



Bachelorarbeit

Die Karhunen-Loève-Zerlegung

Beispiele und Optimalität bezüglich des mittleren
quadratischen Fehlers

eingereicht von

Tim Jaschek

zum Erlangen des akademischen Grades

Bachelor of Science

am

21. Juni 2016

Betreuer: Prof. Dr. Dirk Becherer

Zweitgutachter: Prof. Dr. Markus Reiß

geboren am: 28. Dezember 1994

geboren in: Berlin

Institut für Mathematik
Humboldt Universität zu Berlin

Vorwort

Die Karhunen-Loève-Zerlegung bietet sowohl aus der reinen mathematischen Perspektive als auch im Hinblick auf das Anwendungsspektrum reichlich Potential. Grob vereinfacht ist die Karhunen-Loève-Zerlegung ein Werkzeug, um stochastische Prozesse mit gewissen Eigenschaften auf eine sehr aussagekräftige Weise darzustellen, welche zeitliche und stochastische Abhängigkeiten isoliert. Die Darstellung ist vergleichbar mit einer Fourierreihe. Während eine Fourierreihe aus skalaren Werten und einer orthonormalen Basis aus trigonometrischen Funktionen besteht, sind die Koeffizienten bei der Karhunen-Loève-Entwicklung unkorrelierte Zufallsvariablen und die Basis aus orthonormalen Funktionen ist von dem spezifischen Prozess abhängig. Im Gegensatz zur Fourierreihe besitzt die Karhunen-Loève Entwicklung nicht nur eine, sondern gleich zwei Orthogonalitäten und wird daher als bi-orthogonales System bezeichnet.

Aus der Karhunen-Loève-Zerlegung eines stochastischen Prozesses können zügig gewisse Eigenschaften, wie beispielsweise die totale Varianz des Prozesses, gefolgert werden, die vorher verborgen waren. Das Isolieren der zeitlichen und stochastischen Abhängigkeiten kann genutzt werden, um ein deterministisches Grundverhalten in dem Prozess zu finden. Außerdem zeigt sich, dass die Partialsummen der Karhunen-Loève-Entwicklung unter allen vergleichbaren Reihendarstellungen eines Prozesses den mittleren quadratischen Fehler minimieren. Dies macht sie in Anwendungen und zur Simulation von stochastischen Prozessen besonders nützlich.

Die ersten Arbeiten zu dem Thema entstanden nahezu zeitgleich und unabhängig im Jahre 1947 von dem finnischen Mathematiker Kari Karhunen (1915-1992) in seiner Diplomarbeit „Über lineare Methoden in der Wahrscheinlichkeitsrechnung“[Kar47] und dem israelischen Mathematiker Michael Loève (1907-1979) in „Fonctions aleatoires du second ordre“[Loè48]. Karhunen studierte an der Universität Helsinki und wurde dort auch zum Professor berufen. Loève studierte in Paris bei Paul Lévy und war später Professor an der Universität Berkeley. Zuvor hatte sich der indische Statistiker und Historiker D. D. Kosambi mit dem Thema beschäftigt, weshalb die Thematik in der Literatur auch teilweise als Kosambi-Karhunen-Loève-Zerlegung bezeichnet wird.

Als aktuelle Arbeiten, die eine gute Einführung zur Thematik bieten, beziehungsweise auch weiter darüber hinaus gehen, seien hier Alexanderian [Ale15] und Wang [Wan08] angeführt. Außerhalb der rein theoretischen Ebene reicht das Anwendungsspektrum der Karhunen-

Loève-Zerlegung über viele verschiedene Gebiete. Eine große Rolle spielt hier beispielsweise die Frequenzerkennung aus stark gestörten Signalen. Mit der Karhunen-Loève-Transformation (KLT) können signifikant bessere Ergebnisse erzielt werden, als mit herkömmlichen Methoden wie der schnellen Fourier-Transformation (FFT). Dies zeigen Arbeiten von Szumski [Szu11] von der European Space Agency und von Maccone [Mac08] vom SETI Institut. Die deutlich erhöhte Genauigkeit der Signalerkennung mittels KLT kann in der Raumfahrt- und Satellitenkommunikation von großem Nutzen sein, steht allerdings vor dem Problem eines sehr hohen Rechenaufwandes. Maccone und andere versuchen daher einen Algorithmus zur schnellen Karhunen-Loève-Transformation zu entwickeln. Des Weiteren wird die KLT in einer vereinfachten Form häufig in Bild- und allgemeiner Datenkompression verwendet, siehe dazu beispielsweise Dony [DH95]. Auch zum Entfernen von sogenanntem Bildrauschen, welches bei Bildern aus dem All durch zufällige Störungen entsteht, können bemerkenswerte Ergebnisse erzielt werden. In verschiedenen Artikeln werden vielversprechende Ansätze zur Mustererkennung aus Bild und Filmmaterial, zum maschinellen Lernen, zur Erdbebensimulation, zur Berechnung von maximalen Wellenamplituden und im Finanzbereich diskutiert.

Um die Karhunen-Loève-Zerlegung vollständig zu verstehen, benötigt man ein mathematisches Fundament aus Funktionalanalysis und Stochastik. Die vorliegende Arbeit kann leider unmöglich auf alle Themen im Zusammenhang mit der Karhunen-Loève-Zerlegung eingehen und muss sich auf einige Themenfelder beschränken. Im ersten Kapitel werden funktional-analytische Grundlagen wie das Spektraltheorem für kompakte, selbstadjungierte Operatoren und Mercers Theorem vorgestellt. Außerdem wird die Klasse der Hilbert-Schmidt-Operatoren im Hinblick auf diese Theoreme analysiert. Der wahrscheinlichkeitstheoretische Hintergrund wird im zweiten Kapitel diskutiert. Es stellt sich heraus, dass die Kovarianzfunktion eines stochastischen Prozesses wichtige Informationen über den Prozess liefert und unter bestimmten Voraussetzungen einen Hilbert-Schmidt-Operator induziert. Das dritte Kapitel ist ebenfalls in die Funktionalanalysis einzuordnen. Dort wird das Lebesgue-Integral auf das Bochner-Integral erweitert, welches für das Karhunen-Loève-Theorem eine große Bedeutung hat. Diesem Theorem ist daraufhin das gesamte vierte Kapitel gewidmet. Nach einem ausführlichen Beweis folgen Beispiele, um die Aussagen zu verdeutlichen. Ein weiteres Kapitel beschäftigt sich mit der Optimalität der Zerlegung.

Zu dieser Arbeit wurden außerdem einige Programme in MATLAB entwickelt. Neben einem Algorithmus zur Gewinnung der Zerlegung wurde auch ein Algorithmus zur Signalerkennung und zur Bildkompression mittels KLT implementiert. Außerdem wurden verschiedene Plots erstellt, um für ein besseres Verständnis der Theoreme zu sorgen. Sämtliche Abbildungen in dieser Arbeit sind selbst erstellt und gesondert am Ende der Arbeit zu finden. Die Implementation der Karhunen-Loève-Zerlegung wird im letzten Abschnitt veranschaulicht und diskutiert.

Tim Jaschek, Berlin 21. Juni 2016

Inhaltsverzeichnis

Vorwort	iii
1 Funktionalanalytischer Hintergrund	1
1.1 Spektraltheorie kompakter selbstadjungierter Operatoren	1
1.1.1 Das Spektrum eines beschränkten Operators	1
1.1.2 Normale und selbstadjungierte Operatoren	2
1.1.3 Kompakte Operatoren	3
1.1.4 Das Spektraltheorem	4
1.2 Hilbert-Schmidt-Operatoren und Mercers Theorem	5
2 Stochastische Prozesse zweiter Ordnung	9
2.1 Die Kovarianzfunktion	9
2.2 Konvergenz und Stetigkeit im quadratischen Mittelwert	12
2.3 Hilberträume mit reproduzierendem Kern	14
3 Das Bochner-Integral	17
3.1 Aufbau der allgemeinen Theorie	17
3.2 Anwendung im Kontext dieser Arbeit	22
4 Das Karhunen-Loève-Theorem	25
4.1 Aussage und Beweis des Theorems	25
4.2 Karhunen-Loève-Zerlegung von Gauß-Prozessen	29
4.3 Vorteile der Karhunen-Loève-Zerlegung	31
4.4 Karhunen-Loève-Entwicklung von Zufallsfeldern	32
4.5 Beispiele zur Berechnung der Zerlegung	34
4.5.1 Die Brown'sche Bewegung und die Brown'sche Brücke	34
4.5.2 Das Brown'sche Blatt	35
5 Optimalität der Karhunen-Loève-Zerlegung	37
5.1 Veranschaulichung und empirische Bestätigung	38
5.2 Analytischer Beweis der Optimalität bezüglich des mittleren quadratischen Fehlers	39

6 Implementation der Karhunen-Loève-Transformation	41
6.1 Numerisches Lösen der Fredholm'schen Integralgleichung	41
6.1.1 Quadraturformeln	41
6.1.2 Fourier-Methode	43
6.1.3 Haar-Wavelet-Methode	44
6.2 Anwendungen der Karhunen-Loève-Transformation	45
6.2.1 Signalermittlung	45
6.2.2 Entfernen von Bildrauschen und Bildkompression	46
Schlusswort	47
Abbildungen	49
Literaturverzeichnis	55
Notationsverzeichnis	57
Stichwortverzeichnis	59

Kapitel 1

Funktionalanalytischer Hintergrund

1.1 Spektraltheorie kompakter selbstadjungierter Operatoren

Die Spektraltheorie ist eine Erweiterung der, aus der linearen Algebra bekannten, Theorie von Eigenwerten und Eigenvektoren von Endomorphismen eines endlich-dimensionalen Vektorraumes. Einen guten Überblick zur endlich-dimensionalen Theorie findet man beispielsweise in dem Buch von K. Jänich [Jän00, Kapitel 9 und 10].

In \mathbb{C}^d hat man festgestellt, dass man zu einer hermiteschen Matrix $A \in \mathbb{C}^{d \times d}$ stets reelle Eigenwerte $\lambda_1, \dots, \lambda_d$ und eine orthonormale Basis $\{e_i : i = 1, \dots, d\}$ des \mathbb{C}^d findet, wobei e_i Eigenvektor von A zum Eigenwert λ_i ist, für $i = 1, \dots, d$. Des Weiteren lässt sich die folgende Zerlegung der linearen Abbildung herleiten:

$$Ax = \sum_{i=1}^d \lambda_i \langle x, e_i \rangle e_i. \quad (1.1)$$

Diese Zerlegung von A wird auch als *Spektralzerlegung* bezeichnet. Nun soll eine vergleichbare Darstellung in einem allgemeinen Hilbertraum $(X, \langle \cdot, \cdot \rangle)$, für den Fall $\dim(X) = \infty$ gefunden werden. Die Theorie hierzu wird ausführlich in dem Buch von D. Werner [Wer00, Kapitel VI] beschrieben. In dieser Arbeit soll lediglich ein Überblick zu den wichtigsten Sätzen und Definitionen gegeben werden. Die Arbeit von C. Uebel [Ueb16] liefert eine vertiefende Darstellung zu diesem Thema im selben Kontext wie diese Arbeit.

1.1.1 Das Spektrum eines beschränkten Operators

Sei fortan X ein Hilbertraum und $\mathcal{L}(X, X)$ der Banachraum der linearen beschränkten Operatoren von X in sich selbst mit der Norm

$$\|A\|_{\mathcal{L}(X, X)} := \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|\leq 1} \|Ax\| = \sup_{x \neq x'} \frac{\|Ax - Ax'\|_X}{\|x - x'\|_X}. \quad (1.2)$$

Die zentralen Elemente der Spektraltheorie sind Eigenwerte und Eigenvektoren von Operatoren. In endlich-dimensionalen Räumen sind Eigenwerte dadurch definiert, dass $\lambda - \text{Id}_X$ nicht injektiv ist, was dazu äquivalent ist, dass diese Abbildung nicht surjektiv ist. Im unendlich-dimensionalen Fall gilt diese Äquivalenz nicht mehr und die Definition muss angepasst werden.

Definition 1.1 (Def. VI.1.1 [Wer00]). Sei $A \in \mathcal{L}(X, X)$ und $\lambda - A$ eine abkürzende Schreibweise für $\lambda \text{Id}_X - A$.

- i) Die *Resolventenmenge* von A ist $\rho(A) = \{\lambda \in \mathbb{C} : (\lambda - A)^{-1} \text{ existiert in } \mathcal{L}(X, X)\}$.
- ii) Das *Spektrum* von A ist $\sigma(A) = \mathbb{C} \setminus \rho(A)$.

Jedes $\lambda \in \sigma(A)$ ist entweder im *Punktspektrum* $\sigma_p(A)$, im *stetigen Spektrum* $\sigma_c(A)$ oder im *Residuenspektrum* $\sigma_r(A)$, wobei

$$\begin{aligned}\sigma_p(A) &= \{\lambda \in \mathbb{C} : \ker(\lambda - A) \neq \{0\}\}, \\ \sigma_c(A) &= \{\lambda \in \mathbb{C} : \ker(\lambda - A) = \{0\}, \operatorname{im}(\lambda - A) \neq X, \overline{\operatorname{im}(\lambda - A)} = X\}, \\ \sigma_r(A) &= \{\lambda \in \mathbb{C} : \ker(\lambda - A) = \{0\}, \overline{\operatorname{im}(\lambda - A)} \neq X\}.\end{aligned}$$

Dabei gilt $\sigma(A) = \sigma_p(A) \sqcup \sigma_c(A) \sqcup \sigma_r(A)$. Die Elemente von $\sigma_p(A)$ heißen *Eigenwerte* von A und ein $x \in X \setminus \{0\}$ heißt *Eigenvektor* zum Eigenwert λ , falls $Ax = \lambda x$. Ist X ein Funktionenraum, so spricht man auch von einer *Eigenfunktion*.

Bemerkung 1.2. Laut [Wer00, S. 256] besteht das Punktspektrum typischerweise aus isolierten Punkten, das stetige Spektrum aus einer Vereinigung von Intervallen und das Residuenspektrum ist häufig leer. Dies erklärt die Namensgebung.

1.1.2 Normale und selbstadjungierte Operatoren

Neben Eigenwerten und Eigenvektoren müssen auch die Begriffe der Selbstadjungiertheit und der Hermitizität von Matrizen auf endlich-dimensionalen Vektorräumen zu Operatoren auf allgemeinen Hilberträumen sinnvoll erweitert werden.

Definition und Lemma 1.3 (Def. 8.2 [Per16]). Sei $(X, \langle \cdot, \cdot \rangle)$ ein Hilbertraum und $A \in \mathcal{L}(X, X)$. Dann existiert genau ein $A^* \in \mathcal{L}(X, X)$, sodass für alle $x, y \in X$ gilt:

$$\langle Ax, y \rangle = \langle x, A^*y \rangle \tag{1.3}$$

A^* wird als adjungierter Operator zu A bezeichnet.

Beweis. Der Beweis wurde [Per16] nachempfunden. Für $y \in X$ gilt:

$$|\langle Ax, y \rangle| \stackrel{\text{CSU}}{\leq} \|Ax\| \|y\| \leq \|A\|_{\mathcal{L}(X, X)} \|x\| \|y\|.$$

Daher ist $[x \mapsto \langle Ax, y \rangle] \in X^*$, wobei X^* den Dualraum von X bezeichne. Nach dem Repräsentationstheorem von Riesz für die Dualräume von Hilberträumen existiert dann ein $z \in X$, sodass $\langle Ax, y \rangle = \langle x, z \rangle$ für jedes $x \in X$. Setze $A^*y := z$. Dieser Operator ist linear, da

$$\begin{aligned}\langle Ax, \lambda y + \hat{y} \rangle &= \bar{\lambda} \langle Ax, y \rangle + \langle Ax, \hat{y} \rangle \\ &= \bar{\lambda} \langle x, A^*y \rangle + \langle x, A^*\hat{y} \rangle \\ &= \langle x, \lambda A^*y + A^*\hat{y} \rangle\end{aligned}$$

und beschränkt, da

$$\begin{aligned}\|A^*y\| &= \sup_{\|x\| \leq 1} |\langle x, A^*y \rangle| = \sup_{\|x\| \leq 1} |\langle Ax, y \rangle| \\ &\leq \sup_{\|x\| \leq 1} \|A\| \|x\| \|y\| = \|A\| \|y\|.\end{aligned}$$

Somit ist $A^* \in \mathcal{L}(X, X)$ mit $\langle Ax, y \rangle = \langle x, A^*y \rangle$ für alle $x, y \in X$. \square

Die gezeigte Existenz des adjungierten Operators wird für folgende Definitionen benötigt:

Definition 1.4 (Def. V.5.3 [Wer00]). Sei X ein Hilbertraum und $A \in \mathcal{L}(X, X)$.

- i) A heißt *selbstadjungiert* oder *hermitesch*, falls $A = A^*$.
- ii) A heißt *normal*, falls $AA^* = A^*A$.

Das folgende Lemma fasst wichtige Folgerungen für Eigenwerte und Eigenvektoren selbstadjungierter oder normaler Operatoren zusammen.

Satz 1.5 (Lem. VI.3.1 [Wer00]). *Sei $A \in \mathcal{L}(X, X)$.*

- (a) $\lambda \in \sigma(A)$ genau dann, wenn $\bar{\lambda} \in \sigma(A^*)$.
- (b) Ist A selbstadjungiert und $\lambda \in \sigma_p(A)$, so ist $\lambda \in \mathbb{R}$.
- (c) Ist A normal und x Eigenvektor von A zum Eigenwert λ , so ist x auch Eigenvektor von A^* zum Eigenwert $\bar{\lambda}$.
- (d) Ist A normal, so haben verschiedene Eigenwerte orthogonale Eigenvektoren.

Beweis. Für einen Beweis wird auf Lemma VI.3.1 in [Wer00] verwiesen. \square

1.1.3 Kompakte Operatoren

Um die Existenz von Eigenwerten von A zu erhalten, braucht man im Allgemeinen stärkere Voraussetzungen an den Operator als bisher. Dazu werden nun kompakte Operatoren eingeführt.

Definition 1.6 (Def. II.3.1 [Wer00]). Seien X, Y metrische Räume, $M \subseteq X$ heißt *relativ kompakt in X* , falls der Abschluss von M in X kompakt ist. Ein linearer Operator $T \in \mathcal{L}(X, Y)$ heißt *kompakt*, falls $T(B_X[0, 1])$ relativ kompakt in Y ist, wobei $B_X[0, 1]$ der abgeschlossene Einheitsball in X sei. $\mathcal{K}(X)$ bezeichnet den Raum der kompakten Operatoren von X in sich selbst.

Beispiel 1.7. Falls X endlich-dimensional ist, so ist jede lineare Abbildung $T : X \rightarrow Y$ kompakt. In der Tat: Da X endlich-dimensional ist, ist der Einheitsball $B_X[0, 1]$ kompakt. Die Linearität von T impliziert Stetigkeit und stetige Abbildungen bilden kompakte Mengen stets auf kompakte Mengen ab.

Ist hingegen $X = Y$ und X nicht endlich-dimensional, so ist $B_X[0, 1]$ nicht kompakt. Damit ist klar, dass der Identitätsoperator in diesem Fall nicht kompakt sein kann.

Ein allgemeineres Beispiel für eine ganze Klasse von kompakten Operatoren wird im Abschnitt 1.2 vorgestellt. Es handelt sich dabei um sogenannte Hilbert-Schmidt-Integraloperatoren. Diese Klasse von Operatoren hat eine große Bedeutung für die Karhunen-Loëve-Zerlegung.

Folgendes Lemma liefert ein Kriterium für Kompaktheit, welches im weiteren Verlauf der Arbeit benötigt wird.

Lemma 1.8 (Kor. II.3.3 [Wer00]). *Seien X und Y Banachräume und sei $T \in \mathcal{L}(X, Y)$. Falls eine Folge (T_n) stetiger linearer Operatoren mit endlich-dimensionalem Bild und $\|T_n - T\| \rightarrow 0$ existiert, ist T kompakt.*

Beweis. Für einen Beweis wird auf Kor. II.3.3 in [Wer00] verwiesen. \square

1.1.4 Das Spektraltheorem

Nun kann das anfangs motivierte Spektraltheorem für kompakte normale beziehungsweise kompakte selbstadjungierte Operatoren formuliert werden. Die angegebene Form entspricht Theorem VI.3.2 in [Wer00].

Satz 1.9 (Spektraltheorem). *Sei $T \in \mathcal{K}(X)$ normal (falls $\mathbb{K} = \mathbb{C}$) beziehungsweise selbstadjungiert (falls $\mathbb{K} = \mathbb{R}$). Dann existieren ein (eventuell endliches) Orthonormalsystem e_1, e_2, \dots sowie eine (eventuell abbrechende) Nullfolge $(\lambda_i)_{i \in \mathbb{N}}$ in $\mathbb{K} \setminus \{0\}$, sodass*

$$X = \ker T \oplus \overline{\text{lin}\{e_1, e_2, \dots\}} \quad (1.4)$$

sowie

$$Tx = \sum_{i=1}^{\infty} \lambda_i \langle x, e_i \rangle e_i \quad \forall x \in X; \quad (1.5)$$

und zwar sind die λ_i die von Null verschiedenen Eigenwerte und e_i ist ein Eigenvektor zu λ_i .

Beweis. Einen detaillierten Beweis findet man in der oben angegebenen Quelle. In dieser Arbeit wird darauf wegen anderer Schwerpunktsetzung verzichtet. \square

1.2 Hilbert-Schmidt-Operatoren und Mercers Theorem

Hilbert-Schmidt-Operatoren sind eine Klasse von Integraloperatoren. Sie sollen hier zunächst definiert und dann mit Hilfe des Spektraltheorems analysiert werden. Wie bereits erwähnt, dienen sie hier nicht nur als Beispiel, sondern haben eine große Bedeutung für die Karhunen-Loève-Zerlegung. Diese Operatoren werden auch in [Wer00, Bsp. II.1(m),II.3(c)] besprochen. Im weiteren Verlauf der Arbeit sei $\mathcal{T} \subseteq \mathbb{R}$ kompakt und versehen mit der Borel- σ -Algebra $\mathcal{B}(\mathcal{T})$. Außerdem sei $\lambda = \lambda_{|\mathcal{T}}$ das Lebesgue-Maß auf \mathcal{T} und wie gewohnt $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ der Lebesgue-Raum der messbaren und quadratintegrierbaren Abbildungen $f : \mathcal{T} \rightarrow \mathbb{C}$.

Definition 1.10. Man bezeichnet eine messbare Funktion $K : \mathcal{T} \times \mathcal{T} \rightarrow \mathbb{C}$ als einen *Hilbert-Schmidt-Kern*, falls

$$\begin{aligned} \int_{\mathcal{T}} \int_{\mathcal{T}} |K(s, t)|^2 \lambda(dt) \lambda(ds) &< \infty, & (\text{Integrierbarkeit}) \\ K(s, t) &= \overline{K(t, s)}. & (\text{Hermitizität}) \end{aligned}$$

Die Messbarkeit und Integrierbarkeit zusammen entsprechen $K \in L^2(\mathcal{T} \times \mathcal{T}, \mathcal{B}(\mathcal{T})^{\otimes 2}, \lambda^{\otimes 2})$.

Bemerkung 1.11. In der Literatur wird bei der Definition von Hilbert-Schmidt-Kernen teilweise auf die Bedingung der Hermitizität verzichtet.

Ist K stetig und hermitesch, so ist K ein Hilbert-Schmidt-Operator: Die Kompaktheit von \mathcal{T} impliziert auch Kompaktheit von $\mathcal{T} \times \mathcal{T}$ in der Produkttopologie. Daher wird $|K|$ auf $\mathcal{T} \times \mathcal{T}$ ein Maximum M annehmen und es gilt

$$\int_{\mathcal{T}} \int_{\mathcal{T}} |K(s, t)|^2 \lambda(dt) \lambda(ds) \leq \lambda(\mathcal{T})^2 M^2 < \infty.$$

Im weiteren Verlauf wird oftmals anstelle von $\lambda(dt)$ lediglich dt geschrieben.

Definition 1.12 (vgl. Bsp. II.1(m) [Wer00]). Ein Hilbert-Schmidt-Kern K induziert auf kanonische Weise einen Integraloperator

$$\begin{aligned} T_K : \quad L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda) &\longrightarrow L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda) \\ f &\longmapsto \left[s \mapsto \int_{\mathcal{T}} K(s, t) f(t) dt \right]. \end{aligned}$$

Der Integraloperator T_K heißt *Hilbert-Schmidt-Operator* zum Kern K .

An dieser Stelle muss man sich jedoch die Frage der Wohldefiniertheit stellen. Es ist nicht trivial, dass der Ausdruck auf der rechten Seite integrierbar ist. Der Satz von Fubini liefert aber dank der $(\lambda \otimes \lambda)$ -Integrierbarkeit von K , dass $K(s, \cdot) \in L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ für fast alle $s \in \mathcal{T}$. Da auch $f \in L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$, folgt mit der Hölder'schen Ungleichung

$$\int_{\mathcal{T}} |K(s, t)f(t)| dt \leq \left(\int_{\mathcal{T}} |K(s, t)|^2 dt \right)^{\frac{1}{2}} \left(\int_{\mathcal{T}} |f(t)|^2 dt \right)^{\frac{1}{2}} < \infty,$$

weshalb die Integrale für λ -fast alle $s \in \mathcal{T}$ existieren. Auf der fehlenden Nullmenge setze $T_K f = 0$. Die Abbildungen $T_K f$ sind nach dieser Überlegung messbar für jedes $f \in L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$. Dass T_K sogar nach $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ abbildet, wird sich im Beweis zu Lemma 1.13 zeigen.

Lemma 1.13 (vgl. Bsp. II.1(m) [Wer00]). *Sei T_K ein Hilbert-Schmidt-Operator mit Kern K . Dann gilt*

$$T_K \in \mathcal{L}(L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)).$$

Beweis. Die Linearität von T_K folgt aus der Linearität des Lebesgue-Integrals, denn

$$\begin{aligned} T_K(\alpha f + g) &= \int_{\mathcal{T}} K(s, t)(\alpha f(t) + g(t)) dt \\ &= \alpha \int_{\mathcal{T}} K(s, t)f(t) dt + \int_{\mathcal{T}} K(s, t)g(t) dt = \alpha T_K f + T_K g. \end{aligned}$$

Bezüglich der Beschränktheit stellt man fest:

$$\begin{aligned} \|T_K f\|_{L^2}^2 &= \int_{\mathcal{T}} \left| \int_{\mathcal{T}} K(s, t)f(t) dt \right|^2 ds \\ &\leq \int_{\mathcal{T}} \left(\int_{\mathcal{T}} |K(s, t)| |f(t)| dt \right)^2 ds \\ &\leq \int_{\mathcal{T}} \left(\int_{\mathcal{T}} |K(s, t)|^2 dt \right) \left(\int_{\mathcal{T}} |f(t)|^2 dt \right) ds \\ &= \underbrace{\int_{\mathcal{T}} \int_{\mathcal{T}} |K(s, t)|^2 ds dt}_{<\infty} \|f\|_{L^2}^2. \end{aligned}$$

Die zweite Ungleichung erhält man mit der Hölder'schen Ungleichung. Damit ist gezeigt, dass T_K nach $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ abbildet und beschränkt ist, also $T_K \in \mathcal{L}(L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda))$. \square

Aus dem Beweis folgt, dass stets die Ungleichung $\|T_K\| \leq \|K\|$ gilt. Außerdem folgt analog zur Linearität, dass für zwei solche Operatoren T_K, T_G mit Kernen K und G gilt $T_K \pm T_G = T_{K \pm G}$.

Lemma 1.14. *Sei T_K ein Hilbert-Schmidt-Operator mit Kern K . Dann ist T_K selbstadjungiert.*

Beweis. Lemma und Beweis sind selbst aufgestellt. Es ist gilt für alle $f, g \in L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$:

$$\begin{aligned} \langle T_K f, g \rangle &= \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, t)f(t) dt \overline{g(s)} ds \\ &= \int_{\mathcal{T}} f(t) \int_{\mathcal{T}} \overline{K(t, s)} g(s) ds dt \\ &= \int_{\mathcal{T}} f(t) \overline{\int_{\mathcal{T}} K(t, s) g(s) ds} dt = \langle f, T_K g \rangle. \end{aligned}$$

Der Integraloperator T_K ist folglich selbstadjungiert. \square

Lemma 1.15 (vgl. Bsp. II.3(c) [Wer00]). *Sei T_K ein Hilbert-Schmidt-Operator mit Kern K . Dann gilt*

$$T_K \in \mathcal{K}(L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)).$$

Beweis. Die Linearität und die Beschränktheit wurden bereits gezeigt. Einen schönen Beweis zur fehlenden Eigenschaft der Kompaktheit findet man in [Wer00, S.68]. Da K messbar ist, kann K durch Treppenfunktionen mit Stufen aus messbaren Rechtecken approximiert werden. Diese Treppenfunktionen haben die Form

$$K_n(s, t) := \sum_{i,j=1}^{N^{(n)}} \alpha_{ij}^{(n)} \mathbb{1}_{E_i^{(n)}}(s) \mathbb{1}_{F_j^{(n)}}(t), \quad \text{für } \alpha_{ij}^{(n)} \in \mathbb{C} \text{ und } E_i^{(n)}, F_j^{(n)} \in \mathcal{B}(\mathcal{T}),$$

sodass $\|K_n - K\|_{L^2(\mathcal{T} \times \mathcal{T}, \mathcal{B}(\mathcal{T})^{\otimes 2}, \lambda^{\otimes 2})} \rightarrow 0$. Dann folgt

$$\|T_{K_n} - T_K\| = \|T_{K_n - K}\| \leq \|K_n - K\| \rightarrow 0.$$

Außerdem hat jedes T_{K_n} die Gestalt

$$[T_{K_n} f](s) = \int_{\mathcal{T}} K_n(s, t) f(t) dt = \sum_{i=1}^{N^{(n)}} \underbrace{\left(\sum_{j=1}^{N^{(n)}} \alpha_{ij}^{(n)} \int_{F_j^{(n)}} f(t) dt \right)}_{\in \mathbb{C}} \mathbb{1}_{E_i^{(n)}}(s).$$

Daher gilt

$$T_{K_n} f \in \text{lin}\{\mathbb{1}_{E_1^{(n)}}, \dots, \mathbb{1}_{E_{N^{(n)}}^{(n)}}\} \quad \forall f \in L^2(\mathcal{T}).$$

Die T_{K_n} haben demnach stets endlich-dimensionale Bilder und sind als Hilbert-Schmidt-Operatoren beschränkt und damit stetig. Mit Lemma 1.8 folgt, dass auch deren Grenzwert T_K kompakt ist. \square

Zusammenfassend: Hilbert-Schmidt-Integraloperatoren sind selbstadjungiert und kompakt. Es kann daher das Spektraltheorem 1.9 angewendet werden. Dieses liefert folgende Erkenntnisse:

- 1) Jeder Hilbert-Schmidt-Integraloperator T_K besitzt abzählbar viele Eigenfunktionen. Diese bilden ein Orthonormalsystem $(\phi_i)_{i \in \mathbb{N}}$ von $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ und falls $\ker(T_K) = \{0\}$ ist dieses System vollständig, also eine Basis.
- 2) Sei $(\lambda_i)_{i \in \mathbb{N}}$ die zugehörige Nullfolge der Eigenwerte. Dann gilt

$$T_K f = \sum_{i=1}^{\infty} \lambda_i \langle f, \phi_i \rangle \phi_i = \sum_{i=1}^{\infty} \lambda_i \phi_i \int_{\mathcal{T}} f(t) \overline{\phi_i(t)} dt \quad \text{für alle } f \in L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda). \quad (1.6)$$

$$\text{Das heißt } \lim_{n \rightarrow \infty} \int_{\mathcal{T}} \left| [T_K f](t) - \sum_{i=1}^n \lambda_i \phi_i(t) \langle f, \phi_i \rangle \right|^2 dt \rightarrow 0 \quad \text{für jedes } f. \quad (1.7)$$

Neben der Darstellung für den Hilbert-Schmidt-Operator T_K aus Gleichung 1.6 lässt sich unter einer weiteren Voraussetzung auch für den zugehörigen Kern K eine Reihendarstellung entwickeln. Dazu beachte man die folgende Definition:

Definition 1.16. Sei T_K ein Hilbert-Schmidt-Operator mit Kern K . Der Integraloperator T_K heißt *positiv*, falls für alle $f \in L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ gilt

$$\langle T_K f, f \rangle = \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, t) f(t) \overline{f(s)} ds dt \geq 0. \quad (1.8)$$

Mit dieser Bedingung an T_K hat der britische Mathematiker James Mercer im Jahre 1909 die folgende Reihendarstellung für den Kern K bewiesen:

Satz 1.17 (Mercer's Theorem). *Sei $K : \mathcal{T} \times \mathcal{T} \rightarrow \mathbb{C}$ stetig mit positivem induzierten Hilbert-Schmidt-Operator T_K . Seien weiter $(\lambda_i)_{i \in \mathbb{N}}$ und $(\phi_i)_{i \in \mathbb{N}}$ die Eigenwerte und Eigenfunktionen von T_K . Dann gilt für alle $s, t \in D$*

$$K(s, t) = \sum_{i=1}^{\infty} \lambda_i \phi_i(s) \overline{\phi_i(t)}, \quad (1.9)$$

wobei die obige Reihe absolut und gleichmäßig konvergiert.

Beweis. Für einen Beweis siehe [Wer00, Satz VI.4.2] oder [GGK03, Sec. 5, Thm. 3.1]. □

Die gleichmäßige Konvergenz ist in dem Sinne zu verstehen, dass

$$\lim_{N \rightarrow \infty} \sup_{s, t \in \mathcal{T}} \left| \sum_{i=1}^N \lambda_i \phi_i(s) \overline{\phi_i(t)} - K(s, t) \right| = 0, \quad (1.10)$$

während absolute Konvergenz bedeutet, dass $\lim_{N \rightarrow \infty} \sum_{i=1}^N |\lambda_i| |\phi_i(s)|^2 < \infty$. Bemerkenswert ist, dass die Stetigkeit von K auch Stetigkeit der Eigenfunktionen $(\phi_i)_{i \in \mathbb{N}}$ impliziert.

Um Mercers Theorem anzuwenden, benötigt man die Eigenwerte und Eigenfunktionen von T_K . Dazu hat man die Gleichung $T_K \phi = \lambda \phi$ zu lösen. Substitution der Definition des Integraloperators liefert $\int_{\mathcal{T}} K(s, t) \phi(t) dt = \lambda \phi(s)$. Diese Gleichung wird auch als *Fredholm'sche Integralgleichung 2. Art* bezeichnet und tritt häufig in der Funktionalanalysis auf. Sie steht in engem Zusammenhang zu dem Karhunen-Loève-Theorem.

In Abbildung 1 wurden verschiedene Hilbert-Schmidt-Kerne mittels der Partialsummen der Reihe aus Mercers Theorem approximiert. Schon bei wenigen Summanden ist die Konvergenz erkennbar. Die für die Approximation benötigten Eigenwerte und Eigenfunktionen wurden numerisch approximiert beziehungsweise analytisch bestimmt. Sie sind in Abbildung 2 dargestellt. Wie man die Fredholm'sche Integralgleichung analytisch lösen kann, wird später in Beispiel 4.10 gezeigt. Methoden zum numerischen Lösen werden im Kapitel 6 diskutiert. In dem nächsten Kapitel wird eine große Klasse von Hilbert-Schmidt-Kernen vorgestellt: Die Kovarianzfunktionen.

Kapitel 2

Stochastische Prozesse zweiter Ordnung

Sei $(\Omega, \mathcal{F}, \mathbf{P})$ ein Wahrscheinlichkeitsraum und $L^2(\Omega, \mathcal{F}, \mathbf{P})$ der Raum der Äquivalenzklassen komplexwertiger Zufallsvariablen zweiter Ordnung auf (Ω, \mathcal{F}) bezüglich der Äquivalenzrelation $X \sim Y : \Leftrightarrow X - Y = 0$ fast sicher. Bekanntlich ist $L^2(\Omega, \mathcal{F}, \mathbf{P})$ ein Hilbertraum mit dem inneren Produkt $\langle \cdot, \cdot \rangle_{L^2(\Omega, \mathcal{F}, \mathbf{P})}$ und der davon induzierten Norm $\|\cdot\|_{L^2(\Omega, \mathcal{F}, \mathbf{P})}$ definiert durch

$$\langle X, Y \rangle_{L^2(\Omega, \mathcal{F}, \mathbf{P})} = \mathbf{E} [X \bar{Y}], \quad \|X\|_{L^2(\Omega, \mathcal{F}, \mathbf{P})} = \sqrt{\langle X, X \rangle_{L^2(\Omega, \mathcal{F}, \mathbf{P})}} = (\mathbf{E} [|X|^2])^{\frac{1}{2}}.$$

Falls es in einem gegebenen Zusammenhang klar ist, welche Norm oder welches Skalarprodukt gemeint ist, wird im Folgenden auf den Index verzichtet. Ebenso wird teilweise auf die Klammern des Erwartungswertoperators verzichtet, um die Übersichtlichkeit zu bewahren. In diesem Kapitel werden Familien von L^2 -Zufallsvariablen, sogenannte stochastische Prozesse zweiter Ordnung, diskutiert. Bereits Kolmogorov hat 1941 Folgen von Zufallsvariablen zweiter Ordnung mit Hilbertraummethoden untersucht. Michel Loève stellt in [Loè91] die Theorie von stetigen Prozessen zweiter Ordnung in einer allgemeineren Form dar, die er zum Teil selbst in den Jahren 1945-46 aufgestellt hat. Auch auf [Ale15] wird in diesem Abschnitt verwiesen.

2.1 Die Kovarianzfunktion

Sei wie zuvor $\mathcal{T} \subseteq \mathbb{R}$ kompakt und zusätzlich $\{X_t : t \in \mathcal{T}\}$ ein zentrierter stochastischer Prozess. Dieser Prozess sei *zweiter Ordnung*, das heißt für alle $t \in \mathcal{T}$ gelte $\|X_t\|_{L^2(\Omega, \mathcal{F}, \mathbf{P})} < \infty$. Diese Bedingung impliziert zusammen mit der Hölder'schen Ungleichung für $p = q = \frac{1}{2}$, dass für alle $s, t \in \mathcal{T}$ die Kovarianzen $\text{Cov}(X_s, X_t)$ definiert sind. Daher ist es möglich, zu diesem Prozess eine *Kovarianzfunktion* K zu betrachten, die wiederum definiert ist durch

$$K : \quad \mathcal{T} \times \mathcal{T} \longrightarrow \mathbb{R}$$
$$(s, t) \longmapsto \text{Cov}(X_s, X_t) = \mathbf{E}(X_s - \mathbf{E}X_s)(\overline{X_t - \mathbf{E}X_t}).$$

Gelegentlich wird die Kovarianzfunktion eines Prozesses auch als *Kovarianzkern* bezeichnet. Da der zugrundeliegende Prozess hier im Folgenden stets zentriert sein wird, vereinfacht sich der Ausdruck für die Kovarianz zu $\mathbf{E}X_s\overline{X}_t$.

Beispiel 2.1. Sei $W = \{W_t : t \in [0, 1]\}$ eine Standard-Brown'sche Bewegung. Dann hat W die Kovarianzfunktion K_W mit $K_W(s, t) = s \wedge t$.

Proposition 2.2. *Sei $\{X_t : t \in \mathcal{T}\}$ ein Prozess zweiter Ordnung und K dessen Kovarianzfunktion. Dann erfüllt K die folgenden Eigenschaften:*

- i) $K(s, t) = \overline{K(t, s)}$ und $K(t, t) \geq 0$.
- ii) $K(s, t)$ ist positiv semidefinit auf $\mathcal{T} \times \mathcal{T}$.
- iii) $|K(s, t)| \leq \sqrt{K(s, s)}\sqrt{K(t, t)}$.

Beweis. Die erste Eigenschaft folgt unmittelbar aus der Definition der Kovarianz. Zur zweiten Eigenschaft: Seien $a_1, \dots, a_N \in \mathbb{C}$ und $t_1, \dots, t_N \in \mathcal{T}$ beliebig. Definiert man hilfsweise $Y := \sum_{i=1}^N a_i X_{t_i}$, so sieht man schnell, dass K positiv semidefinit ist, denn

$$\sum_{i=1}^N \sum_{j=1}^N a_i K(t_i, t_j) \overline{a_j} = \mathbf{E}Y\overline{Y} = \mathbf{E}|Y|^2 \geq 0. \quad (2.1)$$

Eigenschaft iii) folgt schließlich mit der Cauchy-Schwarz-Ungleichung. \square

Ist eine Kovarianzfunktion integrierbar, so ist sie ein Hilbert-Schmidt-Kern. Daher ist es naheliegend, die Klasse der stetigen Kovarianzfunktionen näher zu untersuchen, denn diese erfüllen die Bedingung der Integrierbarkeit auf einem Kompaktum.

Angenommen, es ist nun K eine stetige Kovarianzfunktion. Wie im ersten Kapitel gezeigt, ist dann der induzierte Hilbert-Schmidt-Integraloperator T_K kompakt und selbstadjungiert. Um eine Reihendarstellung für K mittels Mercers Theorems zu erhalten, muss gezeigt werden, dass T_K positiv ist. Es gilt ganz allgemein:

Lemma 2.3 (Aufgabe VI.7.16 [Wer00]). *Sei $K \in C(\mathcal{T}^2)$ hermitesch und T_K der zugehörige Hilbert-Schmidt-Integraloperator. Ist K positiv semidefinit, so ist T_K positiv.*

Beweis. Der Beweis wird für den Fall $\mathcal{T} = [0, 1]$ geführt. Es ist zu zeigen, dass für jedes $f \in L^2([0, 1], \mathcal{B}([0, 1]), \lambda)$ gilt $\langle T_K f, f \rangle \geq 0$. Sei zunächst $f \in C([0, 1])$. Dann ist auch die Abbildung $(s, t) \mapsto K(s, t)f(t)\overline{f(s)}$ stetig und damit über der kompakten Menge $[0, 1] \times [0, 1]$ Riemann-integrierbar, wobei das Riemann-Integral mit dem Lebesgue-Integral übereinstimmt. Seien $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{N(n)}^{(n)} = 1$ Partitionen des Intervalls $[0, 1]$ mit $\sup_{i \leq N(n)} |t_i^{(n)} - t_{i-1}^{(n)}| \rightarrow 0$ für $n \rightarrow \infty$. Dann ist das Skalarprodukt der Grenzwert von

Riemann-Summen:

$$\begin{aligned}\langle T_K f, f \rangle &= \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, t) f(t) \overline{f(s)} ds dt \\ &= \lim_{n \rightarrow \infty} \sum_{i,j=1}^{N^{(n)}} K(t_i^{(n)}, t_j^{(n)}) f(t_j^{(n)}) \overline{f(t_i^{(n)})} (t_i^{(n)} - t_{i-1}^{(n)}) (t_j^{(n)} - t_{j-1}^{(n)}) \\ &= \lim_{n \rightarrow \infty} \sum_{i,j=1}^{N^{(n)}} a_i^{(n)} K(t_i^{(n)}, t_j^{(n)}) \overline{a_j^{(n)}} \geq 0.\end{aligned}$$

Mit der Wahl $a_i^{(n)} = f(t_i^{(n)}) (t_i^{(n)} - t_{i-1}^{(n)})$ für $i = 1, \dots, N^{(n)}$ und für alle $n \in \mathbb{N}$. Die letzte Ungleichung folgt, da K positiv semidefinit ist.

Sei nun allgemeiner $f \in L^2([0, 1], \mathcal{B}([0, 1]), \lambda)$. Da $C([0, 1])$ dicht in $L^2([0, 1], \mathcal{B}([0, 1]), \lambda)$ liegt, existiert eine Folge $(f_n)_{n \in \mathbb{N}} \subseteq C([0, 1])$ mit $f_n \rightarrow f$. Da T_K ein linearer beschränkter Operator ist, ist T_K stetig und damit auch die Abbildung Φ definiert durch $\Phi(f) = \langle T_K f, f \rangle$.

$$\langle T_K f, f \rangle = \Phi(f) = \Phi(\lim_{n \rightarrow \infty} f_n) = \lim_{n \rightarrow \infty} \Phi(f_n) = \lim_{n \rightarrow \infty} \langle T_K f_n, f_n \rangle \geq 0 \quad (2.2)$$

impliziert die Behauptung. \square

Bemerkung 2.4. Diese Bemerkung richtet sich nach [Wer00, Bem. Lem. VI.4.5]. Die Eigenschaft der Positiv-Semidefinitheit ist nicht zu verwechseln mit der Eigenschaft Positivität, das heißt $K(s, t) \geq 0$ für alle $s, t \in \mathcal{T}$.

- 1) Ein positiver Kern braucht nicht semidefinit zu sein. Man betrachte den reellwertigen Kern $K(s, t) := |s - t| \geq 0$. Dieser ist nicht positiv semidefinit, da für $a_1 = -1, a_2 = 1$ und $t_1, t_2 \in \mathcal{T}$ gilt: $\sum_{i,j=1}^2 a_i |t_i - t_j| a_j = a_1 a_2 |t_1 - t_2| + a_2 a_1 |t_2 - t_1| = -2|t_1 - t_2|$.
- 2) Ein positiver Kern braucht nicht einen positiven Integraloperator zu induzieren. Angenommen dies wäre der Fall, dann würde für K definiert wie in Punkt 1) Mercers Theorem gelten. Damit wäre

$$\int_{\mathcal{T}} |K(t, t)| dt = \sum_{i=1}^{\infty} \lambda_i \int_{\mathcal{T}} \phi_i(t) \overline{\phi_i(t)} dt = \sum_{i=1}^{\infty} \lambda_i. \quad (2.3)$$

Allerdings ist $K(t, t) = 0$ für jedes $t \in \mathcal{T}$, weshalb $0 = \sum_{i=1}^{\infty} \lambda_i$. Da außerdem T_K positiv ist, gilt $\langle T_K \phi_i, \phi_i \rangle = \lambda_i \|\phi_i\|^2 \geq 0$, weshalb jedes λ_i nicht negativ ist. Dann folgt aber aus dem Spektraltheorem 1.9 $T_K \equiv 0$.

Zu jeder stetigen Kovarianzfunktion K existiert demnach ein kompakter, selbstdadjungierter und positiver Hilbert-Schmidt-Integraloperator T_K . Löst man die Fredholm'sche Integralgleichung, kann man K nach Mercers Theorem durch die Eigenwerte und Eigenfunktionen von T_K darstellen. An dieser Stelle wird erneut auf die Abbildungen 1 und 2 verwiesen. Die dort zugrundeliegenden Hilbert-Schmidt-Kerne sind stets Kovarianzfunktionen.

2.2 Konvergenz und Stetigkeit im quadratischen Mittelwert

In diesem Abschnitt wird ein Kriterium für stochastische Prozesse zweiter Ordnung bewiesen, welches Stetigkeit der Kovarianzfunktion impliziert.

Die Konvergenz bezüglich der Norm von $L^2(\Omega, \mathcal{F}, \mathbf{P})$ wird auch als *Konvergenz im quadratischen Mittelwert*, im Folgenden Konvergenz im q.M., bezeichnet. Man schreibt

$$X_n \xrightarrow{q.M.} X : \|X_n - X\|_{L^2(\Omega, \mathcal{F}, \mathbf{P})} \rightarrow 0 \Leftrightarrow \mathbf{E}|X_n - X|^2 \rightarrow 0. \quad (2.4)$$

Da $L^2(\Omega, \mathcal{F}, \mathbf{P})$ als Hilbertraum vollständig ist, gilt insbesondere

$$X_n \xrightarrow{q.M.} X \Leftrightarrow X_m - X_n \xrightarrow{q.M.} 0 \quad \text{falls } m, n \rightarrow \infty. \quad (2.5)$$

Es seien $(X_n)_{n \in \mathbb{N}}, (Y_n)_{n \in \mathbb{N}} \subseteq L^2(\Omega, \mathcal{F}, \mathbf{P})$ und $X, Y \in L^2(\Omega, \mathcal{F}, \mathbf{P})$. Dann gilt das folgende nützliche Lemma, welches aus [Loè91, S. 135] entnommen und an den Kontext dieser Arbeit angepasst wurde.

Lemma 2.5 (Sec. 37.2 a. [Loè91]). *Falls $X_n \xrightarrow{q.M.} X$ und $Y_n \xrightarrow{q.M.} Y$, so gilt $\mathbf{E}X_n\overline{Y_n} \rightarrow \mathbf{E}X\overline{Y}$.*

Beweis. Der Beweis aus der angegebenen Quelle wurde ebenfalls angepasst. Es ist

$$\mathbf{E}(X_n\overline{Y_n} - X\overline{Y}) = \mathbf{E}(X_n - X)(\overline{Y_n} - \overline{Y}) + \mathbf{E}(X_n - X)\overline{Y} + \mathbf{E}X(\overline{Y_n} - \overline{Y}).$$

Da für $n \rightarrow \infty$ gilt

$$|\mathbf{E}(X_n - X)(\overline{Y_n} - \overline{Y})|^2 \stackrel{\text{CSU}}{\leq} \mathbf{E}|X_n - X|^2 \mathbf{E}|\overline{Y_n} - \overline{Y}|^2 \rightarrow 0,$$

und da mit dem selben Argument auch die beiden verbleibenden Terme auf der rechten Seite verschwinden, folgt die Behauptung. \square

Definition 2.6 (vgl. Sec. 37.2 A. [Loè91]). Ein stochastischer Prozess heißt *stetig im quadratischen Mittelwert* oder L^2 -stetig, falls

$$\lim_{\eta \rightarrow 0} \|X_{t+\eta} - X_t\|_{L^2(\Omega, \mathcal{F}, \mathbf{P})} = 0 \quad \forall t \in \mathcal{T}, t + \eta \in \mathcal{T}$$

Im Folgenden wird die Abkürzung *stetig im q.M.* verwendet.

Bemerkung 2.7. Stetigkeit eines Prozesses X im q.M. entspricht Stetigkeit der Abbildungen $\mathcal{T} \ni t \mapsto X_t \in L^2(\Omega, \mathcal{F}, \mathbf{P})$, welche in Abschnitt 3.2 von Bedeutung sein wird.

Um dies zu sehen, sei $\varepsilon > 0$. Da $\lim_{\eta \rightarrow 0} \|X_{t+\eta} - X_t\|_{L^2(\Omega, \mathcal{F}, \mathbf{P})} = 0$, existiert für alle $k \in \mathbb{N}$ ein $\eta(k) > 0$ derart, dass für alle $\delta < \eta(k)$ gilt $\|X_{t+\eta} - X_t\| < \frac{1}{k}$. Wählt man k_0 derart, dass $\varepsilon > \frac{1}{k_0}$, so folgt für $\delta < \eta(k_0)$, dass für alle $u \in \mathcal{T}$ mit $\|u - t\| < \delta$ gilt $\|X_u - X_t\| < \varepsilon$. Formal gilt damit $\forall \varepsilon > 0 \exists \delta > 0 \forall u \in \mathcal{T} : |u - t| < \delta \Rightarrow \|X_u - X_t\| < \varepsilon$.

Nun kann das anfangs erwähnte Kriterium formuliert werden.

Satz 2.8 (Lem. 4.2. [Ale15]). *Ein stochastischer Prozess $\{X_t : t \in \mathcal{T}\}$ ist stetig im quadratischen Mittelwert genau dann, wenn seine Kovarianzfunktion K stetig auf $\mathcal{T} \times \mathcal{T}$ ist.*

Der folgende Beweis wurde [Ale15] nachempfunden und an den Kontext komplexwertiger Prozesse angepasst.

Beweis. Sei zunächst K stetig auf $\mathcal{T} \times \mathcal{T}$. Für $\varepsilon > 0$ gilt

$$\begin{aligned} \|X_{t+\varepsilon} - X_t\|^2 &= \mathbf{E}(X_{t+\varepsilon} - X_t)(\overline{X_{t+\varepsilon} - X_t}) \\ &= \mathbf{E}X_{t+\varepsilon}\overline{X}_{t+\varepsilon} - \mathbf{E}X\overline{X}_{t+\varepsilon} - \mathbf{E}X_{t+\varepsilon}\overline{X} + \mathbf{E}X\overline{X} \\ &= K(t + \varepsilon, t + \varepsilon) - K(t, t + \varepsilon) - K(t + \varepsilon, t) + K(t, t) \\ &\xrightarrow[\varepsilon \rightarrow 0]{} 0, \end{aligned}$$

womit Stetigkeit von X im q.M. folgt. Umgekehrt gilt für $\varepsilon > 0$ und $\delta > 0$

$$\begin{aligned} |K(s + \delta, t + \varepsilon) - K(s, t)| &= |\mathbf{E}X_{s+\delta}\overline{X}_{t+\varepsilon} - \mathbf{E}X_s\overline{X}_t| \\ &= |\mathbf{E}(X_{s+\delta} - X_s)\overline{(X_{t+\varepsilon} - X_t)} + \mathbf{E}(X_{s+\delta} - X_s)\overline{X}_t + \mathbf{E}X_s(\overline{X_{t+\varepsilon} - X_t})| \\ &\leq |\langle X_{s+\delta} - X_s, X_{t+\varepsilon} - X_t \rangle| + |\langle X_{s+\delta} - X_s, X_t \rangle| + |\langle X_s, X_{t+\varepsilon} - X_t \rangle| \\ &\leq \|X_{s+\delta} - X_s\| \|X_{t+\varepsilon} - X_t\| + \|X_{s+\delta} - X_s\| \|X_t\| + \|X_s\| \|X_{t+\varepsilon} - X_t\|, \end{aligned}$$

wobei man die vorletzte Ungleichung mit der Dreiecksungleichung und die letzte Ungleichung mit der Cauchy-Schwarz'schen Ungleichung erhält. Wegen der Stetigkeit von X im q.M. folgt dann unmittelbar

$$\lim_{(\delta, \varepsilon) \rightarrow (0, 0)} |K(s + \delta, t + \varepsilon) - K(s, t)| = 0.$$

Die Kovarianzfunktion von X ist also stetig in $\mathcal{T} \times \mathcal{T}$. □

Beispiel 2.9. Stetigkeit im q.M. impliziert keineswegs Stetigkeit der Pfade. Um dies zu sehen, sei B eine Prä-Brown'sche Bewegung auf $[0, 1]$. Dann hat B die stetige Kovarianzfunktion $K(s, t) = s \wedge t$, ist also stetig im q.M.. Trotzdem sind Prä-Brown'sche Bewegungen im Allgemeinen nicht stetig.

Als Ergebnis dieses Abschnittes beachte man das folgende Korollar.

Korollar 2.10. *Sei X ein stochastischer Prozess zweiter Ordnung, stetig im quadratischen Mittelwert. Dann ist die Kovarianzfunktion K von X ein Hilbert-Schmidt-Kern und der induzierte Hilbert-Schmidt-Integraloperator T_K ist positiv.*

Beweis. Nach Satz 2.8 ist K stetig. Zusammen mit Bemerkung 1.11 ist K ein Hilbert-Schmidt-Operator. Die Positivität folgt schließlich aus Lemma 2.3. □

2.3 Hilberträume mit reproduzierendem Kern

Zu einem stochastischen Prozess $X = \{X_t : t \in \mathcal{T}\}$ kann man den *von X generierten Hilbertraum* $\mathcal{H}_X := \text{span}\{X_t, t \in \mathcal{T}\}$ betrachten. Dieser Raum ist dabei per Definition des Spans abgeschlossen und ein Teilraum von $L^2(\Omega, \mathcal{F}, \mathbf{P})$ und erbt auch dessen inneres Produkt. In diesem Abschnitt soll der Hilbertraum \mathcal{H}_X näher untersucht werden. Dazu wird ein Exkurs in die Theorie der Hilberträume mit reproduzierendem Kern bereitgestellt, welche eng mit Kovarianzfunktionen von stochastischen Prozessen und den von ihnen generierten Hilberträumen zusammenhängen. Es wird dabei weitestgehend auf Beweise und vertiefende Ausführungen verzichtet, um nicht den Rahmen dieser Arbeit zu überschreiten. Details können in dem umfangreichen Buch von A. Berlinet und C. Thomas-Agnan [BTA04, Kapitel 1 und 2] nachvollzogen werden. Für einige Eigenschaften wird auch [Wan08, Kapitel 1] zitiert. Sei E eine nichtleere Menge und $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ ein Hilbertraum aus Funktionen, die über E definiert seien.

Definition 2.11 (Def. 1 [BTA04]). Eine Funktion $K : E \times E \rightarrow \mathbb{C}$ heißt *reproduzierender Kern* des Hilbertraumes \mathcal{H} genau dann, wenn

- i) $\forall t \in E : K(\cdot, t) \in \mathcal{H}$
- ii) $\forall t \in E, \forall \varphi \in \mathcal{H} : \langle \varphi, K(\cdot, t) \rangle_{\mathcal{H}} = \varphi(t).$ (reproduzierende Eigenschaft)

Der Wert von φ an der Stelle t wird vom inneren Produkt von φ und $K(\cdot, t)$ reproduziert. Ein Hilbertraum komplexwertiger Funktionen, der einen reproduzierenden Kern besitzt, heißt *Hilbertraum mit reproduzierendem Kern* (*Engl. reproducing kernel hilbert space*), abkürzend *RKHS*.

Bemerkung 2.12. Aus i) und ii) folgt für alle $(s, t) \in E \times E : K(s, t) = \langle K(\cdot, t), K(\cdot, s) \rangle_{\mathcal{H}}$.

Einige grundlegende Eigenschaften werden in der folgenden Proposition zusammengetragen.

Proposition 2.13 (Chap. 1 [Wan08]). (a) Falls ein reproduzierender Kern K zum Hilbertraum \mathcal{H} existiert, so ist er eindeutig.

(b) Ein reproduzierender Kern ist positiv semidefinit.

(c) In einem RKHS impliziert Konvergenz in der Norm punktweise Konvergenz.

Beweis. Die Beweise orientieren sich an [Wan08]. (a): Angenommen \hat{K} ist ein weiterer reproduzierender Kern zum Hilbertraum \mathcal{H} . Dann gilt

$$\begin{aligned} 0 &\leq \|K(\cdot, t) - \hat{K}(\cdot, t)\|^2 = \langle K(\cdot, t) - \hat{K}(\cdot, t), K(\cdot, t) \rangle - \langle K(\cdot, t) - \hat{K}(\cdot, t), \hat{K}(\cdot, t) \rangle \\ &= K(t, t) - \hat{K}(t, t) - K(t, t) + \hat{K}(t, t) = 0. \end{aligned}$$

(b): Für $n \in \mathbb{N}$ und $t_1, \dots, t_n \in E$ und $a_1, \dots, a_n \in \mathbb{C}$ gilt:

$$\sum_{i,j=1}^n a_i K(t_j, t_i) \bar{a}_j = \left\langle \sum_{i=1}^n a_i K(\cdot, t_i), \sum_{j=1}^n \bar{a}_j K(\cdot, t_j) \right\rangle = \left\| \sum_{i=1}^n a_i K(\cdot, t_i) \right\|^2 \geq 0. \quad (2.6)$$

(c): Seien $(f_n)_{n \in \mathbb{N}} \subseteq \mathcal{H}$ mit $f_n \rightarrow f \in \mathcal{H}$ für $n \rightarrow \infty$.

$$|f_n(t) - f(t)| = |\langle f_n - f, K(\cdot, t) \rangle| \stackrel{\text{CSU}}{\leq} \|f_n - f\| \|K(\cdot, t)\| \rightarrow 0 \quad \text{für } n \rightarrow \infty. \quad (2.7)$$

□

Der folgende Satz stellt eine wichtige Eins-zu-eins-Beziehung zwischen positiv-semidefiniten Funktionen und RKHSs dar. Er ist in [BTA04, Thm. 3] zu finden und wird dort ausführlich bewiesen.

Satz 2.14 (Moore-Aronszajin). *Zu jedem RKHS \mathcal{H} existiert genau ein reproduzierender Kern $K(\cdot, \cdot)$. Umgekehrt existiert zu jeder positiv semidefiniten Funktion $K(\cdot, \cdot)$ genau ein RKHS \mathcal{H}_K .*

$$\mathcal{H}_K = \text{span}\{K(\cdot, t), t \in E\} \quad (2.8)$$

Beweis. Es wird auf den Beweis von Theorem 3 in [BTA04] verwiesen. □

Bemerkung 2.15. Der Raum \mathcal{H}_K^0 der endlichen Linearkombinationen der $(K(\cdot, t))_{t \in E}$ ist dicht in \mathcal{H}_K und für $\varphi, \phi \in \mathcal{H}_K^0$ ist

$$\langle \varphi, \phi \rangle_{\mathcal{H}_K^0} = \sum_{i=1}^n \sum_{j=1}^m \alpha_i \bar{\beta}_j K(t_i, s_j), \quad \text{wobei } \varphi = \sum_{i=1}^n \alpha_i K(t_i, \cdot) \text{ und } \phi = \sum_{j=1}^m \beta_j K(s_j, \cdot). \quad (2.9)$$

Wichtig für diese Arbeit ist auch der folgende Satz.

Satz 2.16 (Cor. 5 [BTA04]). *Sei \mathcal{H}_K ein RKHS aus Funktionen, die auf einem kompakten metrischen Raum (E, d) definiert sind. Falls K stetig ist, so ist \mathcal{H}_K separabel und ein Raum stetiger Funktionen.*

Beweis. Es wird auf den Beweis von Korollar 5 in [BTA04] verwiesen. □

Soweit zur reinen Theorie der Hilberträume mit reproduzierendem Kern. Um den Bogen zur eigentlichen Aufgabe des Kapitels zu führen, sei nun \mathcal{H}_X der von einem zentrierten Prozess X zweiter Ordnung mit Kovarianzfunktion K generierte Hilbertraum.

Da K positiv semidefinit ist, existiert genau ein RKHS \mathcal{H}_K mit Kern K .

Man kann eine lineare Abbildung $\Theta : \mathcal{H}_K \rightarrow \mathcal{H}_X$ durch

$$\Theta(\varphi) = \Theta \left(\sum_{i=1}^n \alpha_i K(t_i, \cdot) \right) = \sum_{i=1}^n \alpha_i X_{t_i}, \quad \text{wobei } \varphi = \sum_{i=1}^n \alpha_i K(t_i, \cdot) \quad (2.10)$$

definieren. Θ wird als der kanonische isometrische Isomorphismus zwischen \mathcal{H}_K und \mathcal{H}_X bezeichnet.

In der Tat ist Θ eine Isometrie, denn

$$\langle \Theta(\varphi), \Theta(\phi) \rangle_{\mathcal{H}_X} = \left\langle \sum_{i=1}^n \alpha_i X_{t_i}, \sum_{j=1}^m \beta_j X_{s_j} \right\rangle_{\mathcal{H}_X} = \sum_{i=1}^n \sum_{j=1}^m \alpha_i \overline{\beta_j} K(t_i, s_j) = \langle \varphi, \phi \rangle_{\mathcal{H}_K}. \quad (2.11)$$

Da Θ ein isometrischer Isomorphismus ist und nach Satz 2.16 für stetiges K der Raum \mathcal{H}_K separabel ist, folgt:

Korollar 2.17. *Sei $X = \{X_t, t \in \mathcal{T}\}$ ein zentrierter stochastischer Prozess zweiter Ordnung. Ist X stetig im quadratischen Mittelwert, so ist \mathcal{H}_X separabel.*

Dieses Korollar ist im folgenden Kapitel für Fragen im Zusammenhang mit Messbarkeit von Bedeutung.

Des Weiteren kann man aus der Existenz der Isometrie Θ folgern, dass für eine Orthonormalbasis $(\varphi_i)_{i \in \mathbb{N}}$ von \mathcal{H}_K , für $\xi_i := \Theta(\varphi_i)$, damit $(\xi_i)_{i \in \mathbb{N}}$ Orthonormalbasis in \mathcal{H}_X ist. Insbesondere folgt, dass $\xi_i \sim \mathcal{N}(0, 1)$ und

$$X_t = \sum_{i=1}^{\infty} \xi_i \mathbf{E} X_t \overline{\xi_i}. \quad (2.12)$$

mit Konvergenz in $L^2(\Omega, \mathcal{F}, \mathbf{P})$. Dies gilt in der Tat, denn $K(t, \cdot) = \sum_{i=1}^{\infty} \langle K(t, \cdot), \varphi_i \rangle_{\mathcal{H}_K} \varphi_i$ und damit

$$\Theta(K(t, \cdot)) = \sum_{i=1}^{\infty} \langle K(t, \cdot), \varphi_i \rangle_{\mathcal{H}_K} \Theta(\varphi_i) = \sum_{i=1}^{\infty} \langle X_t, \xi_i \rangle_{\mathcal{H}_X} \xi_i = \sum_{i=1}^{\infty} \xi_i \mathbf{E} X_t \overline{\xi_i}. \quad (2.13)$$

Da Θ eine Isometrie ist und K ein reproduzierender Kern von \mathcal{H}_K gilt weiter:

$$\mathbf{E} X_t \overline{\xi_i} = \langle K(\cdot, t), \varphi_i \rangle_{\mathcal{H}_K} = \varphi_i(t). \quad (2.14)$$

Zusammenfassend ergibt sich der folgende Satz:

Satz 2.18 (Thm 3.1.1 [Adl10]). *Ist $(\varphi_i)_{i \in \mathbb{N}}$ vollständiges Orthonormalbasis in \mathcal{H}_K , dann hat X die $L^2(\Omega, \mathcal{F}, \mathbf{P})$ -Repräsentation*

$$X_t = \sum_{i=1}^{\infty} \xi_i \varphi_i(t), \quad (2.15)$$

wobei $(\xi_i)_{i \in \mathbb{N}}$ eine orthonormale Folge zentraler Zufallsvariablen sei, die definiert ist durch $\xi_i := \Theta(\varphi_i)$ für den kanonischen isometrischen Isomorphismus Θ von \mathcal{H}_K und \mathcal{H}_X .

Die Theorie der RKHS umfasst noch viele weitere spannende Zusammenhänge. Insbesondere sind hier Kriterien zu nennen, die garantieren, dass eine Funktion $f \in L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ in einem RKHS \mathcal{H}_K enthalten ist. Einen vollständigen Beweis des Karhunen-Loève-Theorems über die RKHS Theorie findet man in [BTA04, Kapitel 3.2, Korollar 7].

Kapitel 3

Das Bochner-Integral

Sei wie zuvor $(\Omega, \mathcal{F}, \mathbf{P})$ ein Wahrscheinlichkeitsraum, $\mathcal{T} \subseteq \mathbb{R}$ kompakt und $X = \{X_t, t \in \mathcal{T}\}$ ein stochastischer Prozess zweiter Ordnung, stetig im quadratischen Mittelwert. Ziel dieses Abschnittes ist es, Integrale der Form $\int_{\mathcal{T}} X_t \lambda(dt)$ sinnvoll zu definieren. Dies ist keineswegs trivial. Möchte man dies pfadweise tun, so scheitert man schnell, denn für festes $\omega \in \Omega$ ist nicht ohne Weiteres klar, ob $X(\cdot)(\omega)$ in $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda|_{\mathcal{T}})$ liegt.

Eine Variante ist, das Integral ähnlich zu dem sogenannten Riemann-Stieltjes-Integral aufzubauen und als einen Grenzwert im quadratischen Mittelwert von Riemann Stieltjes Summen zu definieren. Diese Strategie verfolgt [Loè91]. Spätestens ab dem Punkt, wo diese Integrale mehr-dimensional werden, ergeben sich dabei Schwierigkeiten.

Abhilfe verschafft die Theorie der Bochner-Integration, die Integrale für banachraumwertige Funktionen bezüglich Maßen definiert. Genau in einem solchen Fall befindet man sich, denn die Funktion $t \mapsto X_t$ bildet in den Hilbertraum $L^2(\Omega, \mathcal{F}, \mathbf{P})$ ab, der selbstverständlich ein Banachraum ist. Von entscheidender Bedeutung wird die Stetigkeit von X im quadratischen Mittelwert sein. Hierzu soll nun eine Einführung und ein Überblick gegeben werden.

3.1 Aufbau der allgemeinen Theorie

Der folgende Abschnitt hält sich sehr eng an das Buch [AB06, Absch. 11.8], welches das Bochner-Integral gut zugänglich einführt.

Sei $(\mathcal{T}, \Sigma, \mu)$ ein messbarer Raum und \mathbf{B} ein Banachraum. Ähnlich wie bei der Theorie zur Lebesgue-Integration beginnt der Aufbau des Integrals mit sogenannten Treppenfunktionen.

Definition 3.1. i) Eine Funktion $\varphi : \mathcal{T} \rightarrow \mathbf{B}$ heißt **\mathbf{B} -Treppenfunktion**, falls sie nur endlich viele Werte x_1, \dots, x_n aus \mathbf{B} annimmt und $\mu(\varphi^{-1}(\{x_i\})) < \infty$ für $x_i \neq 0$.

ii) Setzt man $A_i := \varphi^{-1}(\{x_i\})$, so heißt die Formel

$$\varphi = \sum_{i=1}^n x_i \mathbf{1}_{A_i} \tag{3.1}$$

die *Standard-Repräsentation* von φ .

iii) Das *Bochner-Integral* von φ ist definiert als

$$\int \varphi d\mu := \sum_{i=1}^n x_i \mu(A_i) \in \mathbf{B}. \quad (3.2)$$

Die \mathbf{B} -Treppenfunktionen bilden einen Vektorraum, welcher fortan als $L_{\mathbf{B}}$ bezeichnet wird.

Diese Definition ist unabhängig von der gewählten Repräsentation von φ , denn falls neben der Standard-Repräsentation auch $\varphi = \sum_{i=1}^m y_i \mathbb{1}_{B_i}$ gilt, so zeigt man genau wie im bereits bekannten Falle von \mathbb{R} -Treppenfunktionen, dass $\int \varphi d\mu = \sum_{i=1}^m y_i \mu(B_i)$.

Die eigentliche Problemstellung ist nun aber, das Integral von $L_{\mathbf{B}}$ auf allgemeinere Klassen von Funktionen auszudehnen. Sei die *Normfunktion* $\|f\| : \mathcal{T} \rightarrow \mathbb{R}$ einer Funktion $f : \mathcal{T} \rightarrow \mathbf{B}$ definiert durch $\|f\|(t) := \|f(t)\|_{\mathbf{B}}$ für jedes $t \in \mathcal{T}$. Im Folgenden wird auf den Index der Norm verzichtet.

Lemma 3.2 (Lem. 11.35 [AB06]). *Sei $\varphi \in L_{\mathbf{B}}$ mit Standard-Repräsentation $\varphi = \sum_{i=1}^n x_i \mathbb{1}_{A_i}$. Dann ist die Normfunktion $\|\varphi\|$ eine \mathbb{R} -Treppenfunktion und es gilt*

$$\|\varphi\| = \sum_{i=1}^n \|x_i\| \mathbb{1}_{A_i}. \quad (3.3)$$

Weiter gilt

$$\int \|\varphi\| d\mu = \sum_{i=1}^n \|x_i\| \mu(A_i) \quad \text{und} \quad \left\| \int \varphi d\mu \right\| \leq \int \|\varphi\| d\mu. \quad (3.4)$$

Beweis. O.B.d.A. seien die $(A_i)_{i \in \mathbb{N}}$ disjunkt. Für $t \in \mathcal{T}$ existiert maximal ein $1 \leq j \leq n$, sodass $\mathbb{1}_{A_j}(t) \neq 0$. Dann gilt: $\| \sum_{i=1}^n x_i \mathbb{1}_{A_i}(t) \| = \|x_j\| \mathbb{1}_{A_j}(t) = \sum_{i=1}^n \|x_i\| \mathbb{1}_{A_i}(t)$. Der erste Ausdruck folgt mit der Definition des Lebesgue-Integrals und der zweite aus der Dreiecksungleichung. \square

Definition 3.3 (Def. 11.36 [AB06]). Sei $(\mathcal{T}, \Sigma, \mu)$ ein messbarer Raum und $f : \mathcal{T} \rightarrow \mathbf{B}$. Die Abbildung f heißt *stark μ -messbar*, falls eine Folge $(\varphi_n)_{n \in \mathbb{N}} \subseteq L_{\mathbf{B}}$ mit

$$\lim_{n \rightarrow \infty} \|f(t) - \varphi_n(t)\| = 0 \quad \mu\text{-fast überall} \quad (3.5)$$

existiert. Falls klar ist, welches Maß gemeint ist, nennt man f abgekürzt *stark messbar*.

Die stark messbaren Funktionen können mittels des folgenden Lemmas charakterisiert werden.

Lemma 3.4 (Lem. 11.37 [AB06]). *Eine Funktion $f : \mathcal{T} \rightarrow \mathbf{B}$ ist stark messbar genau dann, wenn sie messbar ist und ihre Werte $f(t)$ für μ -fast alle $t \in \mathcal{T}$ in einem separablen abgeschlossenen Untervektorraum von \mathbf{B} liegen.*

Beweis. (\Rightarrow) : Die Messbarkeit folgt unmittelbar. Da f stark messbar ist, existiert eine μ -Nullmenge \mathcal{N} , sodass $f|_{\mathcal{T} \setminus \mathcal{N}}$ punktweiser Limes einer Folge $(\varphi_n)_{n \in \mathbb{N}} \subseteq L_{\mathbf{B}}$ ist (eingeschränkt auf den Definitionsbereich $\mathcal{T} \setminus \mathcal{N}$). Man stellt leicht fest, dass $A := \bigcup_{n \in \mathbb{N}} \varphi_n(\mathcal{T} \setminus \mathcal{N})$ abzählbar ist, da \mathbf{B} -Treppenfunktionen per Definition nur endlich viele Werte annehmen. Da A dicht in \overline{A} liegt, ist \overline{A} somit separabel. Per Konstruktion gilt $f(\mathcal{T} \setminus \mathcal{N}) \subseteq \overline{A}$.

(\Leftarrow) : Ist umgekehrt f messbar mit μ -fast allen Werten in einem abgeschlossenen separablen Untervektorraum Y von \mathbf{B} , so existiert eine Nullmenge \mathcal{N} derart, dass $f|_{\mathcal{T} \setminus \mathcal{N}}$ punktweiser Limes einer Folge von Funktionen in $L_{\mathbf{B}}$ ist. Dies impliziert starke Messbarkeit von f . \square

Korollar 3.5 (Lem. 11.38 [AB06]). *Sei eine Funktion $f : \mathcal{T} \rightarrow \mathbf{B}$ μ -fast überall punktweiser Grenzwert einer Folge stark messbarer Funktionen. Dann ist f stark messbar.*

Beweis. Die Aussage folgt unmittelbar aus dem vorherigem Lemma und dem Fakt, dass der punktweise Limes messbarer Funktionen messbar ist. \square

Lemma 3.6 (Lem. 11.39 [AB06]). *Ist eine Funktion $f : \mathcal{T} \rightarrow \mathbf{B}$ stark messbar, so ist die Normfunktion $\|f\|$ messbar.*

Beweis. Der Beweis stammt aus der obigen Quelle. Aus der umgekehrten Dreiecksungleichung folgt

$$|\|f(t)\| - \|\varphi_n(t)\|| \leq \|f(t) - \varphi_n(t)\| \quad \mu\text{-fast überall} \quad (3.6)$$

und daher gilt $\|\varphi_n(t)\| \rightarrow \|f(t)\|$ μ -fast überall, wodurch Messbarkeit von $\|f\|$ impliziert wird. \square

Die Menge aller stark messbaren Funktionen von \mathcal{T} nach \mathbf{B} wird mit $\mathcal{M}(\mathcal{T}, \mathbf{B})$ notiert. Man kann zeigen, dass $\mathcal{M}(\mathcal{T}, \mathbf{B})$ ein Vektorraum ist. Es gelten die Vektorrauminklusionen:

$$L_{\mathbf{B}} \subseteq \mathcal{M}(\mathcal{T}, \mathbf{B}) \subseteq \mathbf{B}^{\mathcal{T}}. \quad (3.7)$$

Um das Bochner-Integral auf $\mathcal{M}(\mathcal{T}, \mathbf{B})$ auszudehnen, benötigt man folgendes Lemma:

Lemma 3.7 (Lem. 11.41 [AB06]). *Sei $f \in \mathcal{M}(\mathcal{T}, \mathbf{B})$. Angenommen für zwei Folgen von Treppenfunktionen $(\varphi_n)_{n \in \mathbb{N}}, (\psi_n)_{n \in \mathbb{N}} \subseteq L_{\mathbf{B}}$ sind die reellwertigen und messbaren Funktionen $\|f - \varphi_n\|$ und $\|f - \psi_n\|$ Lebesgue-integrierbar für jedes $n \in \mathbb{N}$ und es gilt*

$$\lim_{n \rightarrow \infty} \int \|\varphi_n\| d\mu = \lim_{n \rightarrow \infty} \int \|\psi_n\| d\mu = 0. \quad (3.8)$$

Dann gilt für jedes $A \in \Sigma$

$$\lim_{n \rightarrow \infty} \int_A \varphi_n d\mu = \lim_{n \rightarrow \infty} \int_A \psi_n d\mu, \quad (3.9)$$

wobei die letzten Grenzwerte bezüglich $\|\cdot\|_{\mathbf{B}}$ zu verstehen sind.

Beweis. Der Beweis stammt aus der obigen Quelle. Seien $(\varphi_n)_{n \in \mathbb{N}}, (\psi_n)_{n \in \mathbb{N}} \subseteq L_{\mathbf{B}}$ mit der gewünschten Eigenschaft. Für ein festes $A \in \Sigma$ und unter der Verwendung von Lemma 3.2 gilt dann

$$\begin{aligned} \left\| \int_A \varphi_n d\mu - \int_A \varphi_m d\mu \right\| &= \left\| \int_A (\varphi_n - \varphi_m) d\mu \right\| \\ &\leq \int \|\varphi_n - \varphi_m\| d\mu \\ &\leq \int \|\varphi_n - f\| d\mu + \int \|f - \varphi_m\| d\mu, \end{aligned}$$

wobei die rechte Seite beliebig klein wird. Da \mathbf{B} vollständig ist, konvergiert die Cauchy-Folge $(\int_A \varphi_n d\mu)_{n \in \mathbb{N}} \subseteq \mathbf{B}$. Analog kann man für $(\int_A \psi_n d\mu)_{n \in \mathbb{N}} \subseteq \mathbf{B}$ argumentieren. Da

$$0 \leq \left\| \int_A \varphi_n d\mu - \int_A \psi_n d\mu \right\| \leq \int \|f - \varphi_n\| d\mu + \int \|f - \varphi_n\| d\mu, \quad (3.10)$$

erhält man letztendlich

$$\lim_{n \rightarrow \infty} \int_A \varphi_n d\mu = \lim_{n \rightarrow \infty} \int_A \psi_n d\mu. \quad (3.11)$$

□

Definition 3.8 (Def. 11.42 [AB06]). Eine stark μ -messbare Funktion $f : \mathcal{T} \rightarrow \mathbf{B}$ heißt *Bochner-integrierbar*, falls eine Folge $(\varphi_n)_{n \in \mathbb{N}} \subseteq L_{\mathbf{B}}$ existiert, sodass die reellen messbaren Funktionen $\|f - \varphi_n\|$ Lebesgue-integrierbar sind für jedes $n \in \mathbb{N}$ und

$$\lim_{n \rightarrow \infty} \int \|f - \varphi_n\| d\mu = 0. \quad (3.12)$$

In diesem Falle ist für jedes $A \in \Sigma$ das *Bochner-Integral* von f über A definiert als

$$\int_A f d\mu = \lim_{n \rightarrow \infty} \int_A \varphi_n d\mu, \quad (3.13)$$

wobei der Grenzwert bezüglich $\|\cdot\|_{\mathbf{B}}$ zu verstehen ist.

Nach dem vorhergehenden Lemma ist das Bochner-Integral wohldefiniert, in dem Sinne, dass der Wert des Integrals nicht von der Wahl der Folge von \mathbf{B} -Treppenfunktionen abhängt.

Satz 3.9 (Thm. 11.43 [AB06]). *Seien f und g Bochner-integrierbare Funktionen und $\alpha, \beta \in \mathbb{R}$, dann ist $\alpha f + \beta g$ Bochner integrierbar und es gilt*

$$\int_A (\alpha f + \beta g) d\mu = \alpha \int_A f d\mu + \beta \int_A g d\mu \quad \text{für jedes } A \in \Sigma. \quad (3.14)$$

Ist μ -fast überall $f \leq g$, dann gilt

$$\int_A f d\mu \leq \int_A g d\mu \quad \text{für jedes } A \in \Sigma. \quad (3.15)$$

Beweis. Anwendung der Definition und einfache Umformungen liefern die Behauptung. \square

Für endliche Maßräume gibt es ein starkes Kriterium für Bochner-Integrierbarkeit. Dieses wird erlauben, wichtige Eigenschaften des Lebesgue-Integrals auf das Bochner-Integral zu übertragen. Das Kriterium ist als der Satz von Bochner bekannt.

Satz 3.10 (Bochner). *Sei $(\mathcal{T}, \Sigma, \mu)$ ein endlicher Maßraum und f eine stark μ -messbare Funktion. Dann ist f Bochner-integrierbar genau dann, wenn seine Normfunktion $\|f\|$ Lebesgue-integrierbar ist.*

Beweis. Der Beweis richtet sich nach [Wit05, Kap. 4, Satz 4.2] und wurde an den allgemeineren Kontext angepasst.

(\Rightarrow) : Da f Bochner-integrierbar ist, existiert eine Folge $(\varphi_n)_{n \in \mathbb{N}} \subseteq L_{\mathbf{B}}$ sodass $\varphi_n \rightarrow f$ in \mathbf{B} fast überall und $\lim_{n \rightarrow \infty} \int_{\mathcal{T}} \|f - \varphi_n\| d\mu = 0$. Es wurde bereits gezeigt, dass dann auch $\|\varphi_n\| \rightarrow \|f\|$ in \mathbb{R} fast überall und $\|f\|$ messbar ist. Nach dem Lemma von Fatou gilt dann

$$\int \|f\| d\mu \leq \liminf_{n \rightarrow \infty} \int \|\varphi_n\| d\mu. \quad (3.16)$$

Außerdem gilt

$$\begin{aligned} \left| \int \|\varphi_n\| d\mu - \int \|\varphi_m\| d\mu \right| &\leq \int |\|\varphi_n\| - \|\varphi_m\|| d\mu \\ &\leq \int \|\varphi_n - \varphi_m\| d\mu \\ &\leq \int \|\varphi_n - f\| d\mu + \int \|\varphi_m - f\| d\mu \rightarrow 0 \quad \text{für } m, n \rightarrow \infty. \end{aligned}$$

Die Folge $(\int \|\varphi_n\| d\mu)_{n \in \mathbb{N}}$ ist demnach cauchy in \mathbb{R} , insbesondere beschränkt. Daher folgt aus 3.16, dass $\|f\|$ Lebesgue-integrierbar ist.

(\Leftarrow) : Nach Voraussetzung ist f stark messbar, weshalb eine Folge $(\varphi_n)_{n \in \mathbb{N}} \subseteq L_{\mathbf{B}}$, die fast sicher in \mathbf{B} gegen f konvergiert, existiert. Definiert man $\psi_n(t) := \varphi_n(t) \mathbf{1}_{\|\varphi_n(t)\| \leq 2\|f(t)\|}$, so gilt weiterhin $(\psi_n)_{n \in \mathbb{N}} \subseteq L_{\mathbf{B}}$ und $\psi_n \rightarrow f$ in \mathbf{B} fast überall. Zusätzlich gilt

$$\|\psi_n - f\| \leq 3\|f\| \leq \infty \quad \text{fast sicher,}$$

weshalb nach dem Satz von der majorisierten Konvergenz

$$\int \|\psi_n - f\| d\mu \rightarrow 0, \quad (3.17)$$

was Bochner-Integrierbarkeit von f impliziert. \square

Der zentrale Satz, auf den dieses Kapitel hingearbeitet hat, kann nun formuliert werden. Er liefert einen äußerst bemerkenswerten Zusammenhang zwischen Bochner-Integrierbarkeit von Funktionen, die Werte in L^p -Räumen annehmen und der Existenz einer außergewöhnlichen Funktion mit gemeinsamer Lebesgue-Messbarkeit.

Satz 3.11 (Absch. III.11 Thm. 17 [DS57]). Seien $(\mathcal{T}, \Sigma, \mu)$ und $(\Omega, \mathcal{F}, \nu)$ endliche Maßräume und $(\mathcal{T} \times \Omega, \Sigma \otimes \mathcal{F}, \mu \otimes \nu)$ deren Produktraum. Sei weiter $1 \leq p \leq \infty$, \mathbf{B} ein reeller oder komplexer Banachraum und $F : \mathcal{T} \rightarrow L^p(\Omega, \mathcal{F}, \nu, \mathbf{B})$ Bochner-integrierbar.

Dann existiert eine $\mu \otimes \nu$ -messbare Funktion $f : \mathcal{T} \times \Omega \rightarrow \mathbf{B}$, die bis auf einer $\mu \otimes \nu$ -Nullmenge eindeutig bestimmt ist und für die gilt $f(t, \cdot) = F(t)$ für μ -fast alle $t \in \mathcal{T}$.

Außerdem ist $f(\cdot, \omega)$ μ -integrierbar auf \mathcal{T} für ν -fast alle $\omega \in \Omega$ und für $\int_{\mathcal{T}} f(t, \omega) \mu(dt)$ als Funktion von ω und $\int_{\mathcal{T}} F(t) \mu(dt)$ als Element in $L^p(\mathcal{T}, \mathcal{F}, \nu, \mathbf{B})$ gilt

$$\int_{\mathcal{T}} f(t, \omega) \mu(dt) = \int_{\mathcal{T}} F(t) \mu(dt). \quad (3.18)$$

Beweis. Ein Beweis kann in [DS57, Absch. III.11 Thm. 17] nachvollzogen werden.

Wie dieser Satz angewendet werden kann, zeigt sich im nun folgenden und letzten Abschnitt vor dem Karhunen-Loëve Theorem.

3.2 Anwendung im Kontext dieser Arbeit

Wie bereits in der Motivation des Bochner-Integrals angesprochen, ist das Ziel des Kapitels, für stochastische Prozesse zweiter Ordnung die stetig im q.M. sind, Integrale der Form $\int_{\mathcal{T}} X_t \lambda(dt)$ als Elemente in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ zu definieren, wobei λ das Lebesgue-Maß auf \mathcal{T} bezeichne. Sei dazu

$$F : \quad \mathcal{T} \longrightarrow L^2(\Omega, \mathcal{F}, \mathbf{P}, \mathbb{C})$$

$$t \longmapsto X_t.$$

Nach Satz 3.10 genügt es zu zeigen:

- 1) F ist stark λ -messbar und 2) $\|F\|$ ist Lebesgue-integrierbar.

Um starke λ -Messbarkeit von F zu zeigen, genügt es nach Lemma 3.4 zu zeigen, dass das Bild von X aufgefasst als Funktion von t fast überall in einem separablen und abgeschlossenen Untervektorraum von $L^2(\Omega, \mathcal{F}, \mathbf{P})$ liegt und dass F messbar ist.

All dies lässt sich aus der Stetigkeit im quadratischen Mittelwert folgern. Gilt nämlich

$$\lim_{\varepsilon \rightarrow 0} \|X_{t+\varepsilon} - X_t\|_{L^2(\Omega, \mathcal{F}, \mathbf{P})} = 0 \quad \text{für alle } t \in \mathcal{T}, \quad (3.19)$$

so gilt wie in Bemerkung 2.7 beschrieben, dass für jedes $t \in \mathcal{T}$ und $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass für jedes $u \in \mathcal{T}$ mit $|u - t| < \delta$ gilt $\|X_u - X_t\|_{L^2(\Omega, \mathcal{F}, \mathbf{P})} < \varepsilon$.

Dies impliziert zwar keine Stetigkeit der Abbildung

$$\begin{aligned}\mathcal{T} &\longrightarrow \mathbb{C} \\ t &\longmapsto X_t(\omega) \quad \text{für } \omega \in \Omega \text{ fixiert,}\end{aligned}$$

aber eben die der Abbildung

$$\begin{aligned}F : \quad \mathcal{T} &\longrightarrow L^2(\Omega, \mathcal{F}, \mathbf{P}, \mathbb{C}) \\ t &\longmapsto X_t\end{aligned}$$

und genau diese wird benötigt, denn Stetigkeit impliziert Messbarkeit.

Das Bild von F ist der von X generierte Hilbertraum \mathcal{H}_X . Nach Lemma 2.17 ist dieser separabel. Somit ist F stark λ -messbar.

Den zweiten Punkt erhält man erneut dank der Stetigkeit von X im q.M.: Es gilt

$$\int_{\mathcal{T}} \|F\| d\lambda = \int_{\mathcal{T}} \mathbf{E}|X_t|^2 \lambda(dt) < \infty, \quad (3.20)$$

da die Abbildung $t \mapsto \mathbf{E}|X_t|^2$ stetig ist und \mathcal{T} kompakt. Die Stetigkeit dieser Abbildung erhält man unmittelbar, da sie eine Verkettung von F und der (stetigen) Normabbildung ist oder indem man feststellt, dass aufgrund der umgekehrten Dreiecksungleichung $\|\|X_u\| - \|X_t\|\| \leq \|X_u - X_t\|$ für alle $u, t \in \mathcal{T}$ gilt.

Eine entscheidene Erkenntnis liefert nun Satz 3.11. Hierbei ist $(\mathcal{T}, \Sigma, \mu) = (\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ und $(\Omega, \mathcal{F}, \nu) = (\Omega, \mathcal{F}, \mathbf{P})$. Außerdem ist $B = \mathbb{C}$ und F wie oben. Da F Bochner-integrierbar ist, existiert eine $(\lambda \otimes \mathbf{P})$ -messbare Funktion $f : \mathcal{T} \times \Omega \rightarrow \mathbb{C}$, die bis auf einer $(\lambda \otimes \mathbf{P})$ -Nullmenge eindeutig bestimmt ist und es gilt $f(t, \cdot) = F(t)$ für fast alle $t \in \mathcal{T}$. Die Abbildung f ist demnach eine Modifikation des Prozesses X . Außerdem ist $f(\cdot, \omega)$ λ -integrierbar auf \mathcal{T} für \mathbf{P} -fast alle $\omega \in \Omega$ und

$$\int_{\mathcal{T}} f(t, \omega) \mu(dt) = \int_{\mathcal{T}} F(t) \mu(dt). \quad (3.21)$$

Resultat:

Ein im quadratischen Mittelwert stetiger stochastischer Prozess besitzt eine $(\lambda \otimes \mathbf{P})$ -messbare Modifikation, ist Lebesgue-integrierbar und das Lebesgue-Integral als Funktion von ω stimmt fast sicher mit dem Bochner-Integral überein.

Kapitel 4

Das Karhunen-Loève-Theorem

Sei $(\Omega, \mathcal{F}, \mathbf{P})$ ein Wahrscheinlichkeitsraum, $\mathcal{T} \subseteq \mathbb{R}$ kompakt und $X = \{X_t : t \in \mathcal{T}\}$ ein zentrierter stochastischer Prozess zweiter Ordnung. Ziel dieses Kapitels ist es, eine gewisse Reihendarstellung für X_t in der Form

$$X_t = \sum_{i=1}^{\infty} Z_i \phi_i(t) \quad \forall t \in \mathcal{T} \tag{4.1}$$

mit Konvergenz in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ zu entwickeln, sodass man den Prozess durch einfachere unkorrelierte Zufallsvariablen $(Z_i)_{i \in \mathbb{N}}$ und orthonormale deterministische Funktionen $(\phi_i)_{i \in \mathbb{N}}$ beschreiben kann. Man spricht daher auch von einem *bi-orthogonal System*, da Unkorreliertheit das wahrscheinlichkeitstheoretische Analogon zur Orthogonalität ist. Im Abschnitt zu Hilberträumen mit reproduzierendem Kern wurde bereits skizziert, dass solche Reihendarstellungen für sämtliche Orthonormalbasen des von der Kovarianzfunktion induzierten RKHS \mathcal{H}_K möglich sind.

Die Besonderheit der Karhunen-Loève-Reihe wird sein, dass die Konvergenz sogar gleichmäßig in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ stattfindet. Außerdem wird sich Optimalität im Bezug auf den mittleren quadratischen Fehler bei abgebrochenen Reihendarstellungen zeigen.

4.1 Aussage und Beweis des Theorems

Das Karhunen-Loève-Theorem wird in der Literatur nicht einheitlich angegeben. Der folgende Satz fasst alle wichtigen Aussagen zusammen und ist in [Wan08] zu finden, allerdings nur im reellwertigen Kontext. Im Wesentlichen gleicht dies der Formulierung von [Loè91], hebt aber den Zusammenhang mit der Fredholm'schen Integralgleichung stärker hervor. Es wird ausdrücklich darauf hingewiesen, dass in beiden Quellen die Integrale der Form $\int X_t dt$ anders aufgefasst werden. Die ursprüngliche Version von Michel Loève verwendet Riemann-Stieltjes-Integrale mit Konvergenz in $L^2(\Omega, \mathcal{F}, \mathbf{P})$, welche allerdings kein Standard in der heutigen Literatur sind. Hier werden diese Integrale als Bochner-Integrale aufgefasst. Aspekte zur Messbarkeit und Existenz wurden im vorherigen Kapitel ausführlich bewiesen.

Satz 4.1 (Karhunen-Loève). Sei $\{X_t : t \in \mathcal{T}\}$ ein zentrierter stochastischer Prozess zweiter Ordnung und stetig im quadratischen Mittelwert. Weiter seien $(\phi_i)_{i \in \mathbb{N}}$ eine Folge orthogonaler Funktionen in $L^2(\mathcal{T}, \lambda)$ und $(\lambda_i)_{i \in \mathbb{N}} \subseteq \mathbb{R}_+$. Dann gelten die zwei Implikationen:

- i) Angenommen λ_i und ϕ_i sind Eigenwerte und Eigenfunktionen des zu K assoziierten Hilbert-Schmidt-Operators und genügen demnach der Fredholm'schen Integralgleichung

$$\int_{\mathcal{T}} K(s, t) \phi_i(t) dt = \lambda_i \phi_i(s) \quad \text{für alle } s \in \mathcal{T}. \quad (4.2)$$

Setzt man

$$\xi_i = \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{T}} X_t \overline{\phi_i(t)} \lambda(dt), \quad (4.3)$$

wobei es sich um ein Bochner-Integral handelt, so sind $(\xi_i)_{i \in \mathbb{N}}$ u.i.v. mit Erwartungswert 0 und Varianz 1 und es gilt

$$X_t = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i \phi_i(t) \quad (4.4)$$

gleichmäßig im quadratischen Mittelwert im Sinne, dass

$$\sup_{t \in \mathcal{T}} \mathbf{E} \left[\left| X_t - \sum_{i=1}^k \sqrt{\lambda_i} \xi_i \phi_i(t) \right|^2 \right] \rightarrow 0 \quad \text{für } k \rightarrow \infty. \quad (4.5)$$

- ii) Ist umgekehrt $X_t = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i \phi_i(t)$ mit Konvergenz in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ und derart, dass $(\xi_i)_{i \in \mathbb{N}}$ paarweise unkorreliert sind mit Erwartungswert 0 und Varianz 1, so gilt die Fredholm'sche Integralgleichung 4.2.

Bemerkung 4.2. i) Die Reihendarstellung von X_t durch $\sum_{i=0}^{\infty} \sqrt{\lambda_i} \xi_i \phi_i(t)$ wird als *Karhunen-Loève-Zerlegung* oder als *Karhunen-Loève-Entwicklung* bezeichnet. In numerischen Anwendungen, in denen die Berechnungen der Koeffizienten mit einer Realisierung der Zufallsvariablen erfolgt, wird der Term *Karhunen-Loève-Transformation* verwendet, im Folgenden abgekürzt durch KLT. Da die Reihe nicht bis ins Unendliche berechnet werden kann und nur bis zu einem festen $k \in \mathbb{N}$ summiert wird, verwendet man häufig den Begriff *abgeschnittene Karhunen-Loève-Transformation*.

- ii) Bemerkenswert ist, dass Satz 4.1 eine Eins-zu-eins-Beziehung zwischen der Fredholm'schen Integralgleichung und der Karhunen-Loève-Zerlegung liefert.
- iii) Falls der Prozess X nicht zentriert ist, so kann man die Karhunen-Loève-Zerlegung von dem Prozess $\hat{X} = \{X_t - \mathbf{E}X_t : t \in \mathcal{T}\}$ bilden und es gilt

$$X_t = \hat{X}_t + \mathbf{E}X_t = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i \phi_i(t) + \mathbf{E}X_t.$$

Der Beweis in [Wan08] geht nicht hinreichend auf die Existenz der Integrale und vor allem auf Vertauschungen der Integrationsreihenfolge ein. In [Loè91] ist der Beweis ausgesprochen kurz gehalten. Hier wird nun eine eigene Version präsentiert, die beide Aspekte vermeidet, sich aber an Ideen beider Beweise bedient.

Beweis. Die Existenz der $(\xi_i)_{i \in \mathbb{N}}$ folgt aus den Überlegungen in Abschnitt 3.2. Da X stetig im q.M. ist und ϕ_i nach Mercers Theorem stetig für jedes $i \in \mathbb{N}$ ist, ist $X \cdot \phi_i$ stetig im q.M.. Zu i): Nach dem Resultat des letzten Kapitels existiert eine Modifikation des Prozesses X , welche bezüglich des Produktmaßes $(\lambda \otimes \mathbf{P})$ -messbar ist. Im Folgenden wird X mit dieser Modifikation identifiziert. Alle folgenden Lebesgue-Integrale sind bezüglich dieser Modifikation zu verstehen. Der Satz von Fubini ist anwendbar, da die betrachteten Maßräume endlich sind und X als komplexwertige Funktion $(\lambda \otimes \mathbf{P})$ -integrierbar ist. Eine einfache Rechnung zeigt damit die gewünschten Eigenschaften der $(\xi_i)_{i \in \mathbb{N}}$:

$$\begin{aligned}\mathbf{E}\xi_i &= \mathbf{E}\left[\frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{T}} X_t \overline{\phi_i(t)} dt\right] \\ &= \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{T}} \mathbf{E}[X_t] \overline{\phi_i(t)} dt \\ &= 0,\end{aligned}$$

da $\mathbf{E}X_t = 0$ für jedes $t \in \mathcal{T}$. Weiter gilt für die Kovarianzen

$$\begin{aligned}\mathbf{E}\xi_i \overline{\xi_j} &= \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_j}} \mathbf{E}\left[\int_{\mathcal{T}} X_s \overline{\phi_i(s)} ds \int_{\mathcal{T}} \overline{X_t \phi_j(t)} dt\right] \\ &= \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_j}} \int_{\mathcal{T}} \int_{\mathcal{T}} \mathbf{E}[X_s \overline{X_t}] \overline{\phi_i(s)} \phi_j(t) ds dt \\ &= \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_j}} \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, t) \phi_j(t) dt \overline{\phi_i(s)} ds.\end{aligned}$$

Nach Voraussetzung gilt die Fredholm'sche Integralgleichung. Daher:

$$\begin{aligned}&= \frac{1}{\sqrt{\lambda_i} \sqrt{\lambda_j}} \int_{\mathcal{T}} \lambda_j \phi_j(s) \overline{\phi_i(s)} ds \\ &= \frac{\lambda_j}{\sqrt{\lambda_i} \sqrt{\lambda_j}} \delta_{ij} \\ &= \delta_{ij}.\end{aligned}$$

Damit ist gezeigt, dass die $(\xi_i)_{i \in \mathbb{N}}$ ein orthonormales System in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ bilden.

Um zu folgern, dass

$$\sup_{t \in \mathcal{T}} \mathbf{E} \left[\left| X_t - \sum_{i=1}^k \sqrt{\lambda_i} \xi_i \phi_i(t) \right|^2 \right] \rightarrow 0 \quad \text{für } k \rightarrow \infty, \tag{4.6}$$

sei $\mathcal{E}_k(t) := \mathbf{E} \left[\left| X_t - \sum_{i=1}^k \sqrt{\lambda_i} \xi_i \phi_i(t) \right|^2 \right]$. Dann gilt

$$\begin{aligned} \mathcal{E}_k(t) &= \mathbf{E} \left[X_t \overline{X}_t \right] - \mathbf{E} \left[X_t \sum_{i=1}^k \sqrt{\lambda_i} \xi_i \phi_i(t) \right] - \mathbf{E} \left[\overline{X}_t \sum_{i=1}^k \sqrt{\lambda_i} \xi_i \phi_i(t) \right] \\ &\quad + \mathbf{E} \left[\sum_{i=1}^k \sum_{j=1}^k \sqrt{\lambda_i} \sqrt{\lambda_j} \xi_i \overline{\xi_j} \phi_i(t) \overline{\phi_j(t)} \right] \\ &= \mathbf{E} \left[X_t \overline{X}_t \right] - \sum_{i=1}^k \sqrt{\lambda_i} \overline{\phi_i(t)} \mathbf{E} \left[X_t \overline{\xi_i} \right] - \sum_{i=1}^k \sqrt{\lambda_i} \phi_i(t) \mathbf{E} \left[\overline{X_t \xi_i} \right] \\ &\quad + \sum_{i=1}^k \sum_{j=1}^k \sqrt{\lambda_i} \sqrt{\lambda_j} \phi_i(t) \overline{\phi_j(t)} \mathbf{E} \xi_i \overline{\xi_j}. \end{aligned}$$

Obige Rechnung hat bereits gezeigt, dass $\mathbf{E} \xi_i \overline{\xi_j} = \delta_{ij}$. Betrachte den Term $\mathbf{E} X_t \overline{\xi_i}$:

$$\begin{aligned} \mathbf{E} X_t \overline{\xi_i} &= \frac{1}{\sqrt{\lambda_i}} \mathbf{E} \left[X_t \int_{\mathcal{T}} X_s \overline{\phi_i(s)} ds \right] \\ &= \frac{1}{\sqrt{\lambda_i}} \mathbf{E} \int_T X_t \overline{X}_s \phi_i(s) ds \\ &= \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{T}} \mathbf{E} \left[X_t \overline{X}_s \right] \phi_i(s) ds \\ &= \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{T}} K(t, s) \phi_i(s) ds \\ &= \frac{1}{\sqrt{\lambda_i}} \lambda_i \phi_i(t) \\ &= \sqrt{\lambda_i} \phi_i(t), \end{aligned}$$

wobei erneut der Satz von Fubini und die Fredholm'sche Integralgleichung genutzt wurden. Zusammen ergibt sich für $\mathcal{E}_k(t)$:

$$\begin{aligned} \mathcal{E}_k(t) &= \mathbf{E} \left[X_t \overline{X}_t \right] - 2 \sum_{i=1}^k \lambda_i \phi_i(t) \overline{\phi_i(t)} + \sum_{i=1}^k \lambda_i \phi_i(t) \overline{\phi_i(t)} \\ &= K(t, t) - \sum_{i=1}^k \lambda_i \phi_i(t) \overline{\phi_i(t)}. \end{aligned}$$

Nach Mercers Theorem gilt $\sum_{i=1}^k \lambda_i |\phi_i(t)|^2 \rightarrow K(t, t)$ gleichmäßig und somit

$$\sup_{t \in \mathcal{T}} \mathcal{E}_k(t) = \sup_{t \in \mathcal{T}} \left| K(t, t) - \sum_{i=1}^k \lambda_i \phi_i(t) \overline{\phi_i(t)} \right| \rightarrow 0. \quad (4.7)$$

Damit ist die gleichmäßige Konvergenz im quadratischen Mittelwert für die Karhunen-Loève-Reihe gezeigt.

Zu ii): Es ist zu zeigen, dass die Fredholm'sche Integralgleichung gilt.

Sei dazu $X_t^{(N)} = \sum_{i=1}^N \sqrt{\lambda_i} \xi_i \phi_i(t)$. Dann gilt

$$\mathbf{E} X_s \overline{X}_t \stackrel{\text{Lem.2.5}}{=} \lim_{N \rightarrow \infty} \mathbf{E} X_s^{(N)} \overline{X}_t^{(N)} = \lim_{N \rightarrow \infty} \sum_{i,j=1}^N \sqrt{\lambda_i \lambda_j} \phi_i(s) \overline{\phi_j(t)} \mathbf{E} \xi_i \overline{\xi}_j = \sum_{j=1}^{\infty} \lambda_j \phi_j(s) \overline{\phi_j(t)}.$$

Setzt man dies in die linke Seite der Fredholm'schen Integralgleichung ein, so erhält man

$$\begin{aligned} \int_{\mathcal{T}} K(s, t) \phi_i(t) dt &= \int_{\mathcal{T}} \mathbf{E} [X_s \overline{X}_t] \phi_i(t) dt \\ &= \int_{\mathcal{T}} \sum_{j=1}^{\infty} \lambda_j \phi_j(s) \overline{\phi_j(t)} \phi_i(t) dt \\ &= \sum_{j=1}^{\infty} \lambda_j \phi_j(s) \int_{\mathcal{T}} \overline{\phi_j(t)} \phi_i(t) dt \\ &= \lambda_i \phi_i(s) \end{aligned}$$

und das Karhunen-Loève-Theorem ist bewiesen. \square

4.2 Karhunen-Loève-Zerlegung von Gauß-Prozessen

Ist der zugrundeliegende Prozess ein Gauß-Prozess, so lassen sich besondere Resultate erzielen. Zum Einen kann eine konkrete Aussage über die Verteilung der $(\xi_i)_{i \in \mathbb{N}}$ getroffen werden und zum Anderen kann gezeigt werden, dass diese Zufallsvariablen unabhängig und identisch verteilt sind. Besonders bei Simulationen im Computer rechtfertigt die Unabhängigkeit einen einfachen Weg der Simulation.

Lemma 4.3. *Sei X ein stochastischer Prozess zweiter Ordnung, zentriert und stetig im quadratischen Mittelwert. Ist X ein Gauß-Prozess, so sind die ξ_i in der Karhunen-Loève-Reihe von X gemeinsam standard-normalverteilt und unabhängig.*

Beweis. Der Beweis wurde selbstständig entwickelt. Es genügt zu zeigen, dass $\xi = (\xi_i)_{i \in \mathbb{N}}$ ebenfalls ein Gauß-Prozess indiziert über \mathbb{N} ist. Seien dazu $d \in \mathbb{N}$, $a \in \mathbb{R}^d$ und $i_1, \dots, i_d \in \mathbb{N}$. Sei Π der Projektionsoperator. Es gilt für $\Pi_{i_1, \dots, i_d}^{\mathbb{N}}(\xi) = (\xi_{i_1}, \dots, \xi_{i_d})$:

$$\langle a, \Pi_{i_1, \dots, i_d}^{\mathbb{N}}(\xi) \rangle = \sum_{k=1}^d a_k \xi_{i_k} = \sum_{k=1}^d a_k \frac{1}{\sqrt{\lambda_{i_k}}} \int_{\mathcal{T}} X_t \phi_{i_k}(t) \lambda(dt).$$

Bei den Integralen handelt es sich um Bochner-Integrale, welche existieren, da $X \phi_{i_k}$ stetig im quadratischen Mittelwert ist. Da $X \phi_{i_k}$ Bochner-integrierbar ist, existiert eine Folge $(\varphi_n)_{n \in \mathbb{N}} \subseteq L_{L^2(\Omega, \mathcal{F}, \mathbf{P})}$, sodass die reellen messbaren Funktionen $\|\phi_{i_k} - \varphi_n\|$ Lebesgue-integrierbar sind für jedes $n \in \mathbb{N}$ und $\lim_{n \rightarrow \infty} \int \|X \phi_{i_k} - \varphi_n\| d\lambda = 0$. Es ist per Definition

$$\int X_t \phi_{i_k}(t) \lambda(dt) = \lim_{n \rightarrow \infty} \int \varphi_n(t) \lambda(dt). \quad (4.8)$$

Wegen der Stetigkeit der Abbildung $\mathcal{T} \ni t \mapsto X_t \in L^2(\Omega, \mathcal{F}, \mathbf{P})$ kann man

$$\varphi_n = \sum_{l=1}^{N^{(n)}} X_{t_l^{(n)}} \phi_{i_k}(t_l^{(n)}) \mathbb{1}_{\{[t_l^{(n)}, t_{l+1}^{(n)}]\}} \quad (4.9)$$

für eine Folge von feiner werdenden Partitionen $\nabla_n = t_1^{(n)}, \dots, t_{N^{(n)}+1}^{(n)}$ von \mathcal{T} wählen. Dann ist nach Definition des Bochner-Integrals

$$\int_T X_t \phi_{i_k}(t) \lambda(dt) = \lim_{n \rightarrow \infty} \sum_{l=1}^{N^{(n)}} X_{t_l^{(n)}} \phi_{i_k}(t_l^{(n)}) (t_{l+1}^{(n)} - t_l^{(n)}) , \quad (4.10)$$

mit Konvergenz in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ und damit

$$\langle a, \Pi_{i_1, \dots, i_d}^{\mathbb{N}}(\xi) \rangle = \lim_{n \rightarrow \infty} \underbrace{\sum_{k=1}^d a_k \frac{1}{\sqrt{\lambda_{i_k}}} \sum_{l=1}^{N^{(n)}} X_{t_l^{(n)}} \phi_{i_k}(t_l^{(n)}) (t_{l+1}^{(n)} - t_l^{(n)})}_{:= S_n} .$$

Da nach Voraussetzung X ein Gauß-Prozess ist, lässt sich für jedes $n \in \mathbb{N}$ folgern, dass $S_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$ für gewisse $\mu_n, \sigma_n^2 \in \mathbb{R}$. Dies ist leicht zu sehen, wenn man die folgende Umformung vornimmt:

$$S_n = \sum_{l=1}^{N^{(n)}} X_{t_l^{(n)}} \sum_{k=1}^d a_k \frac{1}{\sqrt{\lambda_{i_k}}} \phi_{i_k}(t_l^{(n)}) (t_{l+1}^{(n)} - t_l^{(n)}) = \langle c, \Pi_{\nabla_n}^{\mathcal{T}}(X) \rangle$$

wobei $c \in \mathbb{R}^d$ mit $c_k := \sum_{l=1}^d a_k \frac{1}{\sqrt{\lambda_{i_k}}} \phi_{i_k}(t_l^{(n)}) (t_{l+1}^{(n)} - t_l^{(n)})$.

Ganz allgemein gilt aber, dass der Grenzwert in Verteilung von normalverteilten Zufallsvariablen wieder normalverteilt ist.

Demnach ist $\langle a, \Pi_{i_1, \dots, i_d}^{\mathbb{N}}(\xi) \rangle$ normalverteilt und damit ξ ein Gauß-Prozess. \square

Für die Karhunen-Loève-Reihen von Gauß-Prozessen folgen sogar noch weitere sehr starke Eigenschaften. Nach [Adl10] gilt Folgendes:

Satz 4.4 (Thm. 3.3.3 [Adl10]). *Sei $\{X_t : t \in \mathcal{T}\}$ ein zentriertes Gauß-Zufallsfeld mit stetiger Kovarianzfunktion. Dann sind die Pfade des Feldes fast sicher stetig. Außerdem konvergiert die Karhunen-Loève-Reihe nicht nur gleichmäßig in $L^2(\Omega, \mathcal{F}, \mathbf{P})$, sondern zusätzlich fast sicher.*

Beweis. Für einen Beweis wird auf die angegebene Quelle verwiesen. \square

Bemerkung 4.5. Die soeben gezeigten Eigenschaften rechtfertigen eine numerische Simulation von Gauß-Prozessen durch das Sampeln von unabhängigen standard-normalverteilten Zufallsvariablen und dem Berechnen von großen Partialsummen der Karhunen-Loève-Reihe.

Bemerkung 4.6. Man könnte denken, dass eine Reihe – wie die der Karhunen-Loève-Zerlegung – stets für fast alle ω für jede Orthonormalbasis $(\varphi_i)_{i \in \mathbb{N}}$ von $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ konvergiert. Im Kapitel 3 wurde gezeigt, dass die Pfade eines im quadratischen Mittelwert stetigen Prozesses in $L^1(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ liegen, nicht aber im Hilbertraum L^2 . Die Basis aus Eigenfunktionen ist aber eine Orthonormalbasis von $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$, weshalb die Reihe $\sum_{i=1}^{\infty} \langle X(\cdot)(\omega), \varphi_i \rangle_{L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)} \varphi_i(t)$ nicht in L^2 konvergieren muss und insbesondere nicht einmal sinnvoll definiert ist, da das Skalarprodukt nicht existieren muss. Wäre dies dennoch der Fall, so würde die Karhunen-Loève-Reihe fast sicher in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ konvergieren. Dieses Resultat ist nur für Gauß-Prozesse bekannt. Dazu sei erneut auf Theorem 3.3.3 in [Adl10] verwiesen.

4.3 Vorteile der Karhunen-Loève-Zerlegung

Die Reihendarstellung hat den großen Vorteil, dass die Abhängigkeiten des zugrunde liegenden Prozesses X von ω und von t isoliert werden. Verdeutlicht:

$$X(t, \omega) = \sum_{i=1}^{\infty} \underbrace{\sqrt{\lambda_i}}_{\text{Konstante}} \underbrace{\xi_i(\omega)}_{\text{Funktion von } \omega} \underbrace{\phi_i(t)}_{\text{Funktion von } t}.$$

Dabei sollte man die ξ_i nicht einfach als weiteren stochastischen Prozess ansehen, sondern als eine von der Zeit unabhängige Familie von Zufallsvariablen. Betrachtet man beispielsweise ein deterministisches Signal $s(t)$, welches durch ein zufälliges Geräusch G_t gestört wird, so liefern Messungen einen Prozess $X = \{X_t = s(t) + G_t : t \in \mathcal{T}\}$, der zu jedem Zeitpunkt die Summe aus der Frequenz des Signals und der des Geräusches verkörpert. Nun lässt sich aber nicht ohne weiteres einfach das ursprüngliche Signal berechnen, da das Geräusch unvorhersehbar ist und sich beide Größen in $X(t)$ vermischen. KLT erweist sich hier als mächtiges Werkzeug, um rückwirkend das Signal zu ermitteln. Siehe dazu Abschnitt 6.2.1.

Die Karhunen-Loève-Zerlegung minimiert den totalen mittleren quadratischen Fehler unter allen Reihenentwicklungen des Prozesses in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ bezüglich einer Orthonormalbasis des zur Kovarianzfunktion gehörigen Hilbertraumes mit reproduzierendem Kern. Da diese Eigenschaft sehr komplex ist, findet man in dieser Arbeit das Kapitel 5 zur Optimalität bezüglich des mittleren quadratischen Fehlers.

Ein weiterer großer Vorteil der Karhunen-Loève-Zerlegung ist, dass man durch die neue Darstellung des Prozesses gewisse Eigenschaften direkt folgern kann. Dies macht sich besonders bemerkbar, wenn man die Varianz betrachtet. Zunächst gilt für jedes $t \in \mathcal{T}$:

$$\text{Var}(X_t) = \text{Var} \left(\sum_{i=0}^{\infty} \sqrt{\lambda_i} \xi_i \phi_i(t) \right).$$

Da die $(\xi_i)_{i \in \mathbb{N}}$ orthogonal in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ liegen, sind sie insbesondere unkorreliert. Nach der Gleichung von Bienaymé ist damit die Varianz der Summe die Summe der Varianzen. Nutzt

man außerdem die Eigenschaft, dass für alle $i \in \mathbb{N}$ stets $\text{Var}(\xi_i) = 1$, so ergibt sich

$$\text{Var}(X_t) = \sum_{i=0}^{\infty} \lambda_i |\phi_i(t)|^2. \quad (4.11)$$

Darüberhinaus kann auch die totale Varianz sehr effizient angegeben werden. Hierbei nutzt man die Orthonormalität der $(\phi_i)_{i \in \mathbb{N}}$ in $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ und erhält damit

$$\text{Var}(X) = \int_{\mathcal{T}} \text{Var}(X_t) dt = \sum_{i=0}^{\infty} \lambda_i \int_{\mathcal{T}} \phi_i(t) \overline{\phi_i(t)} dt = \sum_{i=0}^{\infty} \lambda_i. \quad (4.12)$$

Dabei kann die Integrationsreihenfolge wegen monotoner Konvergenz vertauscht werden.

4.4 Karhunen-Loève-Entwicklung von Zufallsfeldern

Das Karhunen-Loève-Theorem kann auch für Prozesse indiziert über mehrere Dimensionen genutzt werden. Das bedeutet, dass \mathcal{T} ein Produkt aus kompakten Intervallen ist, der Prozess aber weiterhin Werte in \mathbb{R} oder \mathbb{C} annimmt. Der Beweis kann analog zu dem ein-dimensionalen Fall geführt werden. Ist $d \in \mathbb{N}$, $\mathcal{T} = \mathcal{T}_1 \times \mathcal{T}_2 \times \dots \times \mathcal{T}_d$ und X ein d -dimensionales Zufallsfeld zweiter Ordnung mit stetiger Kovarianzfunktion

$$K(\mathbf{s}, \mathbf{t}) = \text{Cov}[X(s_1, s_2, \dots, s_d), X(t_1, t_2, \dots, t_d)] \quad \text{für } \mathbf{s}, \mathbf{t} \in \mathcal{T}, \quad (4.13)$$

steht man vor der d -dimensionalen Fredholm'schen Integralgleichung $\int_{\mathcal{T}} K(\mathbf{s}, \mathbf{t}) \phi(\mathbf{s}) d\mathbf{t} = \lambda \phi(\mathbf{s})$, oder äquivalent

$$\begin{aligned} \int_{\mathcal{T}_1} \int_{\mathcal{T}_2} \dots \int_{\mathcal{T}_d} \text{Cov}[X(s_1, s_2, \dots, s_d), X(t_1, t_2, \dots, t_d)] \phi(t_1, t_2, \dots, t_d) dt_d \dots dt_2 dt_1 \\ = \lambda \phi(s_1, s_2, \dots, s_d). \end{aligned}$$

Deren Lösungen seien die Eigenfunktionen $(\phi_i)_{i \in \mathbb{N}} \subseteq L^2(\mathcal{T}^d, \mathcal{B}(\mathcal{T})^{\otimes d}, \lambda^{\otimes d})$ und Eigenwerte $(\lambda_i)_{i \in \mathbb{N}} \subseteq \mathbb{R}$. Dann ist

$$X(t_1, t_2, \dots, t_d) = \sum_{i=0}^{\infty} \sqrt{\lambda_i} \xi_i \phi_i(t_1, t_2, \dots, t_d), \quad (4.14)$$

mit gleichmäßiger Konvergenz in $L^2(\Omega, \mathcal{F}, \mathbf{P})$.

Eine Lösung für die d -dimensionale Fredholm'sche Integralrechnung zu finden, ist analytisch sehr schwierig. Schon im Eindimensionalen ist dies nur in den wenigsten Fällen möglich. Eine Ausnahme bietet die folgende Klasse von Kovarianzfunktionen:

Definition 4.7. Sei K Kovarianzfunktion eines d -dimensionalen zentrierten Zufallsfeldes X zweiter Ordnung. K heißt *separabel*, falls $K(\mathbf{s}, \mathbf{t}) = \prod_{i=1}^d K^{(i)}(s_i, t_i)$ für gewisse $K^{(i)}$, die selbst Kovarianzfunktionen von stochastischen Prozessen mit nur einer Zeitdimension sind.

Satz 4.8 (Thm 1.4.1. [Wan08]). Seien X, \mathcal{T} und K wie bisher. Sei K separabel und seien $(\lambda_k^{(i)})_{k \in \mathbb{N}}$ und $(\phi_k^{(i)})_{k \in \mathbb{N}}$ die Eigenwerte und Eigenfunktionen der Integraloperatoren zu den eindimensionalen Kovarianzfunktionen $K^{(i)}$ für $i = 1, \dots, d$. Dann gilt

$$X_{\mathbf{t}} = \sum_{\substack{n \in \mathbb{N}_0^d \\ |n| \geq 1}} \sqrt{\lambda_n} \xi_n \phi_n(\mathbf{t}), \quad (4.15)$$

wobei die Konvergenz gleichmäßig in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ stattfindet und

$$\phi_n(\mathbf{t}) = \prod_{i=1}^d \phi_{n_i}^{(i)}(t_i) \text{ und } \lambda_n = \prod_{i=1}^d \lambda_{n_i}^{(i)} \text{ für alle } n \in \mathbb{N}^d. \quad (4.16)$$

Beweis. Der Beweis richtet sich nach [Wan08, Thm. 1.4.1]. Es muss überprüft werden, dass für jedes $n \in \mathbb{N}$ das oben definierte ϕ_n eine Eigenfunktion des Integraloperators zum Eigenwert λ_n ist. Setzt man ϕ_n in die linke Seite der d -dimensionalen Fredholm'schen Integralgleichung ein, erhält man

$$\begin{aligned} & \int_{\mathcal{T}_1} \int_{\mathcal{T}_2} \dots \int_{\mathcal{T}_d} \text{Cov}[X(s_1, s_2, \dots, s_d), X(t_1, t_2, \dots, t_d)] \phi_n(t_1, t_2, \dots, t_d) dt_d \dots dt_2 dt_1 \\ &= \int_{\mathcal{T}_1} \int_{\mathcal{T}_2} \dots \int_{\mathcal{T}_d} \prod_{i=1}^d K^{(i)}(s_i, t_i) \prod_{i=1}^d \phi_{n_i}^{(i)}(t_i) dt_d \dots dt_2 dt_1 \\ &= \prod_{i=1}^d \int_{\mathcal{T}_i} K^{(i)}(s_i, t_i) \phi_{n_i}^{(i)}(t_i) dt_i \\ &= \prod_{i=1}^d \lambda_{n_i}^{(i)} \phi_{n_i}^{(i)}(s_i) \\ &= \prod_{i=1}^d \lambda_{n_i}^{(i)} \prod_{i=1}^d \phi_{n_i}^{(i)}(t_i) \\ &= \lambda_n \phi_n(\mathbf{s}). \end{aligned}$$

Zur Orthonormalität der (ϕ_n) : Seien $n, m \in \mathbb{N}_0^d$ mit $|n|, |m| \geq 1$. Dann gilt

$$\begin{aligned} \langle \phi_n, \phi_m \rangle &= \int_{\mathcal{T}_1} \int_{\mathcal{T}_2} \dots \int_{\mathcal{T}_d} \phi_n(t_1, t_2, \dots, t_d) \phi_m(t_1, t_2, \dots, t_d) dt_d \dots dt_2 dt_1 \\ &= \int_{\mathcal{T}_1} \int_{\mathcal{T}_2} \dots \int_{\mathcal{T}_d} \prod_{i=1}^d \phi_{n_i}^{(i)}(t_i) \prod_{i=1}^d \phi_{m_i}^{(i)}(t_i) dt_d \dots dt_2 dt_1 \\ &= \prod_{i=1}^d \int_{\mathcal{T}_i} \phi_{n_i}^{(i)}(t_i) \phi_{m_i}^{(i)}(t_i) dt_i \\ &= \prod_{i=1}^d \delta_{n_i, m_i} \\ &= \delta_{n, m}. \end{aligned}$$

Damit konvergiert nach dem Karhunen-Loève-Theorem die obige Reihe gleichmäßig. \square

Bemerkung 4.9. In [Wan08, Kapitel 5] wird außerdem gezeigt, dass man die Karhunen-Loève-Zerlegung auf Prozesse von \mathbb{R}^d -wertigen Zufallsvariablen erweitern kann.

4.5 Beispiele zur Berechnung der Zerlegung

4.5.1 Die Brown'sche Bewegung und die Brown'sche Brücke

Der Einfachheit halber sei nun stets $\mathcal{T} = [0, 1]$.

Beispiel 4.10. Für die Brown'sche Bewegung gilt $K(s, t) = s \wedge t$. Damit hat die Fredholm'sche Integralgleichung die Form

$$\int_0^1 (s \wedge t) \phi(s) ds = \lambda \phi(t) \quad \text{für alle } t \in [0, 1] \quad (4.17)$$

$$(\Leftrightarrow) \quad \int_0^t s \phi(s) ds + t \int_t^1 \phi(s) ds = \lambda \phi(t) \quad \text{für alle } t \in [0, 1]. \quad (4.18)$$

Differentiation liefert die Gleichung

$$\int_t^1 s \phi(s) ds = \lambda \phi'(t) \quad \text{für alle } t \in [0, 1], \quad (4.19)$$

welche nach erneutem Differenzieren die gewöhnliche Differentialgleichung

$$-\phi(t) = \lambda \phi''(t) \quad \text{für alle } t \in [0, 1] \quad (4.20)$$

ergibt. Deren allgemeine Lösung ist von der Form $\phi(t) = A \sin\left(\frac{t}{\sqrt{\lambda}}\right) + B \cos\left(\frac{t}{\sqrt{\lambda}}\right)$, für gewisse Konstanten $A, B \in \mathbb{R}$. Aus Gleichung (4.18) folgt außerdem die Nebenbedingung $\phi(0) = 0$. Damit folgt $B = 0$. Für $t = 1$ folgt aus der Gleichung nach dem ersten Differenzieren $\phi'(1) = 0$ und damit

$$\cos\left(\frac{1}{\sqrt{\lambda}}\right) = 0 \quad \Leftrightarrow \quad \lambda = \frac{4}{(2i-1)^2 \pi^2} \quad \text{für ein } i \in \mathbb{N}. \quad (4.21)$$

Somit sind bereits die Eigenwerte gefunden. Außerdem gilt

$$1 = \int_0^1 \phi_i^2(t) dt = A^2 \int_0^1 \sin^2\left(\frac{t}{\sqrt{\lambda_i}}\right) dt = A^2 \int_0^1 \sin^2\left[\left(i - \frac{1}{2}\right) \pi t\right] dt = \frac{A^2}{2}.$$

Es ist demnach $A = \sqrt{2}$ und die Karhunen-Loève-Entwicklung hat die Gestalt:

$$B_t = \sum_{i=1}^{\infty} \frac{2}{(2i-1)\pi} \sqrt{2} \sin\left[\left(i - \frac{1}{2}\right) \pi t\right] \xi_i. \quad (4.22)$$

Beispiel 4.11. Die Brown'sche Brücke ist ein zentrierter Gauß-Prozess $X = \{X_t : t \in [0, 1]\}$ mit $K(s, t) = s \wedge t - st$. Ist B eine Brown'sche Bewegung, so ist ein Prozess definiert durch $X_t := B_t - tB_1$ eine Brown'sche Brücke. Auf sehr ähnliche Weise, wie beim vorherigen Beispiel,

kann man die Karhunen-Loève-Entwicklung der Brwonschen Brücke finden. Es gilt

$$X_t = \sum_{i=1}^{\infty} \frac{1}{i\pi} \sqrt{2} \sin(i\pi t) \xi_i. \quad (4.23)$$

Die Rechnung soll hier aufgrund der großen Ähnlichkeit nicht erscheinen. Siehe dazu [Wan08, Absch. 1.3, Bsp 2]. In Abbildung 2 sind die Eigenfunktionen und Eigenwerte zu diesen Beispielen dargestellt. Abbildung 3 zeigt verschiedene Partialsummen der Karhunen-Loève-Zerlegung.

4.5.2 Das Brown'sche Blatt

Beispiel 4.12. Das Brown'sche Blatt Z ist ein zentriertes Gauß'sches Zufallsfeld mit Kovarianzfunktion

$$K(\mathbf{s}, \mathbf{t}) = (s_1 \wedge t_1)(s_2 \wedge t_2). \quad (4.24)$$

Diese Kovarianzfunktion ist separabel und die Faktorkovarianzfunktionen sind die einer Brown'schen Bewegung. Zusammen mit Satz (4.8) und dem Beispiel zur Brown'schen Bewegung folgt:

$$Z_{\mathbf{t}} = 2 \sum_{\substack{n \in \mathbb{N}_0^d \\ |n| \geq 1}} \frac{4}{(2n_1 - 1)(2n_2 - 1)\pi^2} \sin \left[\left(n_1 - \frac{1}{2} \right) \pi t_1 \right] \sin \left[\left(n_2 - \frac{1}{2} \right) \pi t_2 \right]. \quad (4.25)$$

Abbildung 7 zeigt einen Pfad dieses Prozesses simuliert durch die Karhunen-Loève-Reihe.

Kapitel 5

Optimalität der Karhunen-Loève-Zerlegung

Nach dem Karhunen-Loève-Theorem kann ein stochastischer Prozess $X = \{X_t : t \in \mathcal{T}\}$ zweiter Ordnung mit stetiger Kovarianzfunktion K in der Darstellung $X_t = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i \phi_i(t)$ mit gleichmäßiger Konvergenz in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ angegeben werden. Eine solche Reihendarstellung kann auch bezüglich anderer Orthonormalbasen $(\varphi_i)_{i \in \mathbb{N}}$ des RKHS \mathcal{H}_K entwickelt werden, wenn man auf die Bedingung der Gleichmäßigkeit verzichtet, siehe dazu erneut Satz 2.18. In diesem Kapitel wird angenommen, dass $(\varphi_i)_{i \in \mathbb{N}}$ eine Orthonormalbasis von $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$ ist, für die gilt

$$X_t = \sum_{i=1}^{\infty} \left[\int_{\mathcal{T}} X_t \overline{\varphi_i(t)} \lambda(dt) \right] \varphi_i(t), \quad (5.1)$$

wobei die Konvergenz der Reihe in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ stattfindet und die Integrale als Bochner-Integrale zu verstehen sind. Es soll gezeigt werden, dass die Basis aus Eigenfunktionen der Karhunen-Loève-Zerlegung eine gewisse Optimalität unter all solchen Basen besitzt. Dazu zunächst folgende Definition:

Definition 5.1. Sei $X \in L^2(\Omega, \mathcal{F}, \mathbf{P})$ und $\hat{X} \in L^2(\Omega, \mathcal{F}, \mathbf{P})$ eine Approximation von X . Der L^2 -Abstand der beiden Zufallsvariablen wird als *mittlere quadratische Abweichung* bezeichnet und mit MSE notiert. Es ist

$$\text{MSE} := \mathbf{E}|X - \hat{X}|^2. \quad (5.2)$$

Sei nun $X = \{X_t : t \in \mathcal{T}\}$ ein stochastischer Prozess und $\hat{X} = \{\hat{X}_t : t \in \mathcal{T}\}$ eine Approximation von X . Ein Maß für die Abweichung der Approximation ist die *totale mittlere quadratische Abweichung*

$$\text{TMSE} = \int_{\mathcal{T}} \mathbf{E}|X_t - \hat{X}_t|^2 dt = \mathbf{E} \left[\int_{\mathcal{T}} |X_t - \hat{X}_t|^2 dt \right]. \quad (5.3)$$

TMSE wird auch als integrierte mittlere quadratische Abweichung bezeichnet.

Zentrales Element dieses Kapitels ist der folgende Satz. Er findet sich mit der gleichen Aussage in [Wan08, Thm. 1.2.2], wird hier jedoch anders formuliert.

Satz 5.2. Sei $(\phi_i)_{i \in \mathbb{N}}$ die Basis aus Eigenfunktionen des von der Kovarianzfunktion induzierten Hilbert-Schmidt-Operators, absteigend sortiert nach der Größe der zugehörigen Eigenwerte. Außerdem sei $(\varphi_i)_{i \in \mathbb{N}}$ eine weitere Basis von $L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$, für die die Reihe 5.1 in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ konvergiert. Für jedes $N \in \mathbb{N}$ sei $TMSE_N((\varphi_i)_{i \in \mathbb{N}})$ definiert als die totale mittlere quadratische Abweichung von X und der N -ten Partialsumme der Reihe 5.1. Dann gilt für jedes $N \in \mathbb{N}$:

$$TMSE_N((\varphi_i)_{i \in \mathbb{N}}) \geq TMSE_N((\phi_i)_{i \in \mathbb{N}}). \quad (5.4)$$

Die Karhunen-Loève Basis minimiert demnach die totale mittlere quadratische Abweichung bei Approximation eines Prozesses mit Reihen der Form aus 5.1.

Ein Beweis wird in Abschnitt 5.2 präsentiert. Zuvor soll der Satz numerisch bestätigt werden.

5.1 Veranschaulichung und empirische Bestätigung

In einem MATLAB-Programm, dass zu dieser Arbeit implementiert wurde, wird eine Brown'sche Bewegung auf $[0, 1]$ bezüglich verschiedener Orthonormalbasen entwickelt, um den totalen mittleren quadratischen Fehler TMSE zu berechnen. Folgende Basen wurden verwendet:

- 1) Die analytisch bestimmten Eigenfunktionen des Integraloperators der Brown'schen Bewegung.
- 2) Numerische Approximationen der Eigenfunktionen des Integraloperators der Brown'schen Brücke.
- 3) Die Haar-Wavelets. Diese Funktionen werden im Abschnitt 6.1.3 näher beschrieben.
- 4) Numerische Approximationen der Eigenfunktionen des Integraloperators des exponentiellen Kernes $K(s, t) = \exp(-|s - t|)$.

Das Programm kann leicht an andere Basen angepasst werden. Die Berechnung von $TMSE = \mathbf{E} \left[\int_{\mathcal{T}} |X_t - \hat{X}_t|^2 dt \right]$ erfolgt mittels der Monte-Carlo-Methode. Dabei wurden die Integrale mit der Trapezregel approximiert. Mehr dazu wird in Kapitel 6 beschrieben. Die Ergebnisse, für die mehrere Minuten Rechenaufwand benötigt wurden, zeigt die folgende Tabelle.

Tabelle 5.1: Approximation des TMSE

Basis	Br. Bewegung	Br. Brücke	Haar Wavelets	Exp. Kern
TMSE	0.0051	0.0131	0.0090	0.0054

Mit 10.000 Sampels für die jeweils 30. Partialsumme der Karhunen-Loève-Zerlegung

Die numerischen Berechnungen stützen die Behauptung des Satzes. Bemerkenswert ist, dass für die Basis des exponentiellen Kernes ein Ergebnis erzielt wurde, welches sehr nahe an dem Optimum liegt. Abbildung 4 zeigt einen Pfad einer Brown'schen Bewegung und die zu den verschiedenen Basen gehörigen Partialsummen mit den ersten 30 Summanden.

5.2 Analytischer Beweis der Optimalität bezüglich des mittleren quadratischen Fehlers

Einen Beweis zu Satz 5.2 findet man in [Wan08, Thm. 1.2.2]. Dieser zeigt allerdings nur, dass es notwendig ist, für ein Minimum von TMSE_N bezüglich der Karhunen-Loève-Basis entwickelt zu haben. Es wird aber nicht hinreichend argumentiert, dass tatsächlich für diese Basis ein Minimum vorliegt. Der hier gegebene Beweis ist selbst entwickelt und verfolgt einen neuen Ansatz zur Lösung dieses Problems. Die Minimierung von TMSE_N wird auf eine Maximierung eines Doppelintegrals zurückgeführt, mit dem sich schon der bedeutende deutsche Mathematiker David Hilbert beschäftigte.

Beweis. Seien $\text{MSE}_N^{(t)}$ und TMSE_N die mittlere bzw. die totale mittlere quadratische Abweichung von X_t bzw. X und der N -ten Partialsumme. Es gilt für $t \in \mathcal{T}$

$$\begin{aligned} \text{MSE}_N^{(t)} ((\varphi_i)_{i \in \mathbb{N}}) &:= \mathbf{E}(X_t - \hat{X}_t)(\overline{X_t - \hat{X}_t}) \\ &= \mathbf{E} \left(X_t - \sum_{i=1}^N \int_{\mathcal{T}} X_s \overline{\varphi_i(s)} ds \varphi_i(t) \right) \left(\overline{X_t} - \sum_{i=1}^N \int_{\mathcal{T}} \overline{X_s} \varphi_i(s) ds \overline{\varphi_i(t)} \right). \end{aligned}$$

Da die Integrale als Bochner-Integrale existieren, folgt mit dem Resultat aus Kapitel 3 ($\lambda \otimes \mathbf{P}$)-Messbarkeit der Integranden, welche in Kombination mit dem Satz von Fubini im Folgenden die Vertauschung von Erwartungswerten und Integralen rechtfertigt:

$$\begin{aligned} \text{MSE}_N^{(t)} ((\varphi_i)_{i \in \mathbb{N}}) &= \mathbf{E}|X_t|^2 + \mathbf{E} \left[\sum_{i=1}^N \sum_{j=1}^N \int_{\mathcal{T}} X_s \overline{\varphi_i(s)} ds \int_{\mathcal{T}} \overline{X_u} \varphi_j(u) du \varphi_i(t) \overline{\varphi_j(t)} \right] \\ &\quad - \mathbf{E} \left[\overline{X_t} \sum_{i=1}^N \int_{\mathcal{T}} X_s \overline{\varphi_i(s)} ds \varphi_i(t) \right] - \mathbf{E} \left[X_t \sum_{i=1}^N \int_{\mathcal{T}} \overline{X_s} \varphi_i(s) ds \overline{\varphi_i(t)} \right] \\ &= \mathbf{E}|X_t|^2 + \sum_{i=1}^N \sum_{j=1}^N \int_{\mathcal{T}} \int_{\mathcal{T}} \mathbf{E} \left[X_s \overline{X_u} \right] \varphi_j(u) \overline{\varphi_i(s)} ds du \varphi_i(t) \overline{\varphi_j(t)} \\ &\quad - \sum_{i=1}^N \int_{\mathcal{T}} \mathbf{E} \left[X_s \overline{X_t} \right] \overline{\varphi_i(s)} ds \varphi_i(t) - \sum_{i=1}^N \int_{\mathcal{T}} \mathbf{E} \left[X_t \overline{X_s} \right] \varphi_i(s) ds \overline{\varphi_i(t)}. \end{aligned}$$

Um TMSE_N zu erhalten, muss obige Formel auf beiden Seiten über \mathcal{T} integriert werden. Hier kommt die Orthonormalität der $(\varphi_i)_{i \in \mathbb{N}}$ ins Spiel. Außerdem kann die Definition der Kovarianzfunktion und deren Hermitizität genutzt werden. Man erhält

$$\begin{aligned} \text{TMSE}_N &= \int_{\mathcal{T}} \mathbf{E}|X_t|^2 dt + \sum_{i=1}^N \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, u) \varphi_i(u) \overline{\varphi_i(s)} ds du - 2 \sum_{i=1}^N \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, t) \varphi_i(t) \overline{\varphi_i(s)} ds dt \\ &= \int_{\mathcal{T}} \mathbf{E}|X_t|^2 dt - \sum_{i=1}^N \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, u) \varphi_i(u) \overline{\varphi_i(s)} ds du. \end{aligned}$$

Eine Minimierung des totalen mittleren quadratischen Fehlers ist daher äquivalent zu einer Maximierung der Summe $\sum_{i=1}^N \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, u) \varphi_i(u) \overline{\varphi_i(s)} ds du$.

Jeder Summand ist dabei positiv, da aufgrund der Positiv-Semidefinitheit der Kovarianzfunktion K die Ungleichung

$$\int_{\mathcal{T}} \int_{\mathcal{T}} K(s, u) \varphi_i(u) \overline{\varphi_i(s)} ds du \geq 0 \quad \text{für alle } i \in \mathbb{N} \quad (5.5)$$

gilt. Die Maximalität der Summe wird nun mittels vollständiger Induktion über N gezeigt:

Induktionsanfang: Für $N = 1$ verbleibt lediglich der erste Summand. Bereits Hilbert hat gezeigt, dass dieses Doppelintegral für die erste Eigenfunktion maximal wird. Dieser Teil orientiert sich an [Tri57, Kap. 3.11]. Sei im Folgenden $\Psi(\varphi) := \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, u) \varphi(s) \overline{\varphi(u)} ds du$. Es gilt nach Mercer's Theorem:

$$\begin{aligned} \Psi(\varphi) &= \int_{\mathcal{T}} \int_{\mathcal{T}} \left(\sum_{k=1}^{\infty} \lambda_k \phi_k(s) \overline{\phi_k(u)} \right) \varphi(u) \overline{\varphi(s)} ds du \\ &= \sum_{k=1}^{\infty} \lambda_k \left(\int_{\mathcal{T}} \phi_k(s) \overline{\varphi(s)} ds \right) \left(\int_{\mathcal{T}} \varphi(u) \overline{\phi_k(u)} du \right) = \sum_{k=1}^{\infty} \lambda_k |\langle \phi_k, \varphi \rangle|^2. \end{aligned}$$

An dieser Stelle erkennt man direkt $\Psi(\phi_1) = \lambda_1$. Dies benötigt man im weiteren Verlauf des Beweises. Unter Verwendung der Ungleichungskette $|\lambda_1| \geq |\lambda_2| \geq \dots$ und mit der aus der Funktionalanalysis bekannten Bessel'schen Ungleichung gilt für den absoluten Wert von $\Psi(\varphi)$:

$$|\Psi(\varphi)| \leq \sum_{k=1}^{\infty} |\lambda_k| |\langle \phi_k, \varphi \rangle|^2 \leq |\lambda_1| \sum_{k=1}^{\infty} |\langle \phi_k, \varphi \rangle|^2 \leq |\lambda_1| \|\varphi\|^2 = |\lambda_1|, \quad (5.6)$$

da die $(\varphi_i)_{i \in \mathbb{N}}$ normiert sind. Da $\Psi(\phi_1) = \lambda_1 \geq 0$ folgt somit für den Fall $N = 1$ die Behauptung.

Induktionsschritt: Angenommen, die Induktionsbehauptung gilt für ein vorgegebenes $N \in \mathbb{N}$. Um TMSE_N zu minimieren ist demnach $\varphi_i = \phi_i$ für $i = 1, \dots, N$, wobei die ϕ_i die Eigenfunktionen des von der Kovarianzfunktion induzierten Integraloperators sind. Dann gilt

$$\begin{aligned} \sum_{i=1}^{N+1} \int_{\mathcal{T}} \int_{\mathcal{T}} K(s, u) \varphi_i(u) \overline{\varphi_i(s)} ds du &= \sum_{i=1}^N \Psi(\phi_i) + \Psi(\varphi_{N+1}) \\ &= \sum_{i=1}^N \lambda_i + \Psi(\varphi_{N+1}). \end{aligned}$$

Außerdem ist erneut unter Hinzunahme der Normiertheit und Bessels Ungleichung

$$|\Psi(\varphi_{N+1})| \leq \sum_{k=1}^{\infty} |\lambda_k| |\langle \phi_k, \varphi_{N+1} \rangle|^2 = \sum_{k=N+1}^{\infty} |\lambda_k| |\langle \phi_k, \varphi_{N+1} \rangle|^2 \leq |\lambda_{N+1}| \sum_{k=1}^{\infty} |\langle \phi_k, \varphi_{N+1} \rangle|^2 \leq |\lambda_{N+1}|. \quad (5.7)$$

Da $\Psi(\phi_{N+1}) = \lambda_{N+1} \geq 0$ folgt die Behauptung und die Optimalität ist bewiesen. \square

Kapitel 6

Implementation der Karhunen-Loève-Transformation

Wie bereits im Abschnitt 4.5 zu beispielhaften Zerlegungen gesehen, besteht die größte Schwierigkeit auf dem Weg zur Darstellung eines Prozesses mit seiner Karhunen-Loève-Reihe im Lösen der Fredholm'schen Integralgleichung. Bei den diskutierten Beispielen war dies zwar mit einiger Arbeit verbunden, führte aber letztendlich zum Erfolg. In vielen Fällen ist die Fredholm'sche Integralgleichung allerdings nicht analytisch lösbar. Außerdem kann man aufgrund der individuellen analytischen Lösungen keinen Algorithmus implementieren, der zu einem allgemeinen Prozess die gesuchte Darstellung findet. Um dennoch sinnvolle Ergebnisse zu erzielen, können die Lösungen der Fredholm'schen Integralgleichung und damit die Karhunen-Loève-Zerlegung des Prozesses approximiert werden.

6.1 Numerisches Lösen der Fredholm'schen Integralgleichung

Es werden verschiedene Verfahren diskutiert, um die Lösungen der Fredholm'schen Integralgleichung zu approximieren. Der Abschnitt richtet sich nach [Wan08, Kapitel 2]. Im Folgenden sei der Einfachheit halber stets $\mathcal{T} = [0, 1]$.

6.1.1 Quadraturformeln

Quadraturformeln sind Verfahren zur numerischen Integration. Dabei werden die Integranden durch Interpolationspolynome approximiert. Letztendlich ergeben sich endliche Summen, die aus Funktionswerten und Wichtungen, abhängig vom Grad des Polynoms, bestehen. Dazu sei $\nabla : 0 = t_0 < \dots < t_{n+1} = 1$ eine äquidistante Partition von \mathcal{T} mit Feinheit $h = \frac{1}{n+1}$. Man erhält Wichtungen w_i für $i = 1, \dots, n + 1$ und eine Approximationen der Form

$$\int_0^1 f(t) dt \approx \sum_{i=0}^{n+1} w_i f(t_i). \quad (6.1)$$

Tabelle 6.1: Überblick über die wichtigsten Quadraturformeln

Bezeichnung	Quadraturformel
summierte uniforme Regel	$\frac{1}{n+2} \sum_{i=0}^{n+1} f(t_i)$
summierte Trapezregel	$\sum_{i=0}^n \frac{h}{2} (f(t_i) + f(t_{i+1}))$
summierte Simpson-Regel (falls $\frac{n-1}{2} \in \mathbb{N}$)	$\sum_{2i=0}^{\frac{n-1}{2}} \frac{h}{3} (f(t_i) + 4f(t_{i+1}) + f(t_{i+2}))$

Tabelle 6.2: Berechnung der Fehlerterme bei verschiedenen Quadraturformeln für $n = 45$

	uniforme Regel		Trapezregel		Simpson-Regel	
	$\epsilon_{\text{abs}}(\lambda_1)$	$\epsilon_{\text{rel}}(\lambda_1)$	$\epsilon_{\text{abs}}(\lambda_1)$	$\epsilon_{\text{rel}}(\lambda_1)$	$\epsilon_{\text{abs}}(\lambda_1)$	$\epsilon_{\text{rel}}(\lambda_1)$
Brown'sche Bewegung	0.00009	0.02 %	0.00004	0.01 %	0.00005	0.01 %
Brown'sche Brücke	0.00210	2.09 %	0.00000	0.04 %	0.00010	0.05 %
	$\epsilon_{\text{abs}}(\phi_1)$	$\epsilon_{\text{rel}}(\phi_1)$	$\epsilon_{\text{abs}}(\phi_1)$	$\epsilon_{\text{rel}}(\phi_1)$	$\epsilon_{\text{abs}}(\phi_1)$	$\epsilon_{\text{rel}}(\phi_1)$
Brown'sche Bewegung	0.0146	1.03 %	0.0294	2.06 %	0.0171	1.21 %
Brown'sche Brücke	0.0945	6.61 %	0.0945	6.61 %	0.0945	6.68 %

Wie diese Verfahren und insbesondere die Wichtungen hergeleitet werden, kann in [SB07] nachgelesen werden. Einen Überblick über die wichtigsten Quadraturformeln liefert Tabelle 6.1. Diese Formeln sollen auf die Fredholm'sche Integralgleichung angewendet werden. Das heißt, man approximiert

$$\int_0^1 K(s, t)\phi(t)dt \approx \sum_{j=0}^{n+1} w_j K(s, t_j)\phi(t_j) = \lambda\phi(s). \quad (6.2)$$

Diese Gleichung generiert das folgende lineare Gleichungssystem

$$\begin{aligned} K(t_0, t_0)\phi(t_0)w_0 + K(t_0, t_1)\phi(t_1)w_1 + \dots + K(t_0, t_{n+1})\phi(t_{n+1})w_{n+1} &= \lambda\phi(t_0) \\ K(t_1, t_0)\phi(t_0)w_0 + K(t_1, t_1)\phi(t_1)w_1 + \dots + K(t_1, t_{n+1})\phi(t_{n+1})w_{n+1} &= \lambda\phi(t_1) \\ &\vdots \\ K(t_{n+1}, t_0)\phi(t_0)w_0 + K(t_{n+1}, t_1)\phi(t_1)w_1 + \dots + K(t_{n+1}, t_{n+1})\phi(t_{n+1})w_{n+1} &= \lambda\phi(t_{n+1}). \end{aligned}$$

Mit den Notationen $K := (K(t_{i-1}, t_{j-1}))_{ij} \in \mathbb{C}^{(n+2) \times (n+2)}$, $\phi := (\phi(t_{i-1}))_i \in \mathbb{C}^{(n+2)}$ und W als Diagonalmatrix der Wichtungen erhält man in Matrixschreibweise $KW\phi = \lambda\phi$, beziehungsweise mit $\psi = W^{\frac{1}{2}}\phi$ das Gleichungssystem $W^{\frac{1}{2}}KW^{\frac{1}{2}}\psi = \lambda\psi$.

Da $W^{\frac{1}{2}}KW^{\frac{1}{2}}$ hermitesch ist, können stets Eigenwerte $\lambda_1, \dots, \lambda_{n+2} \in \mathbb{R}$ und zugehörige Eigenvektoren $\psi^{(1)}, \dots, \psi^{(n+2)} \in \mathbb{C}^{(n+2)}$ gefunden werden. Der Rechenaufwand steigt allerdings mit wachsender Größe des Gleichungssystems massiv an. Löst man dieses Eigenwertproblem und berechnet $\phi^{(i)}$ durch $\phi^{(i)} = W^{-\frac{1}{2}}\psi^{(i)}$, so erhält man diskrete Approximationen der ersten $(n+2)$ Eigenfunktionen und Eigenwerte des Integraloperators T_K .

In Tabelle 6.2 sind die relativen und absoluten Fehler des ersten Eigenwertes und der zugehörigen Eigenfunktion aufgelistet, die sich bei einer Approximation mit $n = 45$ im Verhältnis zur analytischen Lösung ergeben. Die Fehler der Eigenfunktionen wurden bezüglich der Supremumsnorm berechnet. Dies ist sinnvoll, da sie in $C([0, 1])$ liegen. Sowohl bei der Brown'schen Bewegung als auch bei der Brown'schen Brücke liefert die Trapezregel die besten Approximationen des ersten Eigenwertes. Die erste Eigenfunktion wird am Besten mit der uniformen Regel approximiert. Die Simpson-Regel erzielt in beiden Kategorien gute Ergebnisse und ist daher zu bevorzugen. Bei sämtlichen Simulationen im weiteren Verlauf wurde dennoch die Trapez-Regel verwendet, da sie im Gegensatz zur Simpson-Regel für alle $n \in \mathbb{N}$ anwendbar ist.

6.1.2 Fourier-Methode

Eine andere Strategie um die Lösungen der Fredholm'schen Integralgleichung zu approximieren besteht darin, die Eigenfunktionen bezüglich einer vorgegebenen Orthogonalbasis zu entwickeln. In diesem Abschnitt seien dazu für ungerades $M \in \mathbb{N}$ die ersten M Basisfunktionen der Fourierbasis gegeben durch

$$f_1(t) = 1, \quad f_{2i}(t) = \cos(2\pi it), \quad f_{2i+1}(t) = \sin(2\pi it), \quad i = 1, \dots, \frac{M-1}{2}. \quad (6.3)$$

Die Eigenfunktionen werden bezüglich der Fourierbasis bis zum M -ten Term in der Form

$$\phi_i(t) \approx \sum_{k=1}^M d_{ik} f_k(t) = D_i^T f(t), \quad \text{für } i = 1, \dots, M \quad (6.4)$$

entwickelt, wobei $f^T(t) := (f_1(t), \dots, f_M(t))$ und $D_i^T := (d_{i1}, \dots, d_{iM})$. Diesen Ausdruck substituiert man nun in der Fredholm'schen Integralgleichung für die i -te Eigenfunktion und erhält

$$D_i^T \int_0^1 K(s, t) f(t) dt = D_i^T \lambda_i f(s). \quad (6.5)$$

Um im weiteren Verlauf die Orthogonalität zu nutzen, multipliziert man diese Gleichung mit $f(s)$ und integriert beide Seiten nach s . Es ergibt sich

$$D_i^T \int_0^1 \int_0^1 K(s, t) f(t) f(s)^T dt ds = D_i^T \lambda_i \int_0^1 f(s) f^T(s) ds. \quad (6.6)$$

Mit den Notationen für die vier $M \times M$ Matrizen

$$A = (A_{jk})_{jk} = \left(\int_0^1 \int_0^1 K(s, t) f_k(s) f_j(t) ds dt \right)_{jk}, \quad D^T = (D_1, \dots, D_M),$$

$$B = (B_{jk})_{jk} = \left(\int_0^1 f_k(s) f_j(s) ds \right)_{jk} = \left(\delta_{jk} \|f_j\|_{L^2([0,1])}^2 \right)_{jk} \quad \text{und} \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_M)$$

fasst man die ersten M Integralgleichungen zusammen zu dem Gleichungssystem

$$DA = \Lambda DB \Leftrightarrow AD^T = BD^T\Lambda. \quad (6.7)$$

Setzt man nun $E = B^{\frac{1}{2}}D^T$ und multipliziert beide Seiten der vorherigen Gleichung von links mit $B^{-\frac{1}{2}}$, erhält man das stets lösbare hermitesche Eigenwertproblem $B^{-\frac{1}{2}}AB^{-\frac{1}{2}}E = E\lambda$. Man erhält daraufhin D als $D = E^T B^{-\frac{1}{2}}$ und kann damit die Eigenfunktionen berechnen.

6.1.3 Haar-Wavelet-Methode

Ähnlich zur Fourier-Methode kann man auch mit anderen Orthonormalbasen verfahren. Häufig werden dabei die sogenannten Haar-Wavelets verwendet. Die *Haar-Mutterfunktion* ist definiert als

$$H(t) = \mathbb{1}_{[0, \frac{1}{2})}(t) - \mathbb{1}_{[\frac{1}{2}, 1]}(t). \quad (6.8)$$

Für $p \in \mathbb{N}_0$ und $0 \leq p < 2^n$ seien die *Haar-Wavelets* als Verschiebungen und Streckungen der Haar-Mutterfunktion definiert:

$$H_{pm}(t) = \sqrt{2^p} H(2^p t - m) = \begin{cases} \sqrt{2^p}, & t \in [\frac{m}{2^p}, \frac{2m+1}{2^{p+1}}) \\ -\sqrt{2^p}, & t \in [\frac{2m+1}{2^{p+1}}, \frac{m+1}{2^p}) \\ 0, & \text{sonst.} \end{cases} \quad (6.9)$$

Außerdem sei $H_{-10}(t) = \mathbb{1}_{[0,1]}(t)$. Man kann zeigen, dass $\{H_{pm}, p \geq -1, 0 \leq m < 2^p\}$ ein vollständiges Orthonormalsystem von $L^2([0, 1], \mathcal{B}([0, 1]), \lambda)$ ist.

Die Stammfunktionen der Haar-Wavelets heißen *Schauder-Funktionen*. Mit Hilfe der Schauder-Funktionen kann eine Brown'sche Bewegung auf sehr elegante Weise als Reihe mit fast sicherer Konvergenz konstruiert werden. In dem Buch von Peter Mörters [MP10, Kapitel 1.1] wird dies ausführlich beschrieben.

Man erhält $M = 2^p$ orthonormale Basisfunktionen auf $[0, 1]$, indem man

$$H_1 = 1; \quad H_i = H_{pm} \quad i = 2^p + m + 1; \quad p = 0, 1, \dots, n-1; \quad 0 \leq m < 2^p \quad (6.10)$$

setzt. Für diese kann man dann analog wie im vorherigen Abschnitt verfahren, um die Eigenfunktionen zu berechnen.

In der Arbeit von L. Wang [Wan08] wurden alle drei in diesem Abschnitt vorgestellten Methoden verglichen. Die Fourier-Methode lieferte die schlechtesten Approximationen in Relation zum Rechenaufwand. Auch die Wavelet-Methode ist nicht präziser als die simple Trapez-Regel. Dennoch sei hervorgehoben, dass der Vorteil dieser Methoden ist, dass die Eigenfunktionen nicht als diskrete Approximationen berechnet werden, sondern über dem gesamten Intervall definiert sind.

6.2 Anwendungen der Karhunen-Loève-Transformation

6.2.1 Signalermittlung

Ziel dieses Abschnittes ist es, aus einem gestörten Signal die ursprünglichen Frequenzen herauszufiltern. Wie bereits in Abschnitt 4.3 erwähnt, isoliert die Karhunen-Loève-Zerlegung zeitliche und stochastische Abhängigkeiten, weshalb sie hierzu besonders geeignet ist.

Sei $s(t)$ ein deterministisches Signal, welches durch ein zufälliges Geräusch G_t gestört wird. Sowohl Geräusch als auch Signal werden auf dem kompakten Zeitintervall $\mathcal{T} = [0, 1]$ betrachtet. Durch Messungen erhält man einen Prozess $X = \{X_t = s(t) + G_t : t \in [0, 1]\}$.

Einen Sinuston mit einer Frequenz von f Hz kann man realisieren, indem man das Signal als passende Sinuswelle

$$s(t) = y_0 \sin(2\pi ft) \quad (6.11)$$

berechnet. Dabei ist y_0 die Amplitude, die im Folgenden auf Eins gesetzt wird. In der Praxis wird angenommen, dass G_t ein *weißes Rauschen* ist, welches man approximieren kann, indem man für jeden der diskreten Messpunkte unabhängige und identisch normalverteilte und zentrierte Zufallsvariablen simuliert. Der Algorithmus ist aber auch auf andere Geräusche anwendbar.

Das Verhältnis von der Amplitude der Sinuswelle zu der Standardabweichung der Zufallsvariablen, mit denen das Geräusche simuliert wird, wird als *Signal to noise ration*, kurz SNR bezeichnet.

Um aus X rückwirkend s zu gewinnen, kann bei geringen Störungen eine Fourier-Transformation genutzt werden. Mit dem Algorithmus zur schnellen Fourier-Transformation (FFT) können dabei schnelle und gute Ergebnisse berechnet werden. Sobald die SNR aber zu klein wird, bereits ab 0.1, sind die Ergebnisse der FFT nutzlos. Eine Signalermittlung mittels der Karhunen-Loève-Transformation kann auf folgende Art und Weise gewonnen werden:

- 1) Berechnung eines Schätzers Θ_t von $\mathbf{E}X_t$ für jedes $t \in \mathcal{T}$ mittels der Monte-Carlo-Methode.
- 2) Berechnung eines diskreten Schätzers κ der Kovarianzfunktion K des nahezu zentrierten Prozesses $\{X_t - \Theta_t, t \in [0, 1]\}$. Dabei wird $\kappa_{i,j} \approx \mathbf{E}X_i X_j$ mittels der Monte-Carlo-Methode berechnet.
- 3) Bestimmung der Eigenwerte und zugehörigen Eigenfunktionen des Integraloperators T_K unter Nutzung der Trapez-Regel und der Werte von κ . Anschließend werden die Eigenwerte nach der Größe sortiert.
- 4) Der größte Eigenwert λ und die zugehörige Eigenfunktionen ϕ spiegeln das wesentliche zeitabhängige Verhalten von X wieder. Eine FFT von $t \mapsto \phi(t) + \Theta_t$ liefert eine Approximation der Frequenz f . Es können auch weitere Eigenwerte und Eigenfunktionen in die Berechnung einbezogen werden.

Abbildung 5 vergleicht die Ergebnisse anhand von Frequenzspektren. Die Karhunen-Loève-Transformation erzielt signifikant bessere Resultate. Allerdings muss an dieser Stelle der erhebliche Rechenaufwand hervorgehoben werden. Während der FFT Algorithmus nur den Bruchteil einer Sekunde braucht, dauert es fast 10 Sekunden bis KLT das Ergebnis ausgibt. Dies motiviert die Suche nach einem Algorithmus zur schnellen KLT. Besondere Fortschritte hat hier in den letzten Jahren C. Maccone erzielt [Mac08].

6.2.2 Entfernen von Bildrauschen und Bildkompression

Bildrauschen verhält sich im Grunde analog zu dem oben beschriebenen Signalrauschen. Es tritt beispielsweise bei Aufnahmen der Erde aus dem Weltraum auf, da viele kleine Einflüsse wie Wolken, Partikel und Temperaturunterschiede die Bilder beeinflussen. Hat man genügend Aufnahmen, um eine verlässliche Kovarianzmatrix für die Pixel oder für Pixelgruppen des Bildes zu erstellen, so filtert eine Karhunen-Loève-Transformation die wesentlichen Zusammenhänge zwischen den Pixeln heraus und kann das Bildrauschen entfernen.

Auch in der Bildkompression spielt die KLT eine Rolle. Hierbei unterliegt das Bild, welches man komprimieren möchte, zwar keinem zufälligem Einfluss, man kann aber mit sehr ähnlichen Methoden, wie bei der Signalerkennung, die grundlegenden Muster und Zusammenhänge der Pixel erkennen. Da die Notation für diese Problematik sehr umfangreich ist, wird das Problem hier lediglich in Worten zusammengefasst.

- 1) Man betrachtet ein gegebenes Bild derart, dass man jede Zeile der Pixelmatrix aus $m \times n$ Pixeln als Realisation eines Pfades eines stochastischen Prozesses mit n Werten in \mathbb{R} ansieht.
- 2) Mithilfe der m gegebenen Realisationen für jede der n Zufallsvariablen können Schätzer für die Erwartungswerte gefunden werden, welche anschließend spaltenweise von dem ursprünglichen Bild subtrahiert werden, um einen annähernd zentrierten Prozess zu erhalten.
- 3) Nun berechnet man die Kovarianzmatrix des zentrierten Prozesses und deren Eigenwerte und Eigenvektoren. Dieser Schritt entspricht einem Lösen der induzierten Fredholm'schen Integralgleichung.
- 4) Es folgt eine Sortierung der Eigenwerte ihrer Größe nach. Die Eigenvektoren der Kovarianzmatrix sind eine Basis des \mathbb{R}^m und werden auch als Hauptkomponenten des Bildes bezeichnet. Sei H die Matrix der Hauptkomponenten und B die Matrix des Bildes. Es gilt $B = H^T H B$. Um das Bild zu komprimieren, betrachtet man die Matrix \hat{H} , die nur aus den ersten k Spalten von H besteht und speichert anstelle der Matrix B lediglich $\hat{H}B$ und \hat{H}^T . Man berechnet mittels $B \approx \hat{H}^T \hat{H} B$ eine Approximation des Bildes.

Alternativ kann das Bild auch in quadratische Pixelblöcke unterteilt werden, was vermutlich ein deutlich besseres Ergebnis liefert, aber in der Implementierung komplizierter wäre.

In Abbildung 6 wurde eine Aufnahme der Erde aus dem Weltraum mittels Hauptkomponentenanalyse um verschiedene Faktoren komprimiert.

Schlusswort

Das Alleinstellungsmerkmal dieser Arbeit liegt in der umfangreichen Darstellung der Karhunen-Loëve-Zerlegung und der dafür benötigten Grundlagen. In der Literatur findet man häufig nur sehr kurze Abschnitte zu dieser Thematik, die nur einzelne Gebiete beleuchten. Für diese umfangreiche Darstellung war daher eine ebenso umfangreiche Literaturrecherche notwendig. Durch die detailliert beschriebenen Grundlagen in den ersten drei Kapiteln, ist die Karhunen-Loëve-Zerlegung auch für Leser ohne tiefere Kenntnisse von Funktionalanalysis und stochastischen Prozessen zugänglich dargestellt worden.

Neben dem neuen Beweis der im fünften Kapitel besprochenen Optimalität der Karhunen-Loëve-Entwicklung bezüglich des totalen mittleren quadratischen Fehlers, wurden auch an einigen anderen Stellen Zwischenresultate selbst gezeigt. Der Zugang zur Karhunen-Loëve-Zerlegung über Bochner-Integration und die diskutierten Aspekte zur Messbarkeit des Prozesses sind ebenfalls hervorzuheben, da auch diese in der Literatur häufig außen vor gelassen werden. Die Theorie der Bochner-Integrale taucht dabei nur in den wenigsten Quellen auf, ist aber eine sehr gut passende und erkenntnisreiche Interpretation der Integrale, die mit anderen Zugängen kompatibel ist.

Die theoretischen Überlegungen wurden mit selbst erstellten Abbildungen verdeutlicht. Zusätzlich liefern die eigenständig implementierten Programme zur Signalermittlung, zur Bildkompression und zur Berechnung des totalen mittleren quadratischen Fehlers einen Einblick in das Anwendungsspektrum und zeigen das große Potential der Karhunen-Loëve-Zerlegung. In dieser Arbeit wurde eine Erweiterung des Karhunen-Loëve-Theorems auf einen mehrdimensionalen Definitionsbereich des zugrundeliegenden Prozesses in der Zeitkomponente besprochen. Dadurch war es möglich, eine Karhunen-Loëve-Reihe eines Brown'schen Blattes aufzustellen und damit einen Pfad dieses Prozesses zu simulieren. Zusätzlich kann das Theorem auch auf Prozesse mit mehr-dimensionalem Zielraum angepasst werden. Diese Thematik kann in [Wan08, Kapitel 5] nachgelesen werden.

Die verschiedenen zu dieser Arbeit entwickelten Programme haben gezeigt, dass der Aufwand zur Berechnung der Karhunen-Loëve-Transformation erheblich ist. Die Ursache hierfür liegt in der notwendigen Bestimmung der Eigenwerte und Eigenfunktionen des Integraloperators. In [Mac08, Kapitel 10] wird ein neuer Ansatz vorgestellt, der die Berechnung aller Eigenwerte und Eigenfunktionen umgeht. Die Karhunen-Loëve-Transformation erfolgt dort, indem der betragsmäßig größte Eigenwert durch eine Abhängigkeit von der Anzahl der Approximationsschritte berechnet wird. Nach Maccone wird der Rechenaufwand damit deutlich verringert.

Schlusswort

Ein sehr interessanter Beweis für das Karhunen-Loève-Theorem, der die Theorie der Hilberträume mit reproduzierendem Kern nutzt, kann in [BTA04, Kapitel 3] nachvollzogen werden. Der Zugang über RKHS ist sinnvoll, denn damit kann die Karhunen-Loève-Zerlegung als Entwicklung bezüglich einer Orthonormalbasis in \mathcal{H}_K verstanden werden. Außerdem bietet dieses Buch eine sehr detaillierte Beschreibung der RKHS-Theorie, mit der man auch andere Repräsentationen von stochastischen Prozessen finden kann.

Einen Einblick in das vielseitige Anwendungsspektrum der Karhunen-Loève-Zerlegung wurde durch das sechste Kapitel dieser Arbeit bereits gegeben. Erdbebenvorhersage, Mustererkennung, die Anwendung in der Finanzmathematik und die Abschätzung von maximalen Wellenhöhen auf den Ozeanen sind nur einige der weiteren, zahlreichen Anwendungen. Der interessierte Leser findet schnell eine Fülle an Artikeln zu verschiedensten Gebieten, in denen die Karhunen-Loève-Transformation genutzt werden kann.

Abbildungen

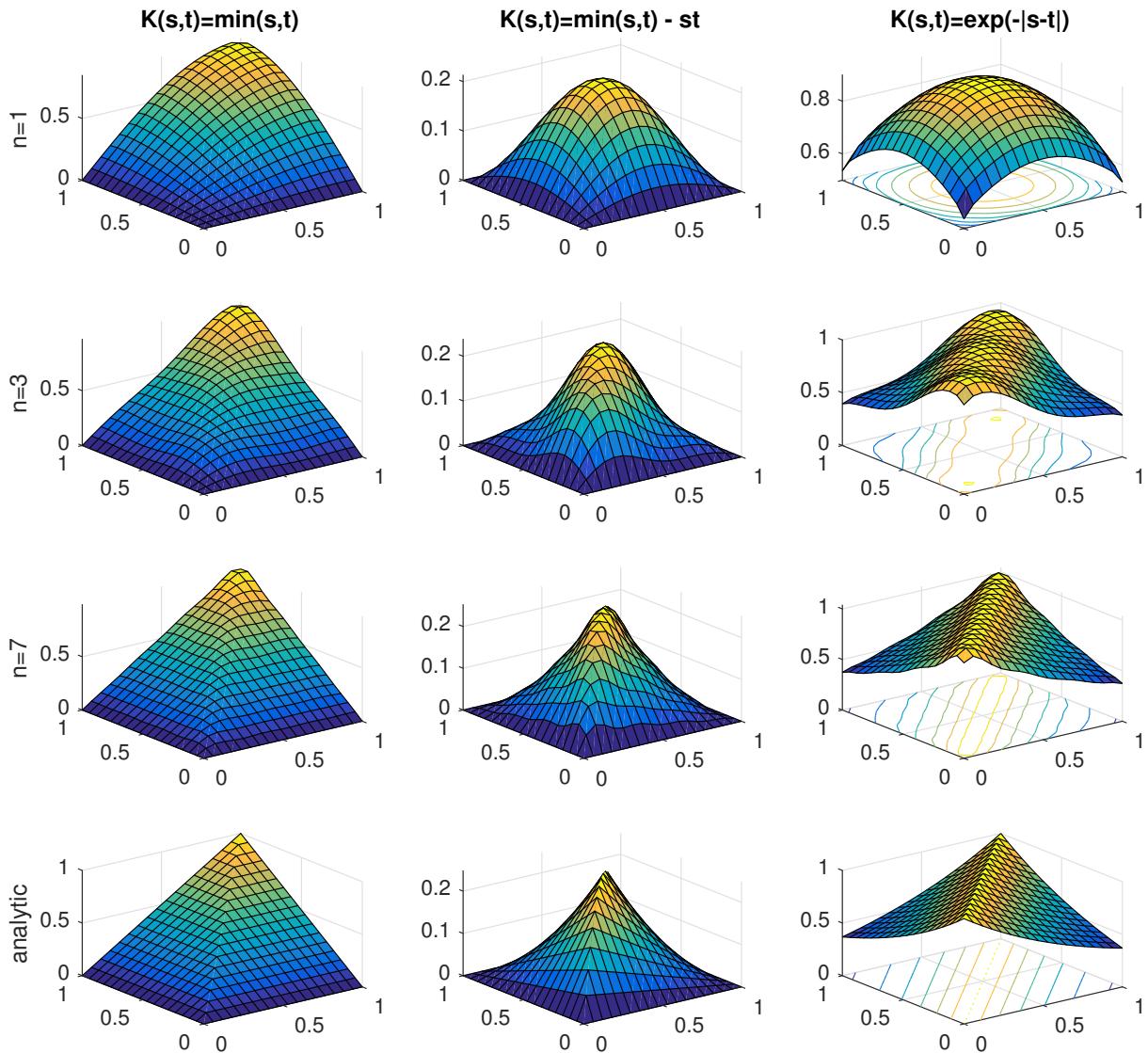


Abbildung 1: Approximation verschiedener Hilbert-Schmidt-Kerne mit Hilfe des Theorems von Mercer. Dabei wurden die Eigenwerte und Eigenvektoren aus der folgenden Abbildung genutzt. Um die Konvergenz zu verdeutlichen sind verschiedene Partialsummen dargestellt. Der erste Kern entspricht dem Kovarianzkern der Brown'schen Bewegung, der Zweite dem der Brown'schen Brücke. Der dritte Kern ist ein Exponentialkern. Er ist Kovarianzkern des Ohrenstein Uhlenbeck Prozesses. Die Grafiken wurden mit eigenständig mit MATLAB generiert.

Abbildungen

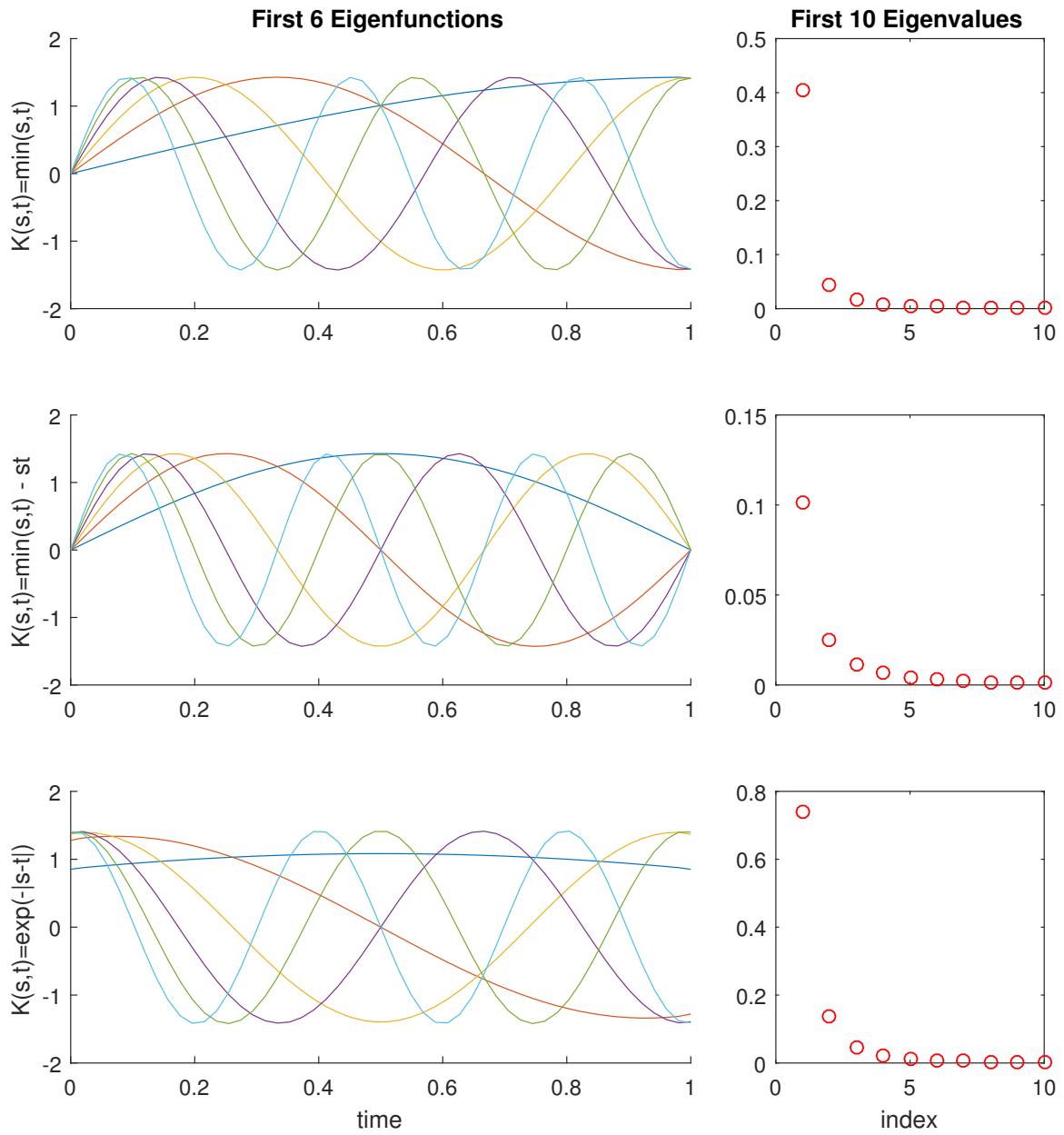


Abbildung 2: Hier sind Approximationen der ersten Eigenfunktionen und Eigenwerte verschiedener Hilbert-Schmidt-Operatoren dargestellt. Dabei wurde die Fredholm'sche Integralgleichung mittels der Trapezregel numerisch gelöst. Der erste Hilbert-Schmidt-Operator entspricht dem, der von der Braunschen Bewegung induziert wird. Der zweite Operator wird von der Braunschen Brücke induziert und der dritte Operator gehört zu einem Exponentialkern. Die Grafiken wurden eigenständig mit MATLAB generiert.

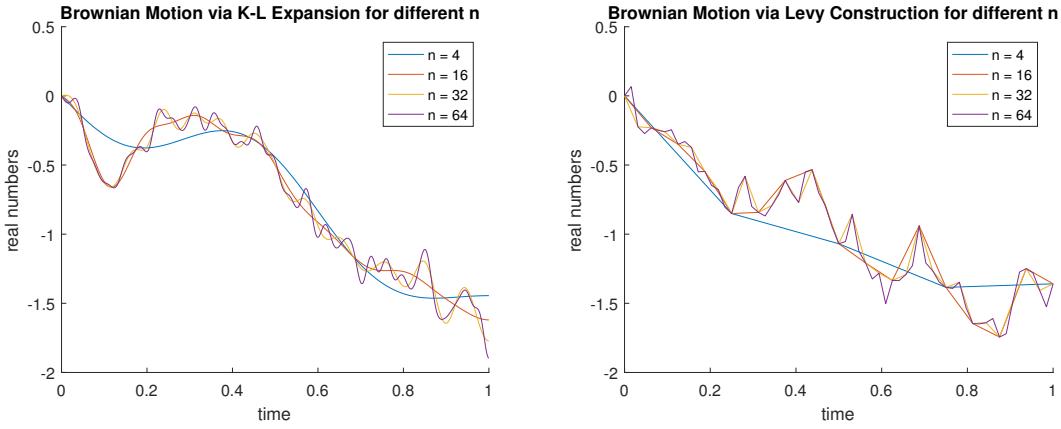


Abbildung 3: Hier sind verschiedene Weisen zur Konstruktion einer Brown'schen Bewegung dargestellt. Zur Linken wurden verschiedene Partialsummen der Karhunen-Loève-Reihe berechnet. Zur Rechten wurde eine Brown'sche Bewegung mittels der Levy-Konstruktion simuliert. Diese basiert auf einer Reihenentwicklung mittels der Schauder-Funktionen. Im Rahmen eines Vortrages zur Levy-Konstruktion der Brown'schen Bewegung wurde diese Abbildung selbst erstellt.

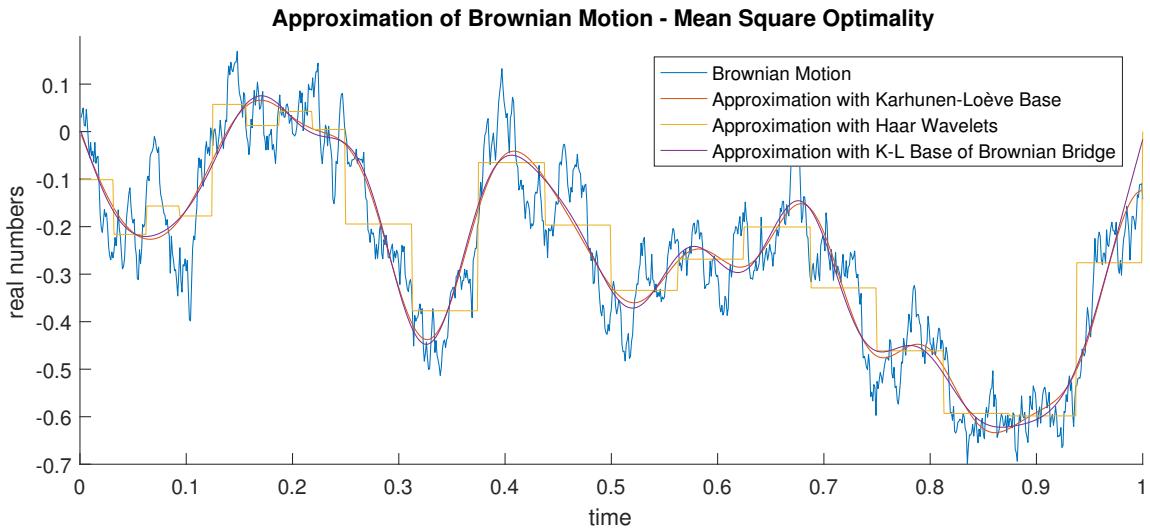


Abbildung 4: Hier wurde ein Pfad einer Brown'schen Bewegung bezüglich verschiedener Orthonormalbasen von $L^2([0, 1])$ entwickelt. Um die verschiedenen Approximationen allgemein vergleichen zu können, wurden die jeweiligen totalen mittleren quadratischen Abweichungen berechnet. Für die Approximation mit der Karhunen-Loève-Basis ergab sich $TMSE = 0.0049$, die Haar-Basis erzielt $TMSE = 0.0099$, die Basis der Brown'schen Brücke $TMSE = 0.0133$. Ein sehr gutes Ergebnis konnte die Basis des Exponentialkernes erzielen. Ihr Wert liegt mit $TMSE = 0.0052$ sehr nahe am Optimum. Plot und Berechnungen sind das Ergebnis eines eigenständig in MATLAB implementierten Programmes.

Abbildungen

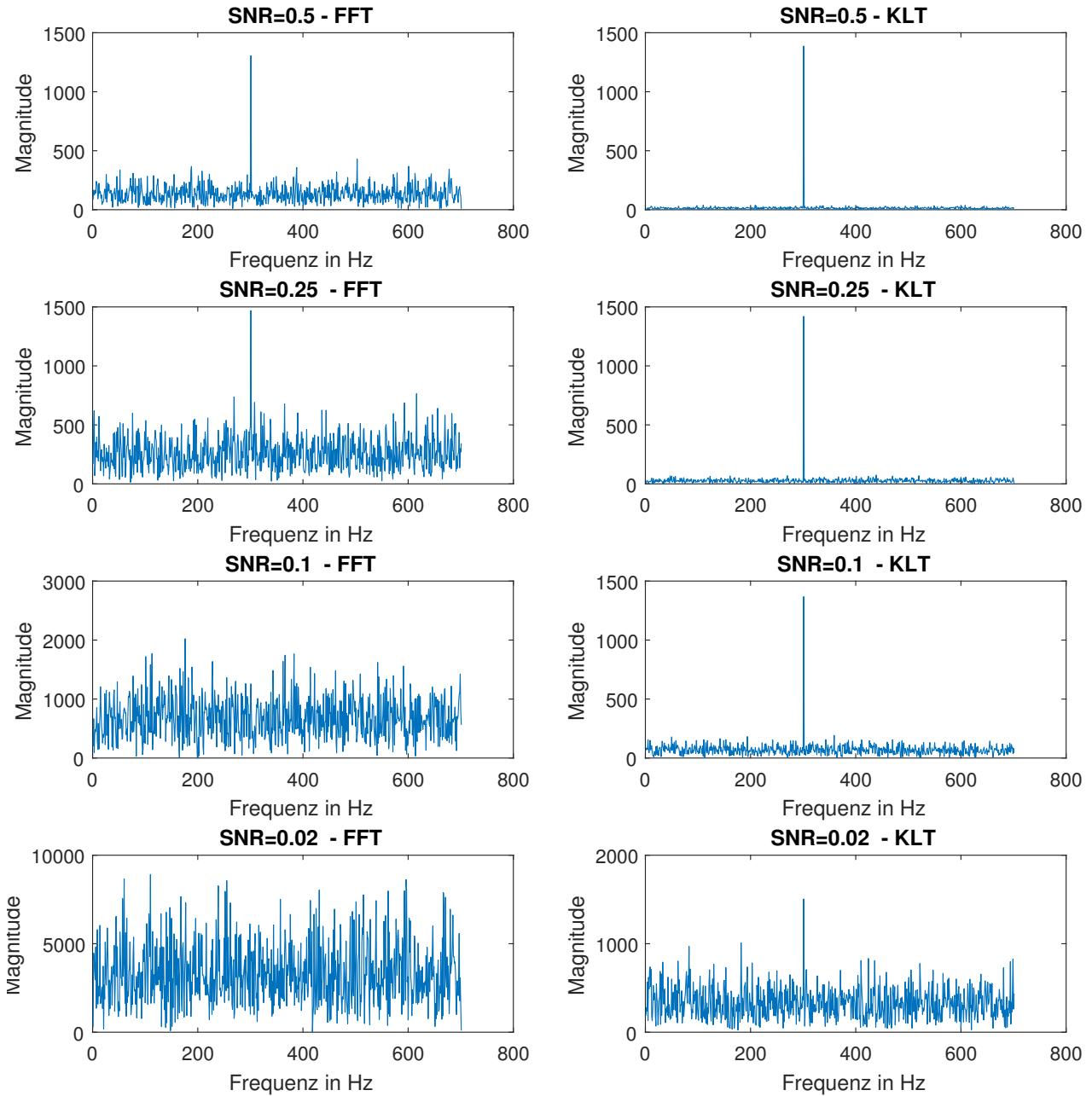


Abbildung 5: In diesen Abbildungen sind Frequenzspektren verschiedener Geräuschmessungen dargestellt. Die Messwerte sind die Summe aus einem deterministischen Ton mit einer Freguenz von 300 Hz und verschiedenen starken Störungen. Zur Ermittlung der Spektren wurde jeweils einmal die Fast-Fourier-Transform (FFT) und einmal die Karhunen-Lòeve Transformation (KLT) angewendet. Für die selbstständige Implementation der Karhunen-Loèvre-Transformation und das Generieren der Plots wurde MATLAB verwendet.

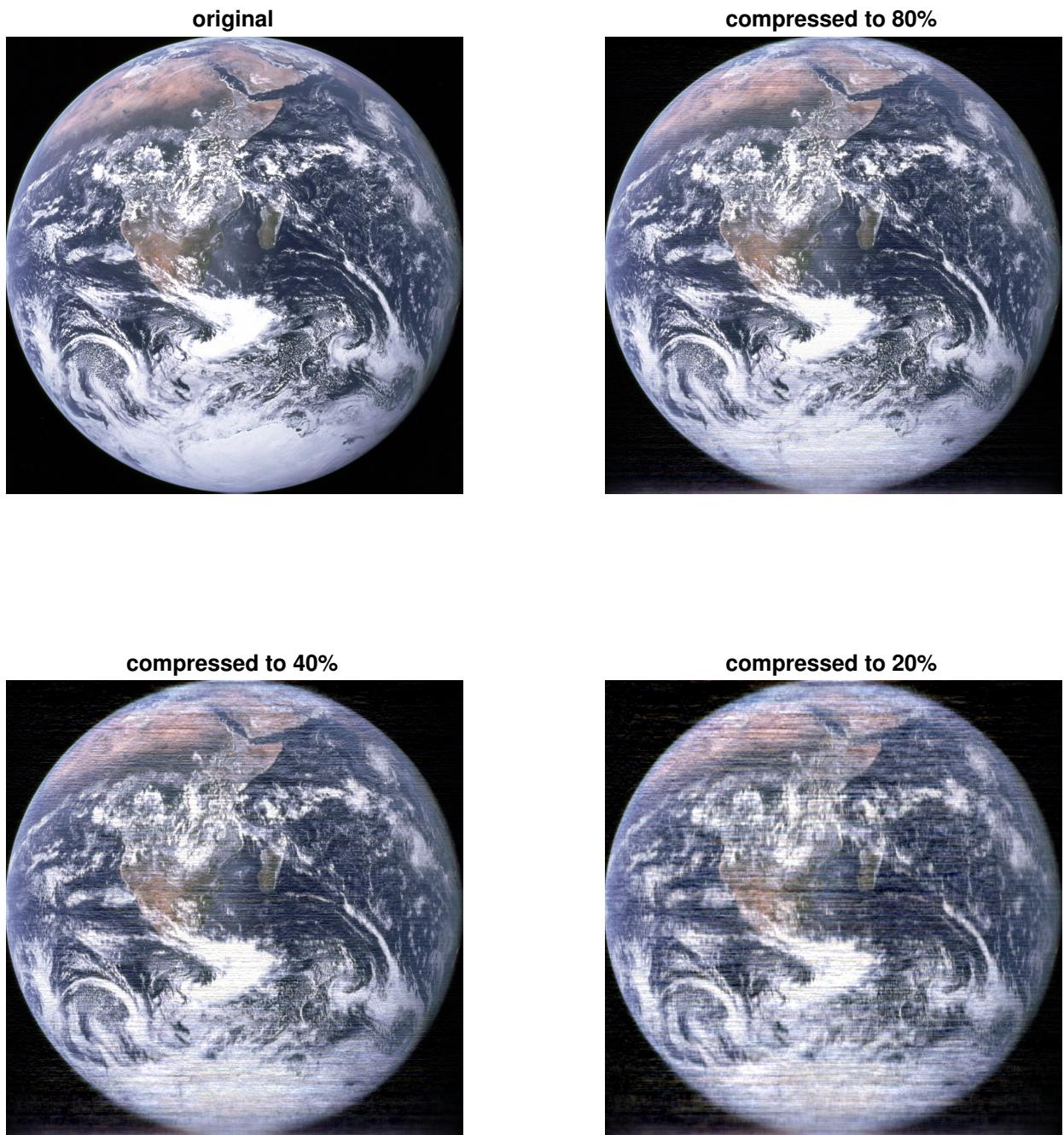


Abbildung 6: Hier sind vier Versionen des berühmten Bildes 'Blue Marbel' dargestellt. Es wurde im Jahre 1972 von der Besatzung der Apollo 17 aufgenommen. Das Bild wurde mittels Karhunen-Loëve-Transformation in verschiedenen Stufen mit einem selbstständig geschriebenen Programm in MATLAB komprimiert. 'Blue Marbel' steht auf der Website der NASA frei zur Verfügung.

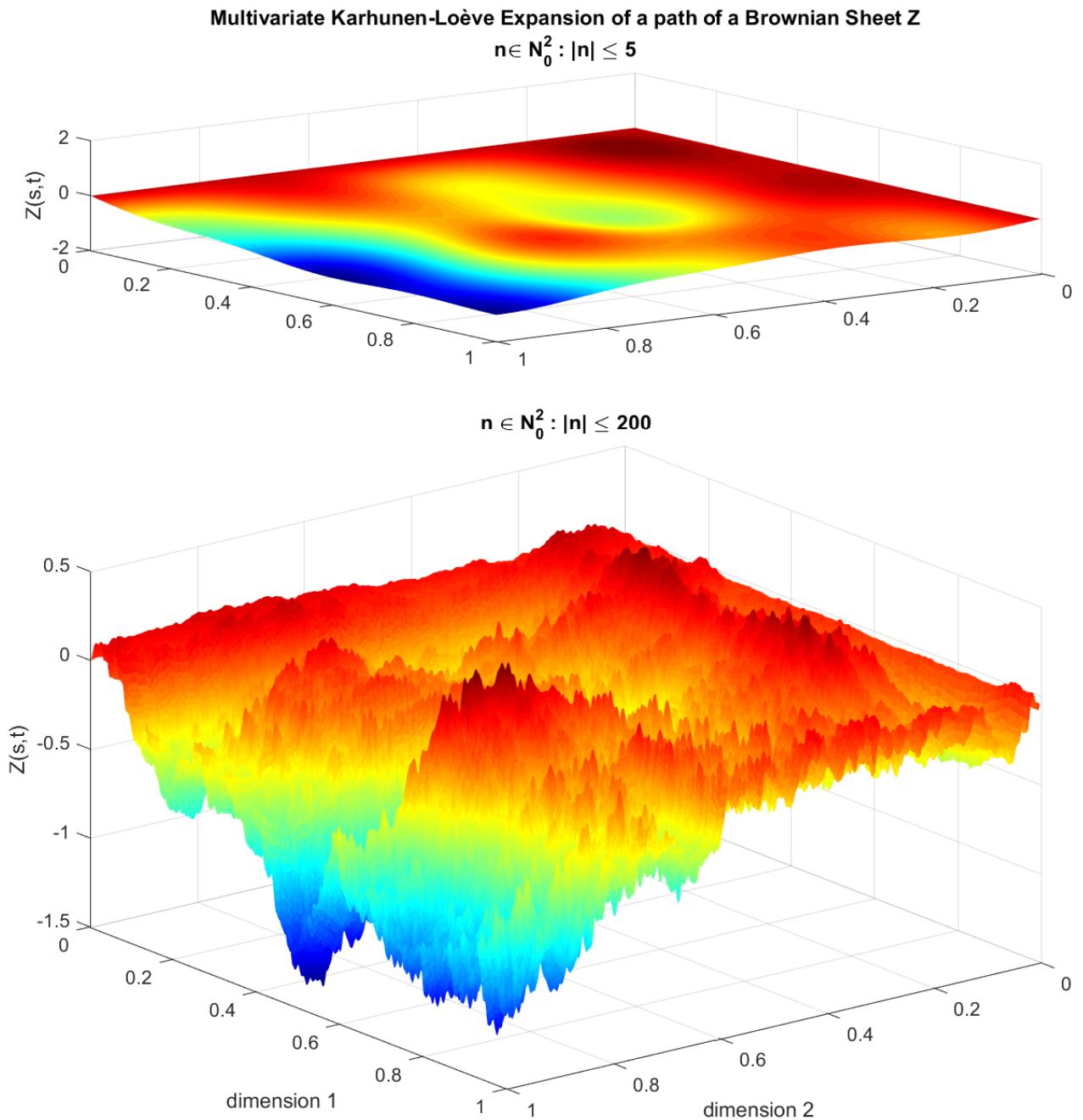


Abbildung 7: Hier ist ein Pfad eines Brown'schen Blattes dargestellt. Die Eigenwerte und Eigenfunktionen wurden in Beispiel 4.12 analytisch bestimmt. Das obere Bild zeigt dabei die Summe über alle $n \in N_0^2$ mit $|n| \leq 5$, also die ersten 25 Summanden. Wesentlich detaillierter ist im zweiten Bild der selbe Pfad mit einer genaueren Approximation dargestellt. Trotz der wenigen Summanden, die für das erste Bild verwendet wurden, ist das wesentliche Höhenprofil bereits dargestellt. Für die Erstellung des zweiten Plots war eine halbe Stunde Rechenzeit notwendig. Die Plots wurden selbstständig mit MATLAB erstellt.

Literaturverzeichnis

- [AB06] Charalambos Aliprantis und Kim Border. *Infinite Dimensional Analysis: A Hitch-hiker's Guide*. 3. Aufl. Berlin Heidelberg New York: Springer-Verlag, 2006.
- [Adl10] Robert J. Adler. *The Geometry of Random Fields*. 2. Aufl. Classics in Applied Mathematics. Philadelphia: SIAM Verlag, 2010.
- [Ale15] Alen Alexanderian. *A brief note on the Karhunen-Loève expansion*. North Carolina State University. Sep. 2015. URL: <http://users.ices.utexas.edu/~alen/articles/KL.pdf> (besucht am 18.06.2016).
- [BTA04] Alain Berlinet und Christine Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. 1. Aufl. New York: Kluwer Academic Publishers, 2004.
- [DH95] Robert D. Dony und Simon Haykin. »Optimally Adaptive Transform Coding«. In: *IEEE Transactions of Image Processing* 4.10 (Okt. 1995), S. 1358–1370.
- [DS57] Nelson Dunford und Jacob Schwartz. *Linear Operators: Part I General Theory*. New York: Interscience Publishers, 1957.
- [GGK03] Israel Gohberg, Seymour Goldberg und Marinus Kaashoek. *Basic Classes of Linear Operators*. Basel: Birkhäuser, 2003.
- [Jän00] Klaus Jänich. *Lineare Algebra*. 8. Aufl. Berlin Heidelberg New York: Springer-Verlag, 2000.
- [Kar47] Kari Karhunen. »Über lineare Methoden in der Wahrscheinlichkeitsrechnung«. In: *Annales Academiae Scientiarum Fennicae. Series A.I.* 37 (1947). U.S. Air Force Translation, Helsinki.
- [Loè48] Michel Loève. *Fonctions aleatoires du second ordre. Processus Stochastiques et Mouvement Brownien*. Hrsg. von Paul Levy in: *Processus stochastiques et mouvement Brownien*. Gauthier-Villars, 1948.
- [Loè91] Michel Loève. *Probability Theory II*. 4. Aufl. New York: Springer-Verlag, 1991.
- [Mac08] Claudio Maccone. *Deep Space Flight and Communications: Exploiting the Sun as a Gravitational Lens*. Berlin Heidelberg New York: Springer-Verlag, Jan. 2008.
- [MP10] Peter Mörters und Yuval Peres. *Brownian Motion*. 1. Aufl. Cambridge: Cambridge University Press, 2010.

- [Per16] Nicolas Perkowski. *Vorlesungsskript: Funktionalanalysis*. Berlin: Humboldt Universität zu Berlin. Wintersemester, 2016/17.
- [SB07] Stoer und Bulirsch. *Numerische Mathematik 1*. Hrsg. von Roland Freund und Roland Hoppe. 10. Aufl. Berlin Heidelberg New York: Springer-Verlag, 2007.
- [Szu11] Arkadiusz Szumski. »Karhunen-Loève Transform as an Istrument to Detect Weak RF signals«. In: *Inside GNSS* (Juni 2011). European Space Agency.
- [Tri57] Francesco Giacomo Tricomi. *Integral Equations*. New York: Courier Corporation, 1957.
- [Ueb16] Cassandra Uebel. »Karhunen-Loève Theorem - Funktionalanalytischer Hintergrund und Beispiele«. Bachelorarbeit. Berlin: Humboldt Universität zu Berlin, Juni 2016.
- [Wan08] Limin Wang. »Karhunen-Loeve Expansions and their Applications«. Diss. London: The London School of Economics und Political Science, März 2008.
- [Wer00] Dirk Werner. *Funktionalanalysis*. 3. Aufl. Berlin Heidelberg New York: Springer-Verlag, Juni 2000.
- [Wit05] Petra Wittbold. *Vorlesungsskript: Differentialgleichungen II*. Technische Universität Berlin. 2005. URL: <http://page.math.tu-berlin.de/~wittbold/Teach/LehreSS05/DGL2Kapitel4.pdf> (besucht am 15.06.2016).

Notationsverzeichnis

A	linearer beschränkter Operator	1
B	Banachraum	17
$\mathcal{B}(\mathcal{T})^{\otimes^2}$	Produkt- σ -Algebra von $\mathcal{B}(\mathcal{T})$ mit sich selbst	5
$\mathcal{B}(\mathcal{T})$	Borel- σ -Algebra auf \mathcal{T}	5
$B_X[0, 1]$	abgeschlossener Einheitsball in X	4
ϵ_{abs}	absoluter Fehler	42
ϵ_{rel}	relativer Fehler	42
E	Erwartungswertoperator	9
\mathcal{H}_K	RKHS zum Kern K	15
\mathcal{H}_X	von X generierter Hilbertraum	14
K	Hilbert-Schmidt-Kern bzw. Kovarianzfunktion eines Prozesses	5, 9
$\mathcal{K}(X)$	Banachraum der kompakten Operatoren von X nach X	4
$L^2(\Omega, \mathcal{F}, \mathbf{P})$	Hilbertraum der Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathbf{P})$ zweiter Ordnung	9
$L^2(\mathcal{T}, \mathcal{B}(\mathcal{T}), \lambda)$	Lebesgue-Raum der messbaren und quadratintegrierbaren Abbildungen $f : \mathcal{T} \rightarrow \mathbb{C}$	5
$(\lambda_i)_{i \in \mathbb{N}}$	Eigenwerte des Integraloperators aus dem Karhunen-Loève Theorem	26
λ	Eigenwert bzw. Lebesguemaß eingeschränkt auf \mathcal{T}	2, 5
$\bar{\lambda}$	komplexe Konjugation von $\lambda \in \mathbb{C}$	3
$(\lambda \otimes \mathbf{P})$	Produktmaß von λ und \mathbf{P}	23
λ^{\otimes^2}	Produktmaß von λ mit sich selbst	5
$L_{\mathbf{B}}$	Vektorraum der \mathbf{B} -Treppenfunktionen	18
$\mathcal{L}(X, X)$	Banachraum der linearen beschränkten Operatoren von X nach X	1
MSE	mittlere quadratische Abweichung	37

Notationsverzeichnis

$\mathcal{M}(\mathcal{T}, \mathbf{B})$	Vektorraum der stark messbaren Funktionen von \mathcal{T} nach \mathbf{B}	19
\mathbb{N}_0^d	Menge der d-dimensionalen Multiindizes mit Einträgen aus \mathbb{N}_0	33
$\ f\ $	Normfunktion	18
$(\Omega, \mathcal{F}, \mathbf{P})$	Wahrscheinlichkeitsraum	9
\oplus	direkte Summe von Vektorräumen	4
$(\phi_i)_{i \in \mathbb{N}}$	Orthonormalbasis von $L^2(\mathcal{T}, \mathcal{B}, \lambda)$ aus Eigenfunktionen eines Hilbert-Schmidt-Integraloperators, Karhunen-Loève-Basis	26
$(\varphi_i)_{i \in \mathbb{N}}$	beliebige Orthonormalbasis von $L^2(\mathcal{T}, \mathcal{B}, \lambda)$	16
Π	Projektionsoperator für Prozesse definiert durch $\Pi_J^I((X_t)_{t \in I}) = (X_t)_{t \in J}$	29
$\xrightarrow{q.M.}$	Konvergenz im quadratischen Mittelwert	12
$\rho(A)$	Resolventenmenge des Operators A	2
$RKHS$	Hilbertraum mit reproduzierendem Kern	14
$\sigma(A)$	Spektrum des Operators A	2
$\sigma_c(A)$	stetiges Spektrum des Operators A	2
$\sigma_p(A)$	Punktspektrum des Operators A	2
$\sigma_r(A)$	Residuenspektrum des Operators A	2
\mathcal{T}	kompakte Teilmenge von \mathbb{R}	5
T	kompakter Operator	4
Θ	kanonischer isometrischer Isomorphismus zwischen \mathcal{H}_K und \mathcal{H}_X	15
T_K	Hilbert-Schmidt-Integraloperator mit Kern K	5
$TMSE$	totale mittlere quadratische Abweichung	37
\sqcup	disjunkte Vereinigung	2
X	Hilbertraum bzw. zentrierter stochastischer Prozess zweiter Ordnung stetig im q.M.	1, 9
$(\xi_i)_{i \in \mathbb{N}}$	Orthonormalbasis von $L^2(\Omega, \mathcal{F}, \mathbf{P})$ aus dem Karhunen-Loève Theorem	26

Stichwortverzeichnis

- adjungierter Operator, 2
- bi-orthogonal System, 25
- Bochner-Integral, 18, 20
- Bochner-integrierbar, 20
- Eigenfunktion, 2
- Eigenvektor, 2
- Eigenwerte, 2
- Fredholm'sche Integralgleichung, 8
- Haar-Mutterfunktion, 44
- Haar-Wavelets, 44
- hermitischer Operator, 3
- Hilbert-Schmidt-Kern, 5
- Hilbert-Schmidt-Operator, 5
- Hilbertraum von Prozess generiert, 14
- Integraloperator, 5
- Karhunen-Loëve-Entwicklung, 26
- Karhunen-Loëve-Theorem, 26
- Karhunen-Loëve-Transformation, 26
- Karhunen-Loëve-Zerlegung, 26
- kompakter Operator, 4
- Konvergenz im q.M., 12
- Kovarianzfunktion, 9
- Kovarianzkern, 10
- Mercers Theorem, 8
- mittlere quadratische Abweichung, 37
- normaler Operator, 3
- Normfunktion, 18
- positiv, 8
- positiv semidefiniter Kern, 10
- positiver Kern, 11
- Prozess zweiter Ordnung, 9
- Punktspektrum, 2
- relativ kompakt, 4
- reproduzierender Kern, 14
- Residuenspektrum, 2
- Resolventenmenge, 2
- RKHS, 14
- Satz von Bochner, 21
- Schauder-Funktionen, 44
- selbstadjungiert, 3
- separable Kovarianzfunkzion, 32
- Signal to noise ration, SNR, 45
- Spektraltheorem, 4
- Spektralzerlegung, endl. dim., 1
- Spektrum, 2
- Standard-Repräsentation, 17
- stark μ -messbar, 18
- stetig im quadratischen Mittelwert, 12
- Stetiges Spektrum, 2
- totale mittlere quad. Abweichung, 37
- totale Varianz, 32
- Treppenfunktion, 17
- weißes Rauschen, 45

Stichwortverzeichnis

Selbstständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig verfasst und noch nicht für andere Prüfungen eingereicht habe. Sämtliche Quellen einschließlich Internetquellen für Texte, Grafiken, Tabellen und Bilder, sind als solche kenntlich gemacht. Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs beziehungsweise Täuschung eingeleitet wird.

Berlin, den

Datum, Unterschrift