \in B_2(U,H)$ and $S \in B_2(H,E)$, then $ST \in B_1(U,E)$ and*

$$\|ST\|_{B_1(U,E)} \le \|S\|_{B_2(H,E)} \|T\|_{B_2(U,H)}.$$

Proof. Let $T \in B_2(U,H)$, $S \in B_2(H,E)$ and $\{f_j\}_{j=1}^{\infty}$ be an orthonormal basis in H. Then

$$STx = \sum_{j=1}^{\infty} (Tx, f_j)_H S f_j = \sum_{j=1}^{\infty} (x, T^* f_j)_U S f_j$$

for all $x \in U$. Thus

$$\|ST\|_{B_1(U,E)} \leq \sum_{j=1}^{\infty} \|Sf_j\|_E \|T^* f_j\|_U \leq \left(\sum_{j=1}^{\infty} \|Sf_j\|_E^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{\infty} \|T^* f_j\|_U^2 \right)^{\frac{1}{2}}$$
$$= \|S\|_{B_2(H,E)} \|T^*\|_{B_2(H,U)} = \|S\|_{B_2(H,E)} \|T\|_{B_2(U,H)}.$$

Therefore $ST \in B_1(U,E)$ and the claimed inequality is valid. □

D.2.1 Trace Class Operators

Let $(H, (\cdot,\cdot)_H)$ be a separable Hilbert space and $\{e_k\}_{k=1}^{\infty}$ an orthonormal basis in H. If $T \in B_1(H)$, we define the *trace* of T by

$$\operatorname{Tr} T := \sum_{j=1}^{\infty} (Te_j, e_j)_H.$$

Proposition D.13. *If $T \in B_1(H)$, then $\operatorname{Tr} T$ is a well defined number independent of the orthonormal basis $\{e_k\}_{k=1}^{\infty}$.*

Proof. Let T be a nuclear operator in H. Then there exist sequences $\{a_j\}_{j=1}^{\infty} \subset H$ and $\{\varphi_j\}_{j=1}^{\infty} \subset H'$ such that T has the representation $Th = \sum_{j=1}^{\infty} a_j \langle h, \varphi_j \rangle$ for all $h \in H$ and $\sum_{j=1}^{\infty} \|a_j\|_H \|\varphi_j\|_{H'} < \infty$. By the Riesz representation theorem for all $j \in \mathbb{N}$ there exists $b_j \in H$ such that $\langle h, \varphi_j \rangle = (h, b_j)_H$ for all $h \in H$ and $\|b_j\|_H = \|\varphi_j\|_{H'}$. Thus

$$\begin{aligned}
\sum_{k=1}^{\infty} |(Te_k, e_k)_H| &= \sum_{k=1}^{\infty} \left| \sum_{j=1}^{\infty} (e_k, b_j)_H (a_j, e_k)_H \right| \leq \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |(e_k, b_j)_H (a_j, e_k)_H| \\
&\leq \sum_{j=1}^{\infty} \left(\sum_{k=1}^{\infty} |(a_j, e_k)_H|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{\infty} |(e_k, b_j)_H|^2 \right)^{\frac{1}{2}} \\
&= \sum_{j=1}^{\infty} \|a_j\|_H \|b_j\|_H < \infty.
\end{aligned} \tag{D.2}$$

Hence the series $\sum_{k=1}^{\infty} (Te_k, e_k)$ converges absolutely and furthermore,

$$\sum_{k=1}^{\infty} (Te_k, e_k)_H = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (e_k, b_j)_H (a_j, e_k)_H = \sum_{j=1}^{\infty} (a_j, b_j)_H.$$

Thus the definition of $\operatorname{Tr} T$ is independent of $\{e_k\}_{k=1}^{\infty}$. □

According to Estimate (D.2),

$$|\operatorname{Tr} T| \leq \|T\|_{B_1(H)} \tag{D.3}$$

for all $T \in B_1(H)$.

Proposition D.14. *A non-negative self-adjoint operator $T \in B(H)$ is nuclear if and only if for an orthonormal basis $\{e_j\}_{j=1}^{\infty}$ in H*

$$\sum_{j=1}^{\infty} (Te_j, e_j)_H < \infty.$$

In addition, $\|T\|_{B_1(H)} = \operatorname{Tr} T$.

Proof. "$\Rightarrow$" If T is nuclear, then $\operatorname{Tr} T < \infty$ by Estimate (D.3).

"$\Leftarrow$" Let T be a non-negative self-adjoint operator such that $\sum_{j=1}^{\infty} (Te_j, e_j)_H < \infty$ for an orthonormal basis $\{e_j\}_{j=1}^{\infty}$ in H. First we show that T is compact. Let $T^{1/2}$ denote the non-negative self-adjoint square root of T [36, Theorem 13.31]. Then $T^{1/2}h = \sum_{j=1}^{\infty} \left(T^{1/2}h, e_j\right)_H e_j$ for all $h \in H$ and

$$\begin{aligned}
&\left\| T^{1/2}h - \sum_{j=1}^{n} \left(T^{1/2}h, e_j\right)_H e_j \right\|_H^2 \\
&= \sum_{j=n+1}^{\infty} \left| \left(T^{1/2}h, e_j\right)_H \right|^2 = \sum_{j=n+1}^{\infty} \left| \left(h, T^{1/2}e_j\right)_H \right|^2 \\
&\leq \|h\|_H^2 \sum_{j=n+1}^{\infty} \left\| T^{1/2}e_j \right\|_H^2 = \|h\|_H^2 \sum_{j=n+1}^{\infty} (Te_j, e_j)_H \longrightarrow 0
\end{aligned}$$

as $n \to \infty$ for all $h \in H$. Hence the operator $T^{1/2}$ is the limit of finite rank operators in the operator norm. Therefore $T^{1/2}$ is compact and $T = T^{1/2}T^{1/2}$ is a compact operator as well.

Let $\{f_k\}_{k=1}^{\infty}$ be the sequence of all normalized eigenvectors of T and $\{\lambda_k\}_{k=1}^{\infty}$ the corresponding sequence of eigenvalues. Then $Th = \sum_{k=1}^{\infty} \lambda_k (h, f_k)_H f_k$ for all $h \in H$ since T is a compact self-adjoint operator [14, Theorem 5.1, pp. 113–115]. Thus

$$\sum_{j=1}^{\infty} (Te_j, e_j)_H = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \lambda_k |(e_j, f_k)_H|^2 = \sum_{k=1}^{\infty} \lambda_k \|f_k\|_H^2 = \sum_{k=1}^{\infty} \lambda_k.$$

Hence

$$\sum_{k=1}^{\infty} \|\lambda_k f_k\|_H \|f_k\|_H = \sum_{k=1}^{\infty} \lambda_k < \infty$$

and therefore T is nuclear. Furthermore, $\operatorname{Tr} T = \sum_{k=1}^{\infty} \lambda_k$. Since $\|T\|_{B_1(H)} \leq \sum_{k=1}^{\infty} \lambda_k$ and $|\operatorname{Tr} T| \leq \|T\|_{B_1(H)}$, we have $\|T\|_{B_1(H)} = \operatorname{Tr} T$. □

Let $T \in B(H)$. Then T^*T is a positive self-adjoint operator in H. Thus there exists positive self-adjoint $R \in B(H)$ such that $R^2 = T^*T$ and $\|Rx\|_H = \|Tx\|_H$ for all $x \in H$ [36, Theorem 12.34]. We define the operator $U : \mathcal{R}(R) \to \mathcal{R}(T)$ by $Ux := Ty$ where $x = Ry$. Then $URx = Tx$ for all $x \in H$ since $\operatorname{Ker}(R) = \operatorname{Ker}(T)$. Thus

$$\|URx\|_H = \|Tx\|_H = \|Rx\|_H$$

for all $x \in H$. Hence U is an isometry from $\mathcal{R}(R)$ to $\mathcal{R}(T)$. Therefore U has a continuous extension to a linear isometry from the closure of $\mathcal{R}(R)$ to the closure of $\mathcal{R}(T)$. Additionally, we define $Ux = 0$ for all $x \in \mathcal{R}(R)^{\perp}$. Hence $U \in B(H)$ and $\|U\|_{B(H)} = 1$. The operators R and U are called the *polar decomposition* of T.

Theorem D.15. *A bounded linear operator $T : H \to H$ is nuclear if and only if*

$$\sum_{k=1}^{\infty} \lambda_k < \infty$$

where $\{\lambda_k\}_{k=1}^{\infty}$ are the eigenvalues of $(T^*T)^{1/2}$.

Proof. We denote $R := (T^*T)^{1/2}$.

"$\Leftarrow$" Let us assume that $\sum_{k=1}^{\infty} \lambda_k < \infty$ where $\{\lambda_k\}_{k=1}^{\infty}$ are the eigenvalues of R. Since R is a non-negative self-adjoint operator in H and $\operatorname{Tr} R = \sum_{k=1}^{\infty} \lambda_k < \infty$, by Proposition D.14 the operator R is nuclear and $\|R\|_{B_1(H)} = \operatorname{Tr} R$. There exists $U \in B(H)$ such that $URx = Tx$ for all $x \in H$ and $\|U\|_{B(H)} = 1$. Thus by Theorem D.10 the operator T is nuclear and

$$\|T\|_{B_1(H)} \leq \|U\|_{B(H)} \|R\|_{B_1(H)} = \operatorname{Tr} R = \sum_{k=1}^{\infty} \lambda_k.$$

"$\Rightarrow$" Let $T \in B_1(H)$. Since $T = UR$ and U is an isometry from the closure of $\mathcal{R}(R)$ to the closure of $\mathcal{R}(T)$, there exists the bounded linear inverse of U from the closure of $\mathcal{R}(T)$ to the closure of $\mathcal{R}(R)$. We define $Vx = U^{-1}x$ for all $x \in \overline{\mathcal{R}(T)}$. Then V is an isometry from the closure of $\mathcal{R}(T)$ to the closure of $\mathcal{R}(R)$. Additionally, we define $Vx = 0$ for all $x \in \mathcal{R}(T)^{\perp}$. Then $V \in B(H)$ and $\|V\|_{B(H)} = 1$. Furthermore, $VTx = Rx$ for all $x \in H$ since $\operatorname{Ker}(T) = \operatorname{Ker}(R)$. Thus by Theorem D.10 the operator R is nuclear and

$$\|R\|_{B_1(H)} \leq \|V\|_{B(H)} \|T\|_{B_1(H)} = \|T\|_{B_1(H)}.$$

Hence $\operatorname{Tr} R = \sum_{k=1}^{\infty} \lambda_k < \infty$ by Proposition D.14. $\square$

Corollary D.16. *Let $T \in B_1(H)$. Then*

$$\|T\|_{B_1(H)} = \operatorname{Tr}(T^*T)^{1/2} = \sum_{k=1}^{\infty} \lambda_k$$

where $\{\lambda_k\}_{k=1}^{\infty}$ are the eigenvalues of $(T^*T)^{1/2}$.

By Theorem D.15 and Corollary D.16 it is justified that the nuclear operators in H are also called *trace class* operators.

Bibliography

[1] S. Agmon, A. Doulig, and L. Nirenberg. Estimates Near the Boundary for Solutions of Elliptic Partial Differential Equations Satisfying General Boundary Conditions, I. *Communications on Pure and Applied Mathematics*, XII:623–727, 1959.

[2] Herbert Amann. Existence and Regularity for Semilinear Parabolic Evolution Equations. *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze. Serie IV*, XI(4):593–676, 1984.

[3] Patrick Billingsley. *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, U.S.A., 1979. ISBN 0-471-03173-9.

[4] Liliana Borcea. Electrical impedance tomography. *Inverse Problems*, 18(6): R99–R136, 2002.

[5] Martin Brühl. Explicit characterization of inclusions in electrical impedance tomography. *SIAM Journal on Mathematical Analysis*, 32(6):1327–1341, 2001.

[6] Martin Brühl and Martin Hanke. Numerical implementation of two noniterative methods for locating inclusions by impedance tomography. *Inverse Problems*, 16(4):1029–1042, 2000.

[7] Raymond Chang. *Physical Chemistry with Applications to Biological Systems*. Collier-Macmillan, New York, U.S.A., 1981. ISBN 0023210400.

[8] E. B. Davies. *One-parameter Semigroups*. Number 15 in London Mathematical Society Monographs. Academic Press, London, Great Britain, 1980. ISBN 0-12-206280-9.

[9] J. Diestel and J. J. Uhl. *Vector Measures*. Number 15 in Mathematical Surveys. American Mathematical Society, U.S.A., 1977. ISBN 0-8218-1515-6.

[10] Nelson Dunford and Jacob T. Schwartz. *Linear Operators Part II*, volume VII of *Pure and Applied Mathematics*. Interscience Publishers Inc., New York, U.S.A., 1964.

[11] Gustav Elfving and Pekka Tuominen. *Todennäköisyyslaskenta II*. Limes ry, 1990. ISBN 951-745-130-X.

[12] D.L. George, J.R. Torczynski, K.A. Shollenberger, T.J. O'Hern, and S.L. Ceccio. Validation of electrical-impedance tomography for measurements of material distribution in two-phase flows. *International Journal of Multiphase Flow*, 26(4):549–581, 2000.

[13] Giuseppe Geymonat and Pierre Grisvard. Alucuni risultati di teoria spettrale per i problemi al limiti lineari ellittici. *Rendiconti del Seminario Matematico della Università di Padova*, 38:121–173, 1967.

[14] Israel Gohberg and Seymour Goldberg. *Basic Operator Theory*. Birkhäuser, Boston, U.S.A., 1980. ISBN 3-7643-3028-7.

[15] Jerome A. Goldstein. *Semigroups of Linear Operators and Applications*. Oxford Mathematical Monographs. Oxford University Press, New York, U.S.A., 1985. ISBN 0-19-503540-2.

[16] Einar Hille and Ralph S. Phillips. *Functional Analysis and Semigroups*, volume XXXI of *Americal Mathematical Society Colloquium Publications*. American Mathematical Society, U.S.A., 1957.

[17] Tuomas Hytönen. R-boundedness and multiplier theorems. Master's thesis, Helsinki University of Technology, Espoo, Finland, 2001.

[18] Jari P. Kaipio, Ville Kolehmainen, Erkki Somersalo, and Marko Vauhkonen. Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography. *Inverse Problems*, 16(5):1487–1522, 2000.

[19] Jari P. Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, Germany, 2004. ISBN 0-387-22073-9.

[20] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Number 113 in Graduate Texts in Mathematics. Springer, New York, U.S.A., 1991. ISBN 0-387-97655-8.

[21] Bernt Øksendal. *Stochastic Differential Equations — An Introduction with Applications*. Springer-Verlag, fifth edition, 1998. ISBN 3-540-63720-6.

[22] Gottfried Köthe. *Topological Vector Spaces II*, volume 237 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, U.S.A., 1979. ISBN 0-387-90440-9.

[23] Hui-Hsiung Kuo. *Gaussian Measures in Banach Spaces*, volume 463 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, Germany, 1975. ISBN 3-540-07173-3.

[24] Kenneth L. Kuttler. *Modern Analysis*. Studies in Advanced Mathematics. CRC Press, 1998. ISBN 0-8493-7166-X.

[25] Sari Lasanen. *Discretizations of Generalized Random Variables with Applications to Inverse Problems*. PhD thesis, University of Oulu, Oulu, Finland, 2002.

[26] J. L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications I*. Number 181 in Die Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, Germany, 1972. ISBN 0-387-05363-8.

[27] R. S. Liptser and A. N. Shiryayev. *Statistics of Random Processes I: General Theory*, volume 5 of *Applications of Mathematics*. Springer-Verlag, New York, U.S.A., 1977. ISBN 0-387-90226-0.

[28] Alessandra Lunardi. *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, volume 16 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Basel, Switzerland, 1995. ISBN 3-7643-5172-1.

[29] Kalle Mikkola. *Infinite-dimensional Linear Systems, Optimal Control and Algebraic Riccatti Equations*. PhD thesis, Helsinki University of Technology, Espoo, Finland, 2002.

[30] Jacques Neveu. *Mathematical Foundations of the Calculus of Probability*. Holden-Day, Inc., San Francisco, U.S.A., 1965.

[31] Amnon Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, U.S.A., 1983. ISBN 0-387-90845-5.

[32] Albrecht Pietsch. *Nuclear Locally Convex Spaces*, volume 66 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, Berlin, Germany, 1972. ISBN 3-540-05644-0.

[33] Hanna Katriina Pikkarainen. The mathematical formulation of the state evolution equation in electric process tomography. In *Proceedings of 3rd World Congress on Industrial Process Tomography*, pages 660–663. The Virtual Centre for Industrial Process Tomography, 2003. ISBN 0-853-16240-9.

[34] Hanna Katriina Pikkarainen. Stochastic approach to electric process tomography. In *Proceedings of Inverse Problems, Desing and Optimization Symposium*, 2004.

[35] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, United Kingdom, 1992. ISBN 0-521-38529-6.

[36] Walter Rudin. *Functional Analysis*. Tata McGraw-Hill Publishing Company Ltd., New Delhi, India, TMH edition, 1974. ISBN 0-07-099558-3.

[37] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill Book Company, Singapore, third edition, international edition, 1987. ISBN 0-07-054234-1.

[38] A.R. Ruuskanen, A. Seppänen, S. Duncan, E. Somersalo, and J.P. Kaipio. Optimal control in process tomography. In *Proceedings of 3rd World Congress on Industrial Process Tomography*, pages 245–251. The Virtual Centre for Industrial Process Tomography, 2003. ISBN 0-853-16240-9.

[39] A. Seppänen, L. Heikkinen, T. Savolainen, E. Somersalo, and J.P. Kaipio. An experimental evaluation of state estimation with fluid dynamical models in process tomography. In *Proceedings of 3rd World Congress on Industrial Process Tomography*, pages 541–546. The Virtual Centre for Industrial Process Tomography, 2003. ISBN 0-853-16240-9.

[40] A. Seppänen, M. Vauhkonen, E. Somersalo, and J.P. Kaipio. Effects of inaccuracies in fluid dynamical models in state estimation of process tomography. In *Process Imaging for Automatic Control*, volume 4188 of *Proceedings of SPIE*, pages 69–80, 2001.

[41] A. Seppänen, M. Vauhkonen, E. Somersalo, and J.P. Kaipio. State space models in process tomography – approximation of state noise covariance. *Inverse Problems in Engineering*, 9:561–585, 2001.

[42] A. Seppänen, M. Vauhkonen, E. Somersalo, and J.P. Kaipio. Inference of velocity fields based on tomographic measurements in process industry. In *4th International Conference on Inverse Problems in Engineering*, 2002.

[43] A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. Fluid dynamical models and state estimation in process tomography: Effect due to inaccuracies in flow fields. *Journal of Electronic Imaging*, 10(3):630–640, 2001.

[44] A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. State estimation in three dimensional impedance imaging – use of fluid dynamical evolution models. In *Proceedings of 2nd World Congress on Industrial Process Tomography*, pages 198–207. The Virtual Centre for Industrial Process Tomography, 2001.

[45] A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. State estimation with fluid dynamical evolution models in process tomography – an application with impedance tomography. *Inverse Problems*, 17:467–484, 2001.

[46] Samuli Siltanen, Jennifer Mueller, and David Isaacson. An implementation of the reconstruction algorithm of A. Nachman for the 2D inverse conductivity problem. *Inverse Problems*, 16(3):681–699, 2000.

[47] E. Somersalo, M. Cheney, D. Isaacson, and E. Isaacson. Layer stripping: a direct numerical method for impedance imaging. *Inverse Problems*, 7(6):899–926, 1991.

[48] Erkki Somersalo, Margaret Cheney, and David Isaacson. Existence and uniqueness for electrode models for electric current computed tomography. *SIAM Journal on Applied Mathematics*, 52(4):1023–1040, 1992.

[49] John Sylvester. A convergent layer stripping algorithm for the radially symmetric impedance tomography problem. *Communications in Partial Differential Equations*, 17(11–12):1955–1994, 1992.

[50] Hiroki Tanabe. *Equations of Evolution.* Pitman, Great Britain, 1979. ISBN 0-273-01137-5.

[51] Hiroki Tanabe. *Functional Analytic Methods for Partial Differential Equations.* Number 204 in Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, New York, U.S.A., 1997. ISBN 0-8247-9774-4.

[52] Francois Treves. *Topological Vector Spaces, Distributions and Kernels*, volume 25 of *Pure and Applied Mathematics.* Academic Press Inc., New York, U.S.A., 1967.

[53] David Williams. *Probability with Martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, Great Britain, 1991. ISBN 0-521-40605-6.

[54] Kôsaku Yosida. *Functional Analysis.* Springer-Verlag, Berlin, Germany, 1965.

(continued from the back cover)

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The list of reports is continued inside. Electronical versions of the reports are available at *http://www.math.hut.fi/reports/* .

A485 Sampsa Pursiainen
Bayesian approach to detection of anomalies in electrical impedance tomography
April 2005

A484 Visa Latvala , Niko Marola , Mikko Pere
Harnack's inequality for a nonlinear eigenvalue problem on metric spaces
March 2005

A482 Mikko Lyly , Jarkko Niiranen , Rolf Stenberg
A refined error analysis of MITC plate elements
April 2005

A481 Dario Gasbarra , Tommi Sottinen , Esko Valkeila
Gaussia Bridges
December 2004

A480 Ville Havu , Jarmo Malinen
Approximation of the Laplace transform by the Cayley transform
December 2004

ISBN 951-22-7651-8
ISSN 0784-3143