SCHOOL OF COMPUTATION,
INFORMATION AND TECHNOLOGY —
INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Informatics

On the Formalization of Martingales

Ata Keskin

SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY — INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Informatics

On the Formalization of Martingales

Eine Formalisierung von Martingalen

Author:	Ata Keskin
Supervisor:	Prof. Tobias Nipkow
Advisor:	M. Sc. Katharina Kreuzer
Submission Date:	15 September 2023

I confirm that this bachelor’s thesis is my own work and I have documented all sources and material used.

Munich, 15 September 2023 Ata Keskin

Acknowledgements

Abstract

This thesis presents a formalization of martingales in arbitrary Banach spaces using Isabelle/HOL. The primary focus lies in the formal construction of conditional expectation in Banach spaces, which extends the existing formulation for real-valued functions. We replicate existing lemmas about martingales from the mathematical proof repository mathlib, which is primarily developed in Lean, based on homotopy type theory (HoTT). While mathlib explores formalization in Lean, we choose Isabelle/HOL as the theorem prover due to its powerful locale system that provides a structured and modular framework for representing these dynamic systems. The formalization of martingales and stochastic processes is achieved through Isabelle's locale system. We define the locale `stochastic_process` to formalize stochastic processes over arbitrary Banach spaces. Similarly, we define adapted, progressively measurable and predictable processes via the locales `adapted_process`, `progressive_process` and `predictable_process`. We also show sublocale relations and simple lemmas concerning vector space operations. Filtered measure spaces and σ-finite variants are introduced with the locales `filtered_measure` and `sigma_finite_filtered_measure`. Similarly, the locales `martingale`, `submartingale` and `supermartingale` are introduced to formalize martingales and related constructs in Banach spaces. Our formalization provides a robust mathematical framework for analyzing random processes.

Contents

1 Introduction

Martingales hold a central position in the theory of stochastic processes, making them a fundamental concept for the working mathematician. They provide a powerful way to study and analyze random phenomena, offering a formal framework for understanding the behavior of random variables over time. In various real-world scenarios, we encounter systems that evolve randomly over time. Representing such systems as martingales, we are able to investigate whether these systems remain bounded or converge to certain values in the long run.

In finance and economics, martingales are an invaluable tool for modeling asset prices [Fam65] and option pricing [MR05]. They provide insights into risk assessment, portfolio management, and the efficient market hypothesis, which postulates that asset prices fully reflect all available information [YB89].

Martingales are also closely related to several important probability limit theorems. These theorems, such as the strong law of large numbers and the central limit theorem, formalize the asymptotic behavior of sample means and sums of random variables. They have profound implications in statistics, allowing us to draw conclusions about large datasets and make predictions based on limited information.

In addition to their relevance in mathematics, martingales find applications in various interdisciplinary fields. Their ability to model randomness and analyze dynamic systems makes them useful in physics [Rol+23], biology, and computer science [MU05], among others.

In the scope of this thesis, we present a formalization of martingales in arbitrary Banach spaces using Isabelle/HOL. The background and related work section examines existing formalizations in two prominent formal proof repositories, mathlib (which uses the Lean theorem prover) and the Archive of Formal Proofs (AFP) (which uses Isabelle). Additionally, we conduct a short review of literature on conditional expectation and martingales in Banach spaces, laying a solid foundation for our research.

The current formalization of conditional expectation in the Isabelle library is limited to real-valued functions. To overcome this limitation, we extend the construction of conditional expectation to general Banach spaces, employing an approach similar to the one in [Hyt+16]. We justify our approach, by comparing it to two alternative constructions of the conditional expectation.

Subsequently, we define stochastic processes and introduce the concepts of adapted, progressively measurable and predictable processes using suitable locale definitions. Most importantly, we provide a generalization for the already present locale `filtration` by introducing the locales `filtered_measure` and `filtered_sigma_finite_measure`. These locales serve to formalize the concept of a filtered measure space. The latter also serves to generalize the locale `sigma_finite_subalgebra` which is necessary for the development of the theory of martingales.

Moving forward, we rigorously define martingales, submartingales, and supermartingales, presenting their first consequences and corollaries. Discrete and continuous time martingales are also covered in the formalization, benefiting from the complex and powerful locale system of Isabelle.

Our formalization fully encompasses the introductory mathlib theory on martingales and offers more generalization.

The thesis further contributes by generalizing concepts in Bochner integration, extending their application to general Banach spaces. Induction schemes for simple, integrable, and Borel measurable functions on Banach spaces are introduced, accommodating scenarios with or without a real vector ordering. These amendments expand the applicability of Bochner integration techniques.

The thesis concludes with reflections on the formalization approach, encountered challenges, and suggests future research directions.

2 Background and Related Work

In the following section, we explore existing formalizations of martingales within the mathematical proof repositories mathlib and AFP. Additionally, we will provide a concise introduction to the theory of integration in Banach spaces, establishing the mathematical foundation that underpins our formalization efforts.

2.1 Existing Formalizations

2.1.1 Lean Mathematical Library

Our main motivation for formalizing a theory of martingales in Isabelle/HOL comes from the existing in-depth formalization of the same subject in mathlib. As stated on their online platform, "The Lean mathematical library, mathlib, is a community-driven effort to build a unified library of mathematics formalized in the Lean proof assistant." The Lean-formalization of martingales consists of six documents. In the introductory Lean document `basic.lean`, fundamentals of the theory of martingales are formalized. The aim of this bachelor's thesis is to reproduce the results contained within this file in Isabelle/HOL. As will become clear in a moment, this is not a straightforward task, since there are a lot of dependencies missing in the Isabelle/HOL libraries.

The file `basic.lean` contains definitions for martingales, submartingales and supermartingales [DY22b]. The main results of this document are

→ `measure_theory.martingale` f $\mathcal{F}$ μ:

f is a martingale with respect to filtration $\mathcal{F}$ and measure μ.

→ `measure_theory.supermartingale` f $\mathcal{F}$ μ:

f is a supermartingale with respect to filtration $\mathcal{F}$ and measure μ.

→ `measure_theory.submartingale` f $\mathcal{F}$ μ:

f is a submartingale with respect to filtration $\mathcal{F}$ and measure μ.

→ `measure_theory.martingale_condexp` f $\mathcal{F}$ μ:

the sequence $(\mu[f|\mathcal{F}_i])_{i\in\mathcal{T}}$ is a martingale with respect to $\mathcal{F}$ and μ, where $\mu[f|\mathcal{F}_i]$ denotes the conditional expectation of f with respect to the subalgebra $\mathcal{F}_i$.

On a first note, we see that this theory relies heavily on the development of a conditional expectation operator. Prior to our development, the only formalization of conditional expectation in Isabelle/HOL was done in the real setting and resides in the theory document `HOL-Probability.Conditional_Expectation`. This formalization was accomplished by Sèbastien Gouëzel, presumably in anticipation of his latter entries [Gou15] and [Gou16]. We will delve further into the existing formalization and how our contribution improves upon it in the upcoming chapter.

Within the mathlib formalization, the majority of lemmata on martingales require the measures in question to be finite. In our formalization of martingales, we will demonstrate that σ-finiteness suffices alone. This approach is also consistent with our generalized formalization of conditional expectation, as it inherits the σ-finiteness requirement from the preexisting formalization in the real setting.

Another short-coming of the mathlib formalization is its treatment of predictable processes. The mathlib formalization contains the definition of adapted processes and progressively measurable processes. No explicit definition of a predictable process is given. Instead predictability is defined only in the discrete case, as a stochastic process which is adapted to the filtration $(\lambda i.\ \mathcal{F}_{i+1})$. In contrast, our formalization defines predictable processes more generally using the concept of a predictable σ-algebra. Additionally, our formalization contains adapted and progressively measurable processes as well. One of the major advantages of our formalization is the use of locales and sublocale relations. Concretely, we show the relationship

$$\texttt{stochastic} \supseteq \texttt{adapted} \supseteq \texttt{progressive} \supseteq \texttt{predictable}$$

Another important point to consider is the restrictions placed on the types in question. In the mathlib formalization, martingales are defined as a family of integrable functions $f : \iota \to \Omega \to E$. The mathlib formalization further requires that

- ι is a preordered set,
- Ω is a measurable space (i.e. a set together with a σ-algebra Σ),
- E is a normed, complete space with an addition operation.

These restrictions are easily replicated in our formalization using type classes and the type $'a$ `measure`. We simply restrict ourselves to functions $f :\ 't\ \to\ 'a\ \texttt{measure}\ \to\ 'b$, where the type $'t$ is an instance of the type class `linorder_topology` and the type $'b$ is an instance of the type class banach. With this specification, our approach mirrors the mathlib formalization, since measure spaces, measurable spaces and σ-algebras are all represented using the type measure in Isabelle/HOL. The additional requirement that $'t$ (equivalently ι in the mathlib case) be linearly ordered is easily justified as well, since in most contexts the index set represents a temporal dimension, which can obviously be linearly ordered. Apart from this, the topology induced must also come from the ordering on $'t$, since otherwise we can't have a useful definition of predictability in the general sense.

The main purpose of the mathlib formalization on martingales is to prove Doob's martingale convergence theorems, which concern discrete time and continuous time martingales (i.e. the naturals or the reals as indices). This justifies their focus on discrete time processes and the formulation of predictability only in the discrete case. More information on the specifics and the development of Doob's martingale convergence theorems is available in [DY22a].

This concludes our review of the mathlib formalization on martingales.

2.1.2 Archive of Formal Proofs

The Archive of Formal Proofs or AFP is a digital repository of formalized proofs and theories developed using the Isabelle theorem prover and proof assistant. The AFP hosts a variety of formalizations and proofs, primarily in the fields of logic, mathematics, and computer science. The repository allows researchers to share their formal proofs, theories, and related materials with the broader community.

The repository offers a search function, which allows us to find if any formalization on martingales has been done previously. A quick search yields the theory file DiscretePricing.Martingale. This entry DiscretePricing, which is attributed to Mnacho Echenim, focuses on the formalization of the Binomial Options Pricing Model in finance [Ech18]. A development of discrete time real-valued martingales is given in order to introduce the concept of risk-neutral measures. Similar to the development on mathlib, the goal of this entry is not to formalize martingales. A partial formalization of martingales is only given as a byproduct. The actual conference paper detailing the formalization can be found here [EP17].

Apart from this entry, no other development on the theory of martingales is present on AFP.

2.2 Mathematical Foundations and Reference Material

The main focus of our project is to formalize martingales in as general of a setting as possible. In this vein, we will study martingales defined on arbitrary Banach spaces, as opposed to the reals only. The main obstacle we will face is the development of conditional expectation in arbitrary Banach spaces. A great resource on this subject is the book *Analysis in Banach Spaces* by Hytönen et al [Hyt+16]. As a primer for the upcoming chapter, we will quickly cover the basics of integration on Banach spaces. The following information can also be found in the aforementioned book.

Remark. For the remainder of this document, unless stated otherwise explicitly, we fix a measure space $M = (\Omega, \Sigma, \mu)$ and a Banach space $(E, \|\cdot\|)$.

Integration on Banach spaces is usually done using the Bochner integral, which is defined similarly to the Lebesgue integral. For M a measure space and E a Banach space, we define the Bochner integral as follows

First, we consider simple functions $s : \Omega \to E$. These are functions which can be expressed μ-a.e. as finite sums of the form

$$s = \sum_{i=1}^{n} \mathbf{1}_{A_i} \cdot_{\mathbb{R}} c_i$$

where $\mathbf{1}_A$ is the indicator function of a set $A \in \Sigma$ and $c_i \in E$. Here $\cdot_{\mathbb{R}}$ denotes the scalar multiplication. We call such a function s Bochner integrable if $\mu(A_i) < \infty$ for all $i \in \{1, \ldots, n\}$. In this case, we define the Bochner integral simply as the sum

$$\int s \, \mathrm{d}\mu = \sum_{i=1}^{n} \mu(A_i) \cdot_{\mathbb{R}} c_i$$

If we replace E with $\mathbb{R}$, we can easily see that Bochner integrable simple functions are exactly those functions, which are Lebesgue integrable and simple.

We call a function $f : \Omega \to E$ strongly measurable, if there exists a sequence $(f_n)_{n \in \mathbb{N}}$ of simple functions converging to f μ-almost everywhere. A strongly measurable function f is called Bochner integrable with respect to μ, if there exists a sequence of Bochner integrable simple functions $f_n : \Omega \to E$ such that

$$\lim_{n \to \infty} \int_{\Omega} \|f - f_n\| \, \mathrm{d}\mu = 0$$

The integral used in this definition is the ordinary Lebesgue integral. This definition makes sense, since $w \mapsto \|f(w) - f_n(w)\|$ is μ-measurable and non-negative.

It can be shown via the triangle inequality that the integrals $\int f_n \, \mathrm{d}\mu$ form a Cauchy sequence. By completeness, this sequence converges to some element $\lim_{n \to \infty} \int f_n \, \mathrm{d}\mu \in E$. This limit is called the Bochner integral of f with respect to the measure μ

$$\int f \, \mathrm{d}\mu = \lim_{n \to \infty} \int f_n \, \mathrm{d}\mu$$

Furthermore, a function f in this setting is Bochner-integrable, if and only if the function $x \mapsto \|f(x)\|$ is integrable.

A formalization of the Bochner integral is available in Isabelle/HOL in the theory file `HOL-Analysis.Bochner_Integration` [HH11]. This formalization, which is due to Johannes Hölzl, has the additional assumption that the space E be second-countable. In the context of a metric space, this is the same as requiring separability.

Remark. One can show that a function f is strongly measurable if and only if it is essentially separably valued and for all $A \in \mathcal{B}(E)$ we have $f^{-1}(A) \in \Sigma$. Here $\mathcal{B}(E)$ denotes the Borel σ-algebra on E. A function is called essentially separably valued if there exists a μ-null set $N \subseteq \Omega$, such that $f(\Omega \setminus N)$ is separable as a subspace of E. Therefore, if E is already a separable Banach space, a function $f : \Omega \to E$ is strongly measurable if and only if it is Σ-measurable.

Consequently, we don't need to concern ourselves with definining strong measurability when working within separable (or equivalently second-countable) Banach spaces.

The book also contains an in depth section on the construction of the conditional expectation operator on Banach spaces. For our purposes, we only need to focus on the case where $f : \Omega \to E$ is a Bochner integrable function. In this case, the conditional expectation can be thought of as a linear operator $\mathbb{E}(\cdot|\mathcal{F}) : L^1(E) \to L^1(E)$ with respect to a sub-σ-algebra. The book contains theorems for the existence and uniqueness of conditional expectations (up to μ-null sets) for functions not only in $L^1(E)$, but also for those in $L^2(E)$ and $L^0(E)$. The latter is the space of strongly measurable functions with codomain E. Unsuprisingly, the definition of conditional expectation in this case is a bit more complicated, since it has to take into account the case where f is not integrable.

For the defining stochastic processes in a general setting, we have used the definitions presented in the books *PDE and Martingale Methods in Option Pricing* by Andrea Pascucci [Pas11] and *Stochastic Calculus and Applications* by Samuel N. Cohen and Robert J. Elliott [EC82]. Apart from these resources, we have made heavy use of the blog *Almost Sure* by George Thowler from the University of Cambridge [Tho].

Another extensive reference regarding martingales in Banach spaces is the book *Martingales in Banach Spaces* by Gilles Pisier [Pis16]. This resource provides an in-depth exploration of the theory of martingales in Banach spaces at a graduate level. Given the limited scope of this thesis, the book serves as a supplementary resource, as only a select few of its results are applicable to our elementary objectives.

3 Conditional Expectation in Banach Spaces

Conditional expectation extends the concept of expected value to situations where we have additional information about the outcomes. In a discrete setting, i.e. when the range of the random variables in question is countable, the setup is quite simple.

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let E be a complete normed vector space, i.e. a Banach space, and $S \subseteq E$ be some countable subset. Let $X : \Omega \to S$ be an integrable random variable and an event $A \in \mathcal{F}$ with $\mu(A) < \infty$. The conditional expectation of X given A, denoted as $\mathbb{E}(X|A)$, represents the expected value of X given that A occurs. In this simple case, we can directly define the conditional expectation as:

$$\mathbb{E}(X|A) = \sum_{w \in S} \frac{\mu(\{X = w\} \cap A)}{\mu(A)} \cdot w$$

Of course, this definition only makes sense if the value on the right hand side is finite and $\mu(A) \neq 0$. Defined this way, the conditional expectation satisfies the following equality

$$\begin{aligned} \int_A X \, \mathrm{d}\mu &= \sum_{w \in S} \mu(\{\mathbf{1}_A \cdot X = w\}) \cdot w \\ &= \mu(A) \cdot \mathbb{E}(X|A) \\ &= \int_A \mathbb{E}(X|A) \, \mathrm{d}\mu \end{aligned}$$

Remark. We use the notation "$c \cdot w$" to denote the scalar multiplication of $c \in \mathbb{R}$ and $w \in E$. When $E = \mathbb{R}$, it is just the standard multiplication on $\mathbb{R}$.

This observation motivates us to generalize the definition of conditional expectation to take into account not just a single event, but a collection of events. Fix $X : \Omega \to E$. Given a sub-σ-algebra $\mathcal{H} \subseteq \mathcal{F}$, we call an $\mathcal{H}$-measurable function $g : \Omega \to E$ a conditional expectation of X with respect to the sub-σ-algebra $\mathcal{H}$, denoted as $\mathbb{E}(X|\mathcal{H})$, if the following equality holds for all $A \in \mathcal{H}$

$$\int_A X \, \mathrm{d}\mu = \int_A g \, \mathrm{d}\mu$$

In the case that $E = \mathbb{R}$, it is straightforward to show that such a function g always exists (via Radon-Nikodym), and is unique up to a μ-null set. Notice that $\mathbb{E}(X|\mathcal{H})$ is a function $\Omega \to E$, as opposed to some value in E.

The suitable setting for defining the conditional expectation is when the sub-σ-algebra $\mathcal{H}$ gives rise to a σ-finite measure space. This is the case when $\mu|_{\mathcal{H}}$, the restriction of μ to $\mathcal{H}$ is a σ-finite measure. To see what goes wrong, consider the trivial sub-σ-algebra $\{\varnothing, \Omega\}$. A function which is measurable with respect to this σ-algebra is necessarily constant. Therefore, if $\mu(\Omega) = \infty$, no conditional expectation can exist, since it would have to be equal to 0 μ-almost everywhere in order to be integrable.

Example. Let $\mathcal{H} \subseteq \mathcal{F}$ be a sub-σ-algebra such that $\mu|_{\mathcal{H}}$ is a σ-finite measure. Given an integrable function $X : \Omega \to \mathbb{R}$, we can define a measure ν on $(\Omega, \mathcal{F})$ via

$$\nu(A) := \int_A X \,\mathrm{d}\mu$$

It is easy to verify that $\mu|_{\mathcal{H}}(A) = 0$ implies $\nu|_{\mathcal{H}}(A) = 0$, i.e. $\nu|_{\mathcal{H}}$ is absolutely continuous with respect to $\mu|_{\mathcal{H}}$. Using the Radon-Nikodym Theorem, we obtain an $\mathcal{H}$-measurable function $g : \Omega \to \mathbb{R}$ such that

$$\nu|_{\mathcal{H}}(A) = \int_A g \,\mathrm{d}\mu|_{\mathcal{H}}$$

Thus for any $A \in \mathcal{H}$, we have

$$\int_A X \,\mathrm{d}\mu = \int_A g \,\mathrm{d}\mu|_{\mathcal{H}} = \int_A g \,\mathrm{d}\mu$$

In the second equality, we use the fact that g is $\mathcal{H}$-measurable. Radon-Nikodym also guarentees that this function g is unique up to a $\mu|_{\mathcal{H}}$-null set. Since all $\mu|_{\mathcal{H}}$-null sets are also μ-null sets, the function g satisfies the requirements of a conditional expectation.

Technicalities aside, this shows that the conditional expectation always exists and is unique up to μ-null set for all $X \in \mathcal{L}^1(\mathbb{R})$. Our job now will be to construct a similar operator on arbitrary Banach spaces using methods from functional analysis and measure theory.

3.1 Preliminaries

In anticipation of our construction, we need to lift some results from the real setting to our more general setting. Our fundamental tool in this regard will be the **averaging theorem**. The proof of this theorem is due to Serge Lang [Lan93]. The theorem allows us to make statements about a function's value almost everywhere, depending on the value it's integral takes on various sets of the measure space.

3.1.1 Averaging Theorem

Before we introduce and prove the averaging theorem, we will first show the following lemma which is crucial for our proof. While not stated exactly in this manner, our proof makes use of the characterization of second-countable topological spaces given in the book General Topology by Ryszard Engelking (Theorem 4.1.15) [Eng89].

Lemma 3.1.1. *Let E be a separable metric space. Then there exists a countable set $D \subseteq E$, such that the set of open balls*

$$\mathcal{B} = \{B_\varepsilon(x) \mid x \in D,\ \varepsilon \in \mathbb{Q} \cap (0, \infty)\}$$

generates the topology on E. Here $B_\varepsilon(x)$ is the open ball of radius ε with centre x.

Proof. In the context of metric spaces, second-countability is equivalent to separability. Consequently, there exists some non-empty countable subset $D \subseteq E$, which is dense in E. We want to show that this D fulfills the statement above. For this end we will use the following equivalence which is valid for any $\mathcal{A} \subseteq \mathcal{P}(E)$

$$\mathcal{A} \text{ is topological basis} \iff \forall \text{open } U.\ \forall x \in U.\ \exists A \in \mathcal{A}.\ x \in A \wedge A \subseteq U$$

Let $U \subseteq E$ be open. Fix $x \in U$. Since U is open and we are working with the metric topology, there is some $\varepsilon > 0$, such that $B_\varepsilon(x) \subseteq U$. Furthermore, we know that a set D is dense if and only if for any non-empty open subset $O \subseteq E$, $D \cap O$ is also non-empty. Therefore, there exists some $y \in D \cap B_{\varepsilon/3}(x)$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists some $r \in \mathbb{Q}$ with $e/3 < r < e/2$. It is easy to check that $x \in B_r(y)$ and $B_r(y) \subseteq U$ with $y \in D$ and $r \in \mathbb{Q} \cap (0, \infty)$. This concludes the proof. □

Now we are ready to state and subsequently prove the averaging theorem.

Theorem 3.1.2. *(Averaging Theorem)*

Let $(\Omega, \mathcal{F}, \mu)$ be some σ-finite measure space. Let $f \in L^1(E)$. Let S be a closed subset of E and assume that for all measurable sets $A \in \mathcal{F}$ with finite and non-zero measure the following holds

$$\frac{1}{\mu(A)} \int_A f \, d\mu \in S$$

Then $f(x) \in S$ for μ-almost all x.

Proof. Without loss of generality we will show the statement assuming $\mu(\Omega) < \infty$. Let $v \in E$ and $v \notin S$.

We show by contradiction that if $B_r(v) \cap S = \varnothing$, then $A := f^{-1}(B_r(v))$, the set of all $x \in \Omega$ such that $f(x) \in B_r(v)$, is a μ-null set. Assume $\mu(A) > 0$. We have

$$
\begin{aligned}
\left\| \frac{1}{\mu(A)} \int_A f \, \mathrm{d}\mu - v \right\| &= \left\| \frac{1}{\mu(A)} \int_A f - v \, \mathrm{d}\mu \right\| \\
&\leq \frac{1}{\mu(A)} \int_A \|f - v\| \, \mathrm{d}\mu \\
&< r
\end{aligned}
$$

The last inequality follows from the fact that $f(x) \in B_r(v)$ for $x \in A$. This contradicts our first assumption. Therefore $\mu(A) = 0$.

Notice that $E \setminus S$ is an open subset of E. By the previous lemma, there exist open balls $B_{r_i}(w_i)$ with $r_i \in \mathbb{Q}_{\geq 0}$, $w_i \in D$ for $i \in \mathbb{N}$ such that $\bigcup_i B_{r_i}(w_i) = -S$. Obviously, $w_i \in E \setminus S$ and $B_{r_i}(w_i) \cap S = \varnothing$ for $i \in \mathbb{N}$. It follows

$$
\begin{aligned}
\mu(f^{-1}(E \setminus S)) &= \mu\left(\bigcup_i f^{-1}(B_{r_i}(w_i))\right) \\
&\leq \sum_i \mu(f^{-1}(B_{r_i}(w_i))) \\
&= 0
\end{aligned}
$$

Thus $\{f \notin S\}$ is a μ-null set, which completes the proof.

□

At the beginning of our proof, we assumed $\mu(\Omega) < \infty$ without loss of generality. This is only possible since we assumed the measure space in question to be σ-finite. To simplify the formalization of proofs employing this argument, we have introduced the following induction scheme

Lemma 3.1.3

```
lemma sigma_finite_measure_induct:
  assumes "⋀ N Ω. finite_measure N
        ⟹ N = restrict_space M Ω
        ⟹ Ω ∈ sets M
        ⟹ emeasure N Ω ≠ ∞
        ⟹ emeasure N Ω ≠ 0
        ⟹ almost_everywhere N Q"
  and "Measurable.pred M Q"
  shows "almost_everywhere M Q"
  ...
```

This induction scheme allows us prove results about a σ-finite measure space M, assuming that we can show the property on arbitrary subspaces of M with finite measure. For increased usability, we include additional assumptions such as `emeasure` $N\ \Omega \neq 0$ which let us to avoid unnecessary trivial cases. The proof of this induction scheme is straightforward.

Proof. Let $M = (\Omega, \Sigma, \mu)$ be a σ-finite measure space. There exists a family of sets with finite measure $(\Omega_i)_{i \in \mathbb{N}}$ such that $\bigcup_{i \in \mathbb{N}} \Omega_i = \Omega$. By assumption, the property Q holds μ-almost everywhere on all Ω_i. Therefore the sets $\Omega_i \cap \{\neg Q\} \in \Sigma|_{\Omega_i} \subseteq \Sigma$ are all μ-null sets. This means that $\bigcup_{i \in \mathbb{N}} (\Omega_i \cap \{\neg Q\}) = \{\neg Q\}$ is also μ-null set, which completes the proof. $\square$

Now that we have the averaging theorem at our disposal, we can lift the following results from the real case, to our more general setting.

Corollary 3.1.4. *Let $f \in L^1(E)$ and $\int_A f\, d\mu = 0$ for all measurable sets $A \subseteq \Omega$. Then $f = 0$ μ-almost everywhere.*

Proof. Apply the averaging theorem with $S = \{0\}$. $\square$

Corollary 3.1.5. *(Uniqueness of Densities)*
Let $f, g \in L^1(E)$ and $\int_A f\, d\mu = \int_A g\, d\mu$ for all measurable sets $A \subseteq \Omega$. Then $f = g$ μ-almost everywhere.

Proof. Follows directly from the previous corollary. $\square$

Corollary 3.1.6. *Let E be linearly orderable. Let $f \in L^1(E)$ and $\int_A f\, d\mu \geq 0$ for all measurable sets $A \subseteq \Omega$. Then f is non-negative μ-almost everywhere.*

Proof. Our first assumption guarantees that $\{y \in E \mid y \geq 0\}$ is a closed subset of E. Applying the averaging theorem on this set, yields the desired result. $\square$

The corollary on the uniqueness of densities is crucial in showing that the conditional expectation is unique as an element of $L^1(E)$.

3.1.2 Diameter Lemma

The goal of this subsection is to prove the diameter lemma, which provides a characterization of Cauchy sequences in metric spaces.

Definition 3.1.7. Let E be a metric space with metric $d : E \times E \to \mathbb{R}$. The diameter of a set A is defined as

$$\operatorname{diam}(A) = \sup_{x,y \in A} d(x, y)$$

Intuitively the diameter of a set A measures how spread out or "large" the set A is with respect to the distance defined by the metric.

Lemma 3.1.8. *(Diameter Lemma)*
Let E be a metric space with metric $d : E \times E \to \mathbb{R}$ and $(s_i)_{i\in\mathbb{N}} \subseteq E$ a sequence. Define $S_n = \{s_i \mid i \geq n\}$. The sequence $(s_i)_{i\in\mathbb{N}}$ is Cauchy, if and only if S_0 is bounded and

$$\lim_{n\to\infty} \operatorname{diam}(S_n) = 0$$

Proof.
First, assume $(s_i)_{i\in\mathbb{N}}$ is Cauchy.

Recall that a set A is bounded if there exists some $x \in E$ and $\varepsilon \in \mathbb{R}$ such that $d(x, y) \leq \varepsilon$ for all $y \in A$. Since $(s_i)_{i\in\mathbb{N}}$ Cauchy, there exists some $N \in \mathbb{N}$ such that $d(s_n, s_m) < 1$ for all $n, m \geq N$. The set $\{s_i \mid i \in \{0, \ldots, N\}\}$ is bounded since it is finite. Thus there exists some $a \in \mathbb{R}$ such that $d(s_N, s_i) < a$ for all $i \in \{0, \ldots, N\}$. Therefore $d(s_N, s_i) < \max(a, 1)$ for all $i \in \mathbb{N}$, which shows that S_0 is bounded.

We know $S_n \subseteq S_m$ for $n \geq m$. Therefore $\operatorname{diam}(S_n) < \infty$ for all $n \in \mathbb{N}$.

Let $\varepsilon > 0$. Then there exists some $N \in \mathbb{N}$ such that $d(s_n, s_m) < \frac{\varepsilon}{2}$ for all $n, m \geq N$. Hence

$$\operatorname{diam}(S_N) = \sup_{x,y\in S_N} d(x, y) \leq \frac{\varepsilon}{2} < \varepsilon$$

Furthermore, we have $\operatorname{diam}(S_n) \leq \operatorname{diam}(S_N)$ for $n \geq N$ because of the subset relation stated above. Thus $\lim_{n\to\infty} \operatorname{diam}(S_n) = 0$.

For the other direction, assume $\lim_{n\to\infty} \operatorname{diam}(S_n) = 0$ and that S_0 is bounded.

Hence $\operatorname{diam}(S_n) < \infty$ for all $n \in \mathbb{N}$ with the same argument as above.

Let $\varepsilon > 0$. There exists some $N \in \mathbb{N}$ such that $\sup_{x,y\in S_n} d(x, y) < \varepsilon$ for all $n \geq N$. Hence $d(x, y) < \varepsilon$ for all $x, y \in S_n$ for $n \geq N$. This implies $d(s_i, s_j) < \varepsilon$ for all $i, j \geq n \geq N$, which shows that $(s_i)_{i\in\mathbb{N}}$ is Cauchy. □

In our construction of the conditional expectation, we will use the diameter lemma to show that the limit of a sequence of simple functions admits a conditional expectation. In anticipation of this, we present the following lemmas concerning measurability and integrability.

Lemma 3.1.9

```
lemma borel_measurable_diameter:
  assumes "⋀x. x ∈ space M ⟹ bounded (range (λi. s i x))"
          "⋀i. (s i) ∈ borel_measurable M"
  shows "(λx. diameter {s i x | i. n ≤ i}) ∈ borel_measurable M"
  ...
```

Lemma 3.1.10

```
lemma integrable_bound_diameter:
  assumes "integrable M f"
          "⋀i. (s i) ∈ borel_measurable M"
          "⋀x i. x ∈ space M ⟹ norm (s i x) ≤ f x"
  shows "integrable M (λx. diameter {s i x | i. n ≤ i})"
  ...
```

The proofs are straightforward and depend on the measurability of the supremum function.

3.1.3 Induction Schemes for Integrable Simple Functions

In the upcoming sections of our work, we will frequently need to prove statements about integrable simple functions. For simple functions $s : \Omega \to \mathbb{R}_{\geq 0} \cup \{\infty\}$, the Isabelle theory `HOL_Analysis.Nonnegative_Lebesgue_Integration` already provides an induction scheme `simple_function_induct`. For our purposes we extend this scheme to cover integrable simple functions $s : \Omega \to E$. Notice that a simple function s is integrable if and only if $\mu(\{s \neq 0\}) < \infty$. The new induction scheme is as follows

Lemma 3.1.11

```
lemma integrable_simple_function_induct[case_names cong indicator add]:
  assumes "simple_function M f" "emeasure M {y ∈ space M. f y ≠ 0} ≠ ∞"
  assumes cong: "⋀f g. simple_function M f ⟹ emeasure M {y ∈ space M. f y ≠ 0} ≠ ∞
                ⟹ simple_function M g ⟹ emeasure M {y ∈ space M. g y ≠ 0} ≠ ∞
                ⟹ (⋀x. x ∈ space M ⟹ f x = g x) ⟹ P f ⟹ P g"
  assumes indicator: "⋀A y. A ∈ sets M ⟹ emeasure M A < ∞
                ⟹ P (λx. indicator A x *R y)"
  assumes add: "⋀f g. simple_function M f ⟹ emeasure M {y ∈ space M. f y ≠ 0} ≠ ∞
                ⟹ simple_function M g ⟹ emeasure M {y ∈ space M. g y ≠ 0} ≠ ∞
                ⟹ (⋀z. z ∈ space M ⟹ norm (f z + g z) = norm (f z) + norm (g z))
                ⟹ P f ⟹ P g ⟹ P (λx. f x + g x)"
  shows "P f"
  ...
```

The idea of the induction scheme is simple. We know f can be represented μ-a.e. as a finite sum $\sum_{i=1}^{n} \mathbf{1}_{A_i} \cdot c_i$ for some collection of measurable sets $(A_i)_{i=1,\ldots,n}$ and elements $c_i \in E$. We do induction on n. In this sense, "indicator" corresponds to the induction basis, while "add" corresponds to the induction step. Since f is representable as a finite sum μ-a.e. we need the additional assumption "cong" to make sure P is a well defined predicate on the space $L^1(E)$.

Remark. To make proving certain properties easier, we have the additional assumption $\|f(x) + g(x)\| = \|f(x)\| + \|g(x)\|$ in the induction step "add". It is easy to see why we can assume this without loss of generality. If we have some simple function $s = \sum_{i=1}^{n} \mathbf{1}_{A_i} \cdot c_i$, we can assume the sets A_i to be pairwise disjoint. Thus, if $x \in A_j$ for some $j \leq n$ we have $\|s(x)\| = \|\mathbf{1}_{A_j}(x) \cdot c_j\| = \sum_{i=1}^{n} \mathbf{1}_{A_i} \cdot \|c_i\|$.

When working with an ordering on E, we may need to concern ourselves with non-negative simple functions. For this goal, we have the following induction scheme.

Lemma 3.1.12

```
lemma integrable_simple_function_induct_nn[case_names cong indicator add]:
  assumes "simple_function M f" "emeasure M {y ∈ space M. f y ≠ 0} ≠ ∞"
          "⋀x. x ∈ space M ⟶ f x ≥ 0"
  assumes cong: "⋀f g. simple_function M f ⟹ emeasure M {y ∈ space M. f y ≠ 0} ≠ ∞
                ⟹ (⋀x. x ∈ space M ⟹ f x ≥ 0)
                ⟹ simple_function M g ⟹ emeasure M {y ∈ space M. g y ≠ 0} ≠ ∞
                ⟹ (⋀x. x ∈ space M ⟹ g x ≥ 0)
                ⟹ (⋀x. x ∈ space M ⟹ f x = g x) ⟹ P f ⟹ P g"
  assumes indicator: "⋀A y. y ≥ 0 ⟹ A ∈ sets M ⟹ emeasure M A < ∞
                     ⟹ P (λx. indicator A x ·R y)"
  assumes add: "⋀f g. simple_function M f ⟹ emeasure M {y ∈ space M. f y ≠ 0} ≠ ∞
               ⟹ (⋀x. x ∈ space M ⟹ f x ≥ 0)
               ⟹ simple_function M g ⟹ emeasure M {y ∈ space M. g y ≠ 0} ≠ ∞
               ⟹ (⋀x. x ∈ space M ⟹ g x ≥ 0)
               ⟹ (⋀z. z ∈ space M ⟹ norm (f z + g z) = norm (f z) + norm (g z))
               ⟹ P f ⟹ P g ⟹ P (λx. f x + g x)"
  shows "P f"
  ...
```

The induction scheme looks complicated and cumbersome, but in essence it is the same induction scheme as the previous one with the added assumption of non-negativity everywhere. The proof is also largely the same. We just need to show that the partial sums stay non-negative all the way through.

3.1.4 Bochner Integration on Linearly Ordered Banach Spaces

When working with the real numbers, the following statement is easy to show.

Let $f, g : \Omega \to \mathbb{R}$ be integrable and $f \geq g$ μ-a.e., then $\int f \, \mathrm{d}\mu \geq \int g \, \mathrm{d}\mu$.

In this subsection, we aim to provide similar results for functions $f, g : \Omega \to E$ with E a linearly ordered Banach space. For the remainder of our discourse, a topological space E is linearly ordered, if there exists a total ordering on E such that the topology on E and the order topology induced by the ordering coincide.

We start with the following lemma

Lemma 3.1.13. *Let $f \in L^1(E)$ and $f \geq 0$ μ-a.e. Then $\int f \, d\mu \geq 0$.*

Proof. Since $f \in L^1(E)$, there exists a sequence of integrable simple functions $(s_n)_{n\in\mathbb{N}}$, such that $\lim_{n\to\infty} s_n(x) = f(x)$ μ-a.e. and $\lim_{n\to\infty} \int s_n \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu$. At first, we have no further information about s_n. However, since we know that $f \geq 0$ μ-a.e, it follows that $f = \max(0, f)$ μ-a.e. Using dominated convergence and the fact that the function $\max(0, \cdot)$ is continuous w.r.t to the order topology on E, we can show

$$\lim_{n\to\infty} \max(0, s_n(x)) = \max(0, f(x)) \ \mu\text{-a.e.}$$

and

$$\lim_{n\to\infty} \int \max(0, s_n) \, \mathrm{d}\mu = \int \max(0, f) \, \mathrm{d}\mu$$

The function $\max(0, s_n)$ is still a simple and integrable function, which has the additional property of being always non-negative.

We will now show that if h is a non-negative simple function, then $\int h \, \mathrm{d}\mu \geq 0$. For this purpose, we will use the induction scheme for non-negative integrable simple functions that we proved in the previous subsection.

Case "cong": Let $h = g$ μ-a.e. and $\int g \, \mathrm{d}\mu \geq 0$. It follows directly

$$\int h \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu \geq 0$$

Case "indicator": Let $h = \mathbf{1}_A \cdot y$ for some measurable set A with finite measure and $y \in E$ with $y \geq 0$. It follows directly

$$\int h \, \mathrm{d}\mu = \mu(A) \cdot y \geq 0$$

Case "add": Let $h = h_1 + h_2$ for some integrable simple functions h_1 and h_2. By the induction hypothesis, we have $\int h_i \, d\mu \geq 0$ for $i = 1, 2$. Therefore

$$\int h \, d\mu = \int h_1 \, d\mu + \int h_2 \, d\mu \geq 0$$

Hence, we know $\int \max(0, s_n) \, d\mu \geq 0$ for all $n \in \mathbb{N}$. Therefore, the same must hold for the limit $\lim_{n\to\infty} \int \max(0, s_n) \, d\mu = \int \max(0, f) \, d\mu$. Since $f = \max(0, f)$ μ-a.e., we have $\int f \, d\mu = \int \max(0, f) \, d\mu$ and the statement follows.

□

Remark. For the proof of this statement, we need the topology on E to coincide with the order topology. Otherwise, we can't guarantee statements such as $(\forall i.\ x_i \geq 0) \implies \lim_{i\to\infty} x_i \geq 0$ or the continuity of the max function.

This lemma entails the following corollary.

Corollary 3.1.14. *Let $f, g \in L^1(E)$ and $f \geq g$ μ-a.e. Then $\int f \, d\mu \geq \int g \, d\mu$.*

In Isabelle, we can replace the assumption $f \in L^1(E)$ with Borel measurability, since a non-integrable function has the value of its integral set to 0 by default. The previous lemma can be stated as

```
lemma integral_nonneg_AE_banach:
  assumes "f ∈ borel_measurable M" and "AE x in M. 0 ≤ f x"
  shows "0 ≤ integral^L M f"
proof (cases "integrable M f")
  ...
qed
```

Furthermore, we have the following two corollaries, that also make use of the lemma on the uniqueness of densities.

Corollary 3.1.15

```
lemma integral_nonneg_AE_eq_0_iff_AE:
  assumes f: "integrable M f" and nonneg: "AE x in M. 0 ≤ f x"
  shows "integral^L M f = 0 ⟷ (AE x in M. f x = 0)"
proof
  assume *: "integral^L M f = 0"
  {
    fix A assume asm: "A ∈ sets M"
    have "0 ≤ integral^L M (λx. indicator A x · f x)" using nonneg by ...
    moreover have "··· ≤ integral^L M f" using nonneg by ...
    ultimately have "set_lebesgue_integral M A f = 0"
        unfolding set_lebesgue_integral_def using * by force
```

```
  }
  thus "AE x in M. f x = 0" by (intro density_zero f, blast)
qed (auto simp add: integral_eq_zero_AE)
```

Corollary 3.1.16

```
lemma integral_eq_mono_AE_eq_AE:
  assumes "integrable M f" "integrable M g"
          "integral^L M f = integral^L M g" "AE x in M. f x ≤ g x"
  shows "AE x in M. f x = g x"
proof -
  define h where "h = (λx. g x - f x)"
  have "AE x in M. h x = 0" unfolding h_def using assms by ...
  then show ?thesis unfolding h_def by auto
qed
```

3.2 Constructing the Conditional Expectation

Before we can talk about *the* conditional expectation, we must define what it means for a function to have *a* conditional expectation. For this purpose we define the following predicate

Definition 3.2.1

```
definition has_cond_exp :: "'a measure ⇒ 'a measure ⇒ ('a ⇒ 'b) ⇒ ('a ⇒ 'b) ⇒ bool" where
  "has_cond_exp M F f g = (∀A ∈ sets F. ∫_A f ∂M = ∫_A g ∂M)
                          ∧ integrable M f
                          ∧ integrable M g
                          ∧ g ∈ borel_measurable F"
```

This predicate precisely characterizes what it means for a function f to have a conditional expectation g w.r.t the measure M and the sub-σ-algebra F. Now we can use Hilbert's ϵ-operator, SOME in Isabelle [NPW02], to define *the* conditional expectation, if it exists.

Definition 3.2.2

```
definition cond_exp :: "'a measure ⇒ 'a measure ⇒ ('a ⇒ 'b) ⇒ ('a ⇒ 'b) ⇒ bool" where
  "cond_exp M F f = (if ∃g. has_cond_exp M F f g then (SOME g. has_cond_exp M F f g) else (λ_. 0))"
```

A major advantage of defining the conditional expectation this way is that it allows us to make statements about its measurability and integrability, without needing to show existence or uniqueness.

Lemma 3.2.3

```
lemma borel_measurable_cond_exp: "cond_exp M F f ∈ borel_measurable F"
  by (metis cond_exp_def someI has_cond_exp_def borel_measurable_const)
```

Lemma 3.2.4

```
lemma integrable_cond_exp: "integrable M (cond_exp M F f)"
  by (metis cond_exp_def has_cond_expD(3) integrable_zero someI)
```

3.2.1 Uniqueness

The conditional expectation of a function is unique up to a μ-null set.

Lemma 3.2.5. *Let $f, g \in L^1(E)$ such that "*`has_cond_exp` $M\ F\ f\ g$*" holds. Then "*`has_cond_exp` $M\ F\ f$ (`cond_exp` $M\ F\ f$)*" and*

$$\texttt{cond\_exp}\ M\ F\ f = g \quad \mu\texttt{-a.e.}$$

Proof. The first statement follows directly from the definition of `cond_exp`. To show `cond_exp` $M\ F\ f = g$ μ-`a.e.` we argue as follows. By the definition of `has_cond_exp` we have for any $A \in \texttt{sets}\ F$

$$\int_A f \,\mathrm{d}\mu = \int_A g \,\mathrm{d}\mu$$

and

$$\int_A f \,\mathrm{d}\mu = \int_A \texttt{cond\_exp}\ M\ F\ f \,\mathrm{d}\mu$$

Together with the lemma on the uniqueness of densities, we have `cond_exp` $M\ F\ f = g$ $\mu|_F$-`a.e.`. The lemma follows from the fact that all $\mu|_F$-null sets are also μ-null sets. □

Hence, we have the following succint characterization.

Lemma 3.2.6

```
lemma cond_exp_charact:
  assumes "⋀ A ∈ sets F. ∫_A f ∂M = ∫_A g ∂M"
          "integrable M f"
          "integrable M g"
          "g ∈ borel_measurable F"
  shows "AE x in M. cond_exp M F f x = g x"
  by (intro has_cond_exp_charact has_cond_expI' assms) auto
```

3.2.2 Existence

Showing the existence is a bit more involved. Specifically, what we aim to show is that has_cond_exp $M\ F\ f$ (cond_exp $M\ F\ f$) holds for any Bochner integrable f. We will use the standard machinery of measure theory. First, we will prove existence for indicator functions. Then we will extend our proof by linearity to simple functions. Finally we use a limiting argument to show that the conditional expectation exists for all Bochner integrable functions.

The conditional expectation operator has already been formalized for real valued functions by Sèbastien Gouëzel via the definition real_cond_exp. The following lemmas show that our definition characterizes the same operator, at least in the real case.

Lemma 3.2.7

```
lemma has_cond_exp_real:
  assumes "integrable M f"
  shows "has_cond_exp M F f (real_cond_exp M F f)"
  by (intro has_cond_expI', auto intro!: real_cond_exp_intA assms)
```

Lemma 3.2.8

```
lemma cond_exp_real:
  assumes "integrable M f"
  shows "AE x in M. cond_exp M F f x = real_cond_exp M F f x"
  using has_cond_exp_charact has_cond_exp_real assms by blast
```

We can now show that the conditional expectation of indicator functions exist.

Lemma 3.2.9. *Let $A \subseteq \Omega$ be measurable with $\mu(A) < \infty$ and $y \in E$. Then*

$$\texttt{has\_cond\_exp}\ M\ F\ (\mathbf{1}_A \cdot y)\ ((\texttt{real\_cond\_exp}\ M\ F\ \mathbf{1}_A) \cdot y)$$

Proof. The statement follows directly from the linearity of the Bochner integral and the previous lemmas. □

Next, we show the following lemma concerning the sum of two conditional expectations.

Lemma 3.2.10. *Assume* has_cond_exp $M\ F\ f\ f'$ *and* has_cond_exp $M\ F\ g\ g'$. *Then*

$$\texttt{has\_cond\_exp}\ M\ F\ (f + g)\ (f' + g')$$

Proof. The statement follows directly from the linearity of the Bochner integral. □

Now, we can show that the conditional expectation of integrable simple functions exist.

Lemma 3.2.11

```
corollary has_cond_exp_simple:
  assumes "simple_function M f" "emeasure M {y ∈ space M. f y ≠ 0} ≠ ∞"
  shows "has_cond_exp M F f (cond_exp M F f)"
  using assms
proof (induction rule: integrable_simple_function_induct)
  case (cong f g)
  then show ?case using has_cond_exp_cong by (metis (no_types, opaque_lifting)
                        Bochner_Integration.integrable_cong has_cond_expD(2)
                        has_cond_exp_charact(1))
next
  case (indicator A y)
  then show ?case using has_cond_exp_charact[OF has_cond_exp_indicator] by fast
next
  case (add u v)
  then show ?case using has_cond_exp_add has_cond_exp_charact(1) by blast
qed
```

Now comes the most difficult part. Given a convergent sequence of integrable simple functions $(s_n)_{n\in\mathbb{N}}$, we must show that the sequence $(\texttt{cond\_exp}\ M\ F\ s_n)_{n\in\mathbb{N}}$ is also convergent. Furthermore, we must show that this limit satisfies the properties of a conditional expectation. Unfortunately, we will only be able to show that this sequence convergence in L^1. Luckily, this is enough to show that the operator $\texttt{cond\_exp}\ M\ F$ preserves limits as a function $L^1(E) \to L^1(E)$. We need the following lemma for this purpose

Lemma 3.2.12. *(Contractivity for Simple Functions)*
Let $f : \Omega \to E$ be an integrable simple function. Then

$$\|\texttt{cond\_exp}\ M\ F\ s\| \leq \texttt{cond\_exp}\ M\ F\ (\lambda x. \|s\ x\|)$$

Proof. In the real case, one can show this property by decomposing a function into positive and negative parts. The statement follows via the induction scheme `integrable_simple_function_induct`. □

The following lemma is the most involved result of our formalization.

Lemma 3.2.13. *Let $f : \Omega \to E$ be an integrable function. Let $(s_n)_{n\in\mathbb{N}}$ be a sequence of integrable simple functions, such that $\lim_{n\to\infty} s_n(x) = f(x)$ and $\forall n.\ \|s_n(x)\| \leq 2 \cdot \|f(x)\|$ for μ-almost all x. Then there exists some subsequence $(s_{r_n})_{n\in\mathbb{N}}$ such that*

$$(\texttt{cond\_exp}\ M\ F\ s_{r_n})_{n\in\mathbb{N}} \text{ is Cauchy } \mu\text{-a.e.}$$

and

$$\texttt{has\_cond\_exp}\ M\ F\ f\ \Big(\lim_{n\to\infty} \texttt{cond\_exp}\ M\ F\ s_{r_n}\Big)$$

Proof. The sequence $(s_n)_{n\in\mathbb{N}}$ is Cauchy μ-a.e. Hence $\lim_{n\to\infty} \operatorname{diam}(S_n(x)) = 0$, with $S_n(x) := \{s_i(x) \mid i \geq n\}$ by the diameter lemma. Furthermore

$$\|\operatorname{diam}(S_n(x))\| \leq 4 \cdot \|f(x)\| \ \mu\text{-a.e.}$$

using the triangle inequality and our second assumption. We have already shown that $\operatorname{diam}(S_n(x))$ is measurable. We apply the dominated convergence theorem and get

$$\lim_{n\to\infty} \int \operatorname{diam}(S_n(x)) \, \mathrm{d}\mu = 0$$

We will now show that $(\texttt{cond\_exp}\ M\ F\ s_n)_{n\in\mathbb{N}}$ is Cauchy in the L^1 norm. Let $\varepsilon > 0$. Hence there is some $N \in \mathbb{N}$ such that $\int \operatorname{diam}(S_n(x)) < \varepsilon$. Thus for any $i, j \geq N$, we have

$$\int \|s_i(x) - s_j(x)\| \, \mathrm{d}\mu \leq \int \operatorname{diam}(S_N(x)) \, \mathrm{d}\mu < \varepsilon$$

by the monotonicity of the integral. Furthermore

$$\begin{aligned}
&\int \|(\texttt{cond_exp}\ M\ F\ s_i)(x) - (\texttt{cond_exp}\ M\ F\ s_j)(x)\| \, \mathrm{d}\mu \\
&= \int \|(\texttt{cond_exp}\ M\ F\ (s_i - s_j))(x)\| \, \mathrm{d}\mu \\
&\leq \int (\texttt{cond_exp}\ M\ F\ (\lambda x.\ \|s_i(x) - s_j(x)\|))(x) \, \mathrm{d}\mu \\
&= \int \|s_i(x) - s_j(x)\| \, \mathrm{d}\mu \\
&< \varepsilon
\end{aligned}$$

since $s_i(x) - s_j(x)$ is an integrable simple function and conditional expectation already exists in the real setting. Hence $(\texttt{cond\_exp}\ M\ F\ s_n)_{n\in\mathbb{N}}$ is Cauchy in the L^1 norm. Therefore, there exists some subsequence $(\texttt{cond\_exp}\ M\ F\ s_{r_n})_{n\in\mathbb{N}}$ that convergences μ-a.e. We have for all $n \in \mathbb{N}$

$$\|(\texttt{cond\_exp}\ M\ F\ s_{r_n})(x)\| \leq \texttt{cond\_exp}\ M\ F\ (\lambda x.\ 2 \cdot \|f(x)\|) \ \mu\text{-a.e.}$$

Together with the dominated convergence theorem, this implies that
$\lim_{n\to\infty}(\texttt{cond\_exp}\ M\ F\ s_{r_n})$ is integrable.
As the limit of F-measurable functions, $\lim_{n\to\infty}(\texttt{cond\_exp}\ M\ F\ s_{r_n})$ is also F-measurable.
Finally, we have for $A \in \texttt{sets}\ F$

$$\begin{aligned}
\int_A (\lim_{n\to\infty}(\texttt{cond_exp}\ M\ F\ s_{r_n}))(x) \, \mathrm{d}\mu &= \lim_{n\to\infty} \int_A (\texttt{cond_exp}\ M\ F\ s_{r_n})(x) \, \mathrm{d}\mu \\
&= \lim_{n\to\infty} \int_A s_{r_n}(x) \, \mathrm{d}\mu \\
&= \int_A f(x) \, \mathrm{d}\mu
\end{aligned}$$

In the first and last equality we have again used the dominated convergence theorem. The statement follows from the definition of has_cond_exp. □

At one point in the proof of our lemma, we have used the fact that a convergent sequence in L^1 admits a subsequence which is convergent in the underlying norm μ-a.e. This result is stated in Isabelle as follows

```
proposition tendsto_L1_AE_subseq:
  fixes u :: "nat ⇒' a ⇒' b"
  assumes "⋀n. integrable M (u n)"
      and "(λn. (∫ norm (u n x) ∂M)) ⟶ 0"
  shows "∃r :: nat ⇒ nat. strict_mono r ∧ (AE x in M. (λn. u (r n) x) ⟶ 0)"
  ...
```

In our case, we can't easily formulate the convergence of $(\texttt{cond\_exp}\ M\ F\ s_n)_{n\in\mathbb{N}}$ in the L^1 norm in the manner stated above. Therefore we have introduced the following lemma which is more flexible. Mathematically, the underlying argument is the same.

Lemma 3.2.14

```
lemma cauchy_L1_AE_cauchy_subseq:
  fixes s :: "nat ⇒' a ⇒' b"
  assumes "⋀n. integrable M (s n)"
      and "⋀e. e > 0 ⟹ ∃N. ∀i ≥ N.∀j ≥ N. LINT x|M. norm (s i x − s j x) < e"
  obtains r where "strict_mono r" "AE x in M. Cauchy (λi. s (r i) x)"
  ...
```

The main result of this subsection is formalized in Isabelle as follows

Corollary 3.2.15

```
corollary has_cond_expI:
  assumes "integrable M f"
  shows "has_cond_exp M F f (cond_exp M F f)"
proof -
  obtain s where s_is: "⋀i. simple_function M (s i)"
                       "⋀i. emeasure M {y ∈ space M. s i y ≠ 0} ≠ ∞"
                       "⋀x. x ∈ space M ⟹ (λi. s i x) ⟶ f x"
                       "⋀x i. x ∈ space M ⟹ norm (s i x) ≤ 2 · norm (f x)"
             using integrable_implies_simple_function_sequence[OF assms] by blast
  show ?thesis using has_cond_exp_simple_lim[OF assms s_is] has_cond_exp_charact(1) by metis
qed
```

3.2.3 Properties of the Conditional Expectation

Identity on F-measurable functions

If an integrable function f is already F-measurable, then cond_exp M F $f = f$ μ-a.e. This is a corollary of the lemma on the characterization of cond_exp.

Corollary 3.2.16

```
corollary cond_exp_F_meas:
  assumes "integrable M f"
          "f ∈ borel_measurable F"
  shows "AE x in M. cond_exp M F f x = f x"
  by (rule cond_exp_charact, auto intro: assms)
```

Tower Property

The following property is called the *tower property* of the conditional expectation.

Lemma 3.2.17. *Let F and G be nested sub-σ-algebras, i.e. $F \subseteq G \subseteq \Sigma$. Then, for any $f \in L^1(E)$, we have*

$$\texttt{cond\_exp}\ M\ F\ (\texttt{cond\_exp}\ M\ G\ f) = \texttt{cond\_exp}\ M\ F\ f\ \mu\text{-a.e.}$$

Proof. For any $A \in F$, we have

$$\begin{aligned}\int_A \texttt{cond\_exp}\ M\ G\ f\ \mathrm{d}\mu &= \int_A f\ \mathrm{d}\mu \\ &= \int_A \texttt{cond\_exp}\ M\ F\ f\ \mathrm{d}\mu\end{aligned}$$

since A is also in G. The characterization lemma yields the result. □

Contractivity

A linear operator $L : V \to W$ between normed vector spaces V and W is called a *contraction* if its operator norm

$$\|L\|_{\texttt{op}} = \inf\{c \geq 0 : \|Lv\|_W \leq c\|v\|_V \text{ for all } v \in V\}$$

is less than or equal to 1. Such an operator always preserves limits and has other useful properties in functional analysis [Sz-+10].

Lemma 3.2.18. *(Contractivity)*
Let $f \in L^1(E)$. Then

$$\|\texttt{cond\_exp}\ M\ F\ f\| \leq \texttt{cond\_exp}\ M\ F\ (\lambda x. \|f(x)\|)$$

Proof. We have already shown contractivity in the case of simple functions. Since f is integrable, there exists a sequence of simple functions $(s_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} s_n = f \ \mu\text{-a.e.}$$

and

$$\|s_n(x)\| \leq 2 \cdot \|f(x)\| \ \mu\text{-a.e. for all } n \in \mathbb{N}$$

Using the results of the previous subsection, we obtain a subsequence $(s_{r_n})_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} (\texttt{cond\_exp}\ M\ F\ s_{r_n}) = \texttt{cond\_exp}\ M\ F\ f \quad \mu\text{-a.e.}$$

With the exact same arguments applied to the sequence of simple functions $(\lambda x. \|s_{r_n}(x)\|)_{n\in\mathbb{N}}$, we obtain a sub-subsequence $(s_{r_{r'_n}})_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} (\texttt{cond\_exp}\ M\ F\ (\lambda x. \|s_{r_{r'_n}}(x)\|)) = \texttt{cond\_exp}\ M\ F\ (\lambda x. \|f(x)\|) \quad \mu\text{-a.e.}$$

Furthermore, we have

$$\|(\texttt{cond\_exp}\ M\ F\ s_{r_{r'_n}})(x)\| \leq (\texttt{cond\_exp}\ M\ F\ (\lambda x. \|s_{r_{r'_n}}(x)\|))(x) \quad \mu\text{-a.e.}$$

for all $n \in \mathbb{N}$, since the functions in question are simple. Taking the limits on both sides and using the continuity of the norm yields the result. $\square$

Corollary 3.2.19. *The linear operator* $\texttt{cond\_exp}\ M\ F : L^1(E) \to L^1(E)$ *is a contraction.*

Proof. Let $f \in L^1(E)$. From the previous lemma we have

$$\begin{aligned}
\|\texttt{cond_exp}\ M\ F\ f\|_1 &= \int \|\texttt{cond_exp}\ M\ F\ f\| \,\mathrm{d}\mu \\
&\leq \int \texttt{cond_exp}\ M\ F\ (\lambda x. \|f(x)\|) \,\mathrm{d}\mu \\
&= \int \|f\| \mathrm{d}\mu = \|f\|_1
\end{aligned}$$

Hence $\|\texttt{cond\_exp}\ M\ F\|_{\texttt{op}} \leq 1$ $\square$

Pulling Out What's Known

The following property of the conditional expectation is called "pulling out what's known".

Lemma 3.2.20. *Let $f : \Omega \to \mathbb{R}$ be an F-measurable function. Let $g \in L^1(E)$ and $f \cdot g \in L^1(E)$. Then*

$$\texttt{cond\_exp}\ M\ F\ (f \cdot g) = f \cdot \texttt{cond\_exp}\ M\ F\ g \quad \mu\text{-a.e.}$$

Proof. The proof of this lemma is involved as well. Therefore we will only focus on the core idea of the proof. We will also assume that the result already holds in the real setting. We show the following seemingly less general statement for $z : \Omega \to \mathbb{R}$ F-measurable and $z \cdot g \in L^1(E)$:

$$\int z \cdot g \; \mathrm{d}\mu = \int z \cdot \texttt{cond\_exp}\ M\ F\ g \; \mathrm{d}$$

The result will follow by taking $z = f \cdot \mathbf{1}_A$ for $A \in F$. Since z is measurable, there exists some sequence of simple functions $(s_n)_{n \in \mathbb{N}}$ such that

$$\lim_{n \to \infty} s_n = z \ \mu\text{-a.e.}$$

and

$$|s_n(x)| \leq 2 \cdot |z(x)| \ \mu\text{-a.e. for all } n \in \mathbb{N}$$

In this case one can easily check that

$$\int s_n \cdot g \; \mathrm{d}\mu = \int s_n \cdot \texttt{cond\_exp}\ M\ F\ g \; \mathrm{d}$$

for all $n \in \mathbb{N}$

By our additional assumption that the result already holds in the real case, we have

$$|z \cdot \texttt{cond\_exp}\ M\ F\ (\lambda x. \|g(x\|))| = \texttt{cond\_exp}\ M\ F\ (\lambda x. |z(x) \cdot \|g(x)\||)$$

Using the contractivity of the conditional expectation and the above bound on s_n, it follows that

$$\|s_n \cdot \texttt{cond\_exp}\ M\ F\ g\| \leq 2 \cdot \texttt{cond\_exp}\ M\ F\ (\lambda x. |z(x) \cdot \|g(x)\||)$$

Applying the dominated convergence theorem twice, we get

$$\lim_{n \to \infty} \int s_n \cdot g \; \mathrm{d}\mu = \int z \cdot g \; \mathrm{d}\mu$$

and

$$\lim_{n \to \infty} \int s_n \cdot \texttt{cond\_exp}\ M\ F\ g \; \mathrm{d}\mu = \int z \cdot \texttt{cond\_exp}\ M\ F\ g \; \mathrm{d}\mu$$

Since the sequence on the right hand side are equal, the statement follows from the fact that limits are unique. ☐

3.3 Conditional Expectation on Linearly Ordered Banach Spaces

In the presence of an ordering, we can prove certain monotonicity properties of the conditional expectation. We start with the following two lemmas

Lemma 3.3.1. *Let $f \in L^1(E)$. Assume $f \geq c$ μ-a.e. for some $c \in E$. Then*

$$\texttt{cond\_exp}\ M\ F\ f \geq c \quad \mu\text{-a.e.}$$

Proof. We will show the statement using the averaging theorem. Let $A \in F$ be a measurable set with $\mu(A) < \infty$. Then

$$\begin{aligned} c &= \frac{1}{\mu(A)} \int_A c \, \mathrm{d}\mu \\ &\leq \frac{1}{\mu(A)} \int_A f \, \mathrm{d}\mu \\ &= \frac{1}{\mu(A)} \int_A \texttt{cond\_exp}\ M\ F\ f \, \mathrm{d}\mu \\ &= \frac{1}{\mu(A)} \int_A \texttt{cond\_exp}\ M\ F\ f \, \mathrm{d}\mu|_F \end{aligned}$$

Hence $\int_A \texttt{cond\_exp}\ M\ F\ f \, \mathrm{d}\mu|_F \in \{x \in E \mid x \geq c\}$. The statement follows from the fact that $\{x \in E \mid x \geq c\}$ is closed. □

Lemma 3.3.2. *Let $f \in L^1(E)$. Assume $f > c$ μ-a.e. for some $c \in E$. Then*

$$\texttt{cond\_exp}\ M\ F\ f > c \quad \mu\text{-a.e.}$$

Proof. The averaging theorem is not applicable in this case since $\{x \in E \mid x > c\}$ is not closed. Therefore, we argue as follows.

Let $S = \{\texttt{cond\_exp}\ M\ F\ f \leq c\}$. The conditional expectation $\texttt{cond\_exp}\ M\ F\ f$ is F-measurable, hence $S \in F$. Since F is a σ-finite sub-σ-algebra, we can assume without loss of generality that $\mu(S) < \infty$. The assumption $f > c$ μ-a.e. implies

$$\int_S f \, \mathrm{d}\mu \geq \int_S c \, \mathrm{d}\mu$$

Furthermore, by the definition of S

$$\begin{aligned} \int_S c \, \mathrm{d}\mu &\geq \int_S \texttt{cond\_exp}\ M\ F\ f \, \mathrm{d}\mu \\ &= \int_S f \, \mathrm{d}\mu \end{aligned}$$

Hence $\int_S f \, \mathrm{d}\mu = \int_S c \, \mathrm{d}\mu$. By Corollary 3.1.16, we have

$$\mathbf{1}_S \cdot f = \mathbf{1}_S \cdot c \quad \mu\text{-a.e.}$$

Because of our assumption $f > c$ μ-a.e., this can only be the case if S is a μ-null set, which completes the proof. □

The corresponding lemmas for $(\leq)$ and $(<)$ are simple corollaries.

Corollary 3.3.3

```
corollary cond_exp_le_c:
  assumes "integrable M f" "AE x in M. f x ≤ c"
  shows "AE x in M. cond_exp M F f x ≤ c"
  ...
```

Corollary 3.3.4

```
corollary cond_exp_less_c:
  assumes "integrable M f" "AE x in M. f x < c"
  shows "AE x in M. cond_exp M F f x < c"
  ...
```

Finally, we can demonstrate the operator's monotonicity and that it preserves the pointwise maximum and minimum of two integrable functions.

Corollary 3.3.5

```
corollary cond_exp_mono:
  assumes "integrable M f" "integrable M g" "AE x in M. f x ≤ g x"
  shows "AE x in M. cond_exp M F f x ≤ cond_exp M F g x"
  using cond_exp_le_c[OF Bochner_Integration.integrable_diff, OF assms(1,2), of 0]
        cond_exp_diff[OF assms(1,2)] assms(3) by auto
```

Corollary 3.3.6

```
corollary cond_exp_max:
  assumes "integrable M f" "integrable M g"
  shows "AE x in M. cond_exp M F (λx. max (f x) (g x)) x = max (cond_exp M F f x) (cond_exp M F g x)"
  ...
```

Corollary 3.3.7

```
corollary cond_exp_min:
  assumes "integrable M f" "integrable M g"
  shows "AE x in M. cond_exp M F (λx. min (f x) (g x)) x = min (cond_exp M F f x) (cond_exp M F g x)"
  ...
```

Apart from some auxillary lemmas that we left out on purpose, this wraps up our overview of the formalization of the conditional expectation operator.

4 Stochastic Processes

It wouldn't make sense to talk about martingales without introducing stochastic processes first. In standard terminology, a stochastic process is a collection of random variables defined on the same probability space. The indexing set often represents time, and each random variable in the collection corresponds to an outcome at a specific point in that set. These processes are fundamental in understanding how random processes evolve over time.

Take the example of stock price movement, where each day's stock price is a random variable influenced by a variety of uncertain factors. This sequence of prices forms a stochastic process, describing the stock's behavior. Another instance is the Poisson process, which models events like customer arrivals at a service center. This process captures the randomness in the timing of arrivals, aiding in optimizing resource allocation and enhancing customer service. In physics, Brownian motion characterizes the unpredictable and continuous trajectory followed by particles suspended in a medium due to random collisions with surrounding molecules, which is again modelled as a stochastic process. The theory of stochastic processes is the cornerstone for analysing randomness and building models that mirror real-world uncertainties.

Keeping this in consideration, we aim to build a comprehensive foundation for a theory of stochastic process in Isabelle. Since the definition is so straightforward, it usually suffices to just consider a collection of measurable functions to make formal statements about stochastic processes. There is not much to gain from making an explicit definition on its own. Nonetheless, we must create a framework to discuss stochastic processes that can afterwards be broadened to formalize concepts like adaptedness and predictability. Locales present themselves as the solution we are looking for.

The locale system in Isabelle is useful for managing large formal developments, as it promotes modularity and reusability. It allows us to define generic theorems and structures in one place and then reuse them in multiple contexts without duplicating efforts. For instance, when defining filtered measure spaces in the following section, we will need to have an element act as the de facto bottom element of an index type. Locales allow us to easily fix such an element for this purpose.

We start with the following locale definition.

Definition 4.0.1

```
locale stochastic_process =
  fixes M t₀
    and X :: "'b :: {second_countable_topology, linorder_topology} ⇒ 'a ⇒ 'c"
  assumes random_variable[measurable]: "⋀i. t₀ ≤ i ⟹ X i ∈ borel_measurable M"
```

The measure M represents the underlying measure space on which the stochastic process is defined. The index t_0 represents the initial point in time beyond which the process X should be defined. As such, this locale formalizes a stochastic process defined on the interval $[t_0, \infty)$.

We have the following lemmas to introduce "constant" stochastic processes.

Lemma 4.0.2

```
lemma stochastic_process_const_fun:
  assumes "f ∈ borel_measurable M"
  shows "stochastic_process M t₀ (λ_. f)"
  using assms by (unfold_locales)

lemma stochastic_process_const:
  shows "stochastic_process M t₀ (λi _. c i)"
  by (unfold_locales) simp
```

For sake of completeness, we also provide the following collection of lemmas which show that stochastic processes are stable under various operations on a vector space.

Lemma 4.0.3

```
lemma compose:
  assumes "⋀i. t₀ ≤ i ⟹ f i ∈ borel_measurable borel"
  shows "stochastic_process M t₀ (λi x. (f i) (X i x))"
  by (unfold_locales) (intro measurable_compose[OF random_variable assms])
```

Lemma 4.0.4

```
lemma norm: "stochastic_process M t₀ (λi x. norm (X i x))"
  by (fastforce intro: compose)
```

Lemma 4.0.5

```
lemma scaleR_right:
  assumes "stochastic_process M t₀ Y"
  shows "stochastic_process M t₀ (λi x. (Y i x) *R (X i x))"
  using stochastic_process.random_variable[OF assms] random_variable
  by (unfold_locales) simp
```

```
lemma scaleR_right_const_fun:
  assumes "f ∈ borel_measurable M"
  shows "stochastic_process M t₀ (λi x. f x *R (X i x))"
  by (unfold_locales) (intro borel_measurable_scaleR assms random_variable)

lemma scaleR_right_const: "stochastic_process M t₀ (λi x. c i *R (X i x))"
  by (unfold_locales) simp
```

Lemma 4.0.6

```
lemma add:
  assumes "stochastic_process M t₀ Y"
  shows "stochastic_process M t₀ (λi x. X i x + Y i x)"
  using stochastic_process.random_variable[OF assms] random_variable
  by (unfold_locales) simp
```

Lemma 4.0.7

```
lemma diff:
  assumes "stochastic_process M t₀ Y"
  shows "stochastic_process M t₀ (λi x. X i x - Y i x)"
  using stochastic_process.random_variable[OF assms] random_variable
  by (unfold_locales) simp
```

Lemma 4.0.8

```
lemma uminus: "stochastic_process M t₀ (-X)" using scaleR_right_const[of "λ_. -1"]
    by (simp add: fun_Compl_def)
```

Lemma 4.0.9

```
lemma partial_sum: "stochastic_process M t₀ (λn x. ∑i∈{t₀..<n}. X i x)"
    by (unfold_locales) simp

lemma partial_sum': "stochastic_process M t₀ (λn x. ∑i∈{t₀..n}. X i x)"
    by (unfold_locales) simp
```

Lemma 4.0.10

```
lemma stochastic_process_sum:
  assumes "⋀i. i ∈ I ⟹ stochastic_process M t₀ (X i)"
  shows "stochastic_process M t₀ (λk x. ∑i∈I. X i k x)"
  using assms[THEN stochastic_process.random_variable] by (unfold_locales, auto)
```

We also specify the following sublocales to easily make statements about discrete-time and continuous-time stochastic processes.

Definition 4.0.11

```
locale nat_stochastic_process = stochastic_process M "0 :: nat" X for M X
locale real_stochastic_process = stochastic_process M "0 :: real" X for M X
```

By explicitly designating an element t_0 to be the bottom element, we can formalize continuous-time stochastic processes, i.e. $(X_t)_{t \in \mathbb{R}_{\geq 0}}$, without the need for introducing a new type for non-negative real numbers.

Remark. Moving forward, we will define the concepts of adaptedness, progressive measurability and predictability. In our formalization, we have introduced analogous lemmas and sublocales for these process varieties as well. To avoid repeating ourselves, we will only reiterate these statements, if the proofs become non-trivial or if the assumptions change.

Before presenting the remaining process varities, we must introduce the concept of a filtered measure space.

4.1 Filtered Measure Spaces

A filtered measure space is a measure space equipped with a sequence of increasing sub-σ-algebras, called a *filtration* that represents the accumulation of information over time.

Concretely, let M be a measure space. Assume we have a sequence of sigma-algebras $(F_n)_{n \in \mathbb{N}}$ where

$$F_0 \subseteq F_1 \subseteq F_2 \subseteq \dots$$

This sequence forms a filtration on M. Intuitively, each F_n represents the information available up to time n. In general, the index set does not need to be countable. We only need it to be totally ordered, so that two indices are always comparable with one another. In Isabelle, we define the following locale to capture this concept.

Definition 4.1.1

```
locale filtered_measure =
  fixes M F and t₀ :: "'b :: {second_countable_topology, linorder_topology}"
  assumes subalgebra: "⋀i. t₀ ≤ i ⟹ subalgebra M (F i)"
    and sets_F_mono: "⋀i j. t₀ ≤ i ⟹ i ≤ j ⟹ sets (F i) ⊆ sets (F j)"
```

with the predicate subalgebra in HOL-Probability.Conditional_Expectation defined via

```
definition subalgebra::"'a measure ⇒' a measure ⇒ bool" where
  "subalgebra M F = ((space F = space M) ∧ (sets F ⊆ sets M))"
```

Remark. In Isabelle the `measure` type is used to represent both measure spaces and σ-algebras. The latter is achieved by only considering the underlying σ-algebra via the projection `sets`.

In general, a type with an ordering does not necessarily inhabit a bottom element, i.e. an element that is lesser than any other element. In the next section, we will see how the existence of a bottom element lets us make easy statements about what constitutes an adapted process and what not. From a practical point of view, this is not too much to assume, since all random processes one encounters in the real world must start at some fixed point in time (or at least that assumption can be made for practical purposes).

The keen reader might have noticed that we need a little bit more to define martingales properly. Namely, the sub-σ-algebras that comprise the filtration $(F_n)_{n\in\mathbb{N}}$ must all be σ-finite. Otherwise, we can't make use of our lemmas concerning the conditional expectation. We introduce the following locale to adress this issue.

Definition 4.1.2

```
locale sigma_finite_filtered_measure = filtered_measure +
  assumes sigma_finite: "sigma_finite_subalgebra M (F t₀)"
```

Remark. Since we artifically designated an element t_0 to represent the least index in consideration, we only need to show σ-finiteness for the sub-σ-algebra F_{t_0}. σ-finiteness of all other sub-σ-algebras follows from the monotonicity of the filtration.

For the sake of completeness, we also introduce a local covering the case where the measure space is finite.

Definition 4.1.3

```
locale finite_filtered_measure = filtered_measure + finite_measure

sublocale finite_filtered_measure ⊆ sigma_finite_filtered_measure
  using subalgebra by (unfold_locales,
                       blast,
                       meson dual_order.refl finite_measure_axioms finite_measure_def
                             finite_measure_restr_to_subalg subalgebra
                             sigma_finite_measure.sigma_finite_countable)
```

In order to make the ideas in this section a bit more concrete, we present the following two filtrations as examples.

Lemma 4.1.4

```
lemma filtered_measure_constant_filtration:
  assumes "subalgebra M F"
  shows "filtered_measure M (λ_. F) t₀"
  using assms by (unfold_locales) (auto simp add: subalgebra_def)

sublocale sigma_finite_subalgebra ⊆ constant_filtration:
        sigma_finite_filtered_measure M "(λ_. F)" t₀
  using subalg by (unfold_locales) (auto simp add: subalgebra_def)
```

If we have some sub-σ-algebra $F \subseteq \Sigma$, then we can trivially take as our filtration $F_i = F$ for all $i \in [t_0, \infty)$. If we additionally know that we are working with a σ-finite subalgebra, then this yields a trivial σ-finite filtration on M. This choice of filtration is called a **constant filtration**.

Remark. In the above lemma, both entries convey the same information. The first one is stated in terms of premises and results, the latter in the language of locales. The notion of a σ-algebra being a subalgebra is formalized via the predicate `subalgebra`. Had the formalization been done in the language of locales, we could replace the first statement with an equivalent sublocale relation.

Preparing for our next example, we introduce a formalization for the notion of a σ-algebra generated by a family of functions.

Definition 4.1.5

```
definition sigma_gen :: "'a set ⇒ 'b measure ⇒ ('a ⇒ 'b) set ⇒ 'a measure" where
  "sigma_gen Ω N S ≡ sigma Ω (⋃f ∈ S. {(f -` A) ∩ Ω | A. A ∈ N})"
```

Given two measure spaces $(V, \mathcal{A})$ and $(W, \mathcal{B})$, it is a well known fact that a function $f : V \to W$ is measurable, if and only if the generated σ-algebra $\sigma(f)$ is a subalgebra of $\mathcal{A}$. This result is captured for families of functions in the following lemma.

Lemma 4.1.6

```
lemma measurable_family_iff_contains_sigma_gen:
  shows "(S ⊆ M →M N) ⟷ S ⊆ (space M → space N) ∧ sigma_gen (space M) N S ⊆ M"
  ...
```

Now, we can introduce our more interesting example, the **natural filtration**.

Definition 4.1.7

```
definition natural_filtration :: "'a measure ⇒ 'b ⇒ ('b ⇒ 'a ⇒ 'c) ⇒ 'b ⇒ 'a measure" where
  "natural_filtration M t₀ Y = (λt. sigma_gen (space M) borel {Y i | i. i ∈ {t₀..t}})"
```

The natural filtration with respect to a stochastic process Y is the filtration generated by all events involving the process up to the time index t, i.e. $F_t = \sigma(\{Y_i \mid i.\ i_{,} \le t\})$. Assuming that Y is a stochastic process, i.e. Y_i is M-measurable for all $i \ge t_0$, the definition indeed provides a filtration. The following sublocale relation formalizes this.

Lemma 4.1.8

```
sublocale stochastic_process ⊆ filtered_measure_natural_filtration:
    filtered_measure M "natural_filtration M t₀ X" t₀
    by (unfold_locales) (intro subalgebra_natural_filtration,
                         simp only: sets_natural_filtration,
                         intro sigma_sets_subseteq, force)
```

The natural filtration contains information concerning the process's past behavior at each point in time. The natural filtration is essentially the simplest filtration for studying a process. However, the natural filtration is not always σ-finite. In order to show that the natural filtration constitutes a sigma finite filtered measure, we need to provide a countable exhausting set in the preimage of X_{t_0}.

Lemma 4.1.9

```
lemma (in sigma_finite_measure) sigma_finite_filtered_measure_natural_filtration:
  assumes "stochastic_process M t₀ X"
  and exhausting_set: "countable A" "(⋃ A) = space M"
          "⋀a. a ∈ A ⟹ emeasure M a ≠ ∞"
          "⋀a. a ∈ A ⟹ ∃b ∈ borel. a = ((X t₀) -` b) ∩ space M"
  shows "sigma_finite_filtered_measure M (natural_filtration M t₀ X) t₀"
  ...
```

This concludes our development of filtered measure spaces.

4.2 Adapted Processes

We call a stochastic process $(X_t)_{t\in[t_0,\infty)}$ is *adapted to the filtration* $(F_t)_{t\in[t_0,\infty)}$ if, for every index $t \ge t_0$, the random variable X_t is measurable with respect to the σ-algebra F_t. This means that the value of X_t depends only on the information available up to time t. In other words, the process "adapts" to the information in a way that it cannot anticipate future values based on events that have not occurred yet. We introduce the following locale.

Lemma 4.2.1

```
locale adapted_process = filtered_measure M F t₀ for M F t₀
    and X :: "_ ⇒ _ ⇒ _ :: {second_countable_topology, banach}" +
  assumes adapted[measurable]: "⋀i. t₀ ≤ i ⟹ X i ∈ borel_measurable (F i)"
```

The properties we have shown concerning stochastic processes also hold for adapted processes. Although in some cases, for example in the following statement, we need to modify the measurability assumptions we make. Here, we see how constraining ourselves to an index set bounded from below helps make the assumption simpler.

Lemma 4.2.2

```
lemma (in filtered_measure) adapted_process_const_fun:
  assumes "f ∈ borel_measurable (F t₀)"
  shows "adapted_process M F t₀ (λ_. f)"
  ...
```

Furthermore, in the presence of a discrete index set, we have the following additional lemma concerning partial sums.

Lemma 4.2.3

```
lemma (in nat_adapted_process) partial_sum_Suc:
    "nat_adapted_process M F (λn x.∑i < n. X (Suc i) x)"
proof (unfold_locales)
  fix i
  have "X j ∈ borel_measurable (F i)" if "j ≤ i" for j using that adaptedD by blast
  thus "(λx. ∑i < n. X (Suc i) x) ∈ borel_measurable (F i)" by auto
qed
```

An adapted process is necessarily a stochastic process. This follows directly from the fact that $F_t \subseteq \Sigma$ for all $t \geq t_0$.

Lemma 4.2.4

```
sublocale adapted_process ⊆ stochastic_process
  using measurable_from_subalg subalgebra adapted by (unfold_locales) blast

sublocale nat_adapted_process ⊆ nat_stochastic_process ..
sublocale real_adapted_process ⊆ real_stochastic_process ..
```

In the other direction, a stochastic process is always adapted to the natural filtration it generates.

Lemma 4.2.5

```
sublocale stochastic_process ⊆ adapted_natural:
    adapted_process M "natural_filtration M t₀ X" t₀ X
  by (unfold_locales) (auto simp add: natural_filtration_def
                             intro: random_variable measurable_sigma_gen)
```

Adapted processes are crucial for defining martingales. A martingale is by definition an adapted process. In the following section, we will explore progressively measurable processes, even though they are not directly relevant to our formalization of martingales. This serves two purposes: first, to replicate the corresponding results on mathlib, and second, to establish a solid foundation for future projects to build upon.

4.3 Progressively Measurable Processes

The definition of a progressively measurable process is more intricate.

Definition 4.3.1. Let $(F_t)_{t\in[t_0,\infty)}$ be a filtration of the measure space M. A stochastic process $(X_t)_{t\in[t_0,\infty)}$ is called progressively measurable (or simply *progressive*) if, for every index $t \geq t_0$, the map $[t_0, t] \times \Omega \to E$ defined by $(i, w) \mapsto X_i(w)$ is measurable with respect to the σ-algebra $\mathcal{B}([t_0, t]) \otimes F_t$. Here $\mathcal{B}([t_0, t])$ denotes the Borel σ-algebra on $[t_0, t]$ induced by the order topology.

The formalized version is as follows.

Definition 4.3.2

```
locale progressive_process = filtered_measure M F t₀ for M F t₀
    and X :: "_ ⇒ _ ⇒ _ :: {second_countable_topology, banach}" +
  assumes progressive[measurable]: "⋀t. t₀ ≤ t
     ⟹ (λ(i,x). X i x) ∈ borel_measurable (restrict_space borel {t₀..t} ⊗M (F t))"
```

Notice that the measurability assumption we make here is on the entire map $(i, w) \mapsto X_i(w)$ instead of being "pointwise" as in the previous two sections. As a side effect, the stochastic process defined by $X_i = c(i)$ for some $c : [t_0, \infty) \to E$ is progressively measurable, only if the function c is Borel measurable. Previously, this assumption was not required.

Lemma 4.3.3

```
lemma (in filtered_measure) progressive_process_const:
  assumes "c ∈ borel_measurable borel"
  shows "progressive_process M F t₀ (λi _. c i)"
  using assms by (unfold_locales)
                 (auto simp add: measurable_split_conv
                  intro!: measurable_compose[OF measurable_fst]
                  measurable_restrict_space1)
```

Similarly, we must modify the premise of the lemma `compose` in order to reflect this change.

Lemma 4.3.4

```
lemma compose:
  assumes "(λ(i, x). f i x) ∈ borel_measurable borel"
  shows "progressive_process M F t0 (λi x. (f i) (X i x))"
  ...
```

A progressively measurable process is necessarily adapted. The proof is trivial and arises from the fact that the injection $y \mapsto (t, y)$ is measurable as a function $\Omega \to [t_0, t] \times \Omega$ for fixed $t \geq t_0$.

Lemma 4.3.5

```
sublocale progressive_process ⊆ adapted_process
  using measurable_compose_rev[OF progressive measurable_Pair1']
  unfolding prod.case by (unfold_locales) simp
```

On a more interesting note, progressive measurability is equivalent to adaptedness in the discrete setting. The following lemma demonstrates this.

Lemma 4.3.6. *Let $(X_i)_{i\in\mathbb{N}}$ be an adapted process with respect to the filtration $(F_i)_{i\in\mathbb{N}}$. Then it is also progressively measurable.*

Proof. Let S be an open set in E. Then $X_j^{-1}(S) \in F_i$ for all $j \leq i \in \mathbb{N}$, since $(X_i)_{i\in\mathbb{N}}$ is adapted by assumption. Let $\psi : \{0, \ldots, i\} \times \Omega \to E$ with $\psi(j, x) = X_j(x)$. Then, we have

$$\psi^{-1}(S) \cap \{j\} \times \Omega = \{j\} \times X_j^{-1}(S) \in \mathcal{B}(\{0, \ldots, i\}) \otimes F_i$$

since the order topology on $\mathbb{N}$ is discrete. Furthermore

$$\psi^{-1}(S) = \bigcup_{j\leq i} \psi^{-1}(S) \cap \{j\} \times \Omega$$

Since the set $\{0, \ldots, i\}$ is countable, it follows that $\psi^{-1}(S) \in \mathcal{B}(\{0, \ldots, i\}) \otimes F_i$, since it is expressable as the union of a countable family of measurable sets. □

Subsequently we express this fact in the language of locales.

Lemma 4.3.7

```
sublocale nat_adapted_process ⊆ nat_progressive_process
  ...
```

Now comes the most challenging portion of this chapter.

4.4 Predictable Processes

Before defining predictable processes in full generality, we will introduce them in the discrete setting, where the definition is easier to grasp.

Definition 4.4.1. A discrete-time stochastic process $(X_i)_{i\in\mathbb{N}}$ is called *predictable* with respect to a filtration $(F_i)_{i\in\mathbb{N}}$, if X_{i+1} is F_i-measurable for all $i \in \mathbb{N}$.

This means that the value of the process in the future, X_{i+1}, can be "predicted" using the information available up to time i. This definition is a special case of the following more general definition for arbitrary index sets.

Definition 4.4.2. Let $(F_t)_{t\in[t_0,\infty)}$ be a filtration of the measure space M. We define the *predictable σ-algebra* Σ_P as follows.

$$\Sigma_P = \sigma(\{(s,t] \times A \mid A \in F_s \ \wedge\ t_0 \le s \ \wedge\ s < t\} \cup \{\{t_0\} \times A \mid A \in F_{t_0}\})$$

A stochastic process $(X_t)_{t\in[t_0,\infty)}$ is called *predictable* if the map $[t_0,\infty) \times \Omega \to E$ defined by $(t,x) \mapsto X_t(x)$ is measurable with respect to this σ-algebra.

At first glance, it is difficult to make intuitive sense of this definition. Investigating properties of predictable processes in arbitrary settings is well beyond the scope of this thesis. However, we will make the following remark.

Remark. One can show that the σ-algebra Σ_P coincides with the σ-algebra generated by all left-continuous adapted processes. A stochastic process is called left-continuous, if the sample paths $t \mapsto X_t(x)$ are left-continuous for μ-almost all $x \in \Omega$. Right-continuity is similarly defined.

The corresponding locale is easy to define.

Definition 4.4.3

```
locale predictable_process = filtered_measure M F t₀ for M F t₀
   and X :: "_ ⇒ _ ⇒ _ :: {second_countable_topology, banach}" +
  assumes progressive[measurable]: "(λ(t,x). X t x) ∈ borel_measurable Σ_P"
```

In the previous section, our results concerning progressively measurable processes all made use of the fact that the projection functions $\pi_1 : [t_0,\infty) \times \Omega \to [t_0,\infty)$ and $\pi_2 : [t_0,\infty) \times \Omega \to \Omega$ are measurable with respect to the underlying σ-algebra. In that setting, this was a triviality, since the σ-algebra in question was the product σ-algebra $\mathcal{B}([t_0,t]) \otimes F_t$ for some $t \ge t_0$, which has many nice properties. We wish to show a similar statement for the projection functions π_i when $[t_0,\infty) \times \Omega$ is equipped with the σ-algebra Σ_P. We have come up with a sufficient condition on the index set $[t_0,\infty)$ that guarantees this.

Lemma 4.4.4. *Assume there exists some countable family of sets $\mathcal{I} \subseteq \{(s, t] \mid t_0 \leq s \wedge s < t\}$ such that $(t_0, \infty) \subseteq (\bigcup \mathcal{I})$. Let $\pi_1 : [t_0, \infty) \times \Omega \to [t_0, \infty)$ and $\pi_2 : [t_0, \infty) \times \Omega \to \Omega$ be projections onto respective components. Then, π_1 is Σ_P-Borel-measurable and π_2 is Σ_P-F_{t_0}-measurable.*

Proof. We first show that π_1 is Σ_P-Borel-measurable.

π_1 is trivially $(\mathcal{B}([t_0, \infty)) \otimes \sigma(\varnothing))$-Borel-measurable. Hence, if we can show

$$(\mathcal{B}([t_0, \infty)) \otimes \sigma(\varnothing)) \subseteq \Sigma_P$$

then this implies that π_1 is Σ_P-Borel-measurable. For this, we will show that the Borel σ-algebra $\mathcal{B}([t_0, \infty))$ coincides with the σ-algebra generated by the set $\{(s, t] \mid t_0 \leq s \wedge s < t\}$.

Since the ordering on $[t_0, \infty)$ is linear, the set of open rays $\{(s, \infty) \mid s \geq t_0\}$ generates the order topology on $\mathcal{B}([t_0, \infty))$. This also depends on the premise that the order topology is second-countable. Let $c \geq t_0$. We have

$$\begin{aligned}(c, \infty) &= (c, \infty) \cap (\bigcup \mathcal{I}) \\ &= (\bigcup_{I \in \mathcal{I}} I \cap (c, \infty))\end{aligned}$$

From the assumptions, we know that

$$I \cap (c, \infty) \in \{(s, t] \mid t_0 \leq s \wedge s < t\}$$

for $I \in \mathcal{I}$. Hence $(c, \infty) \in \sigma(\{(s, t] \mid t_0 \leq s \wedge s < t\})$, since $\mathcal{I}$ is countable. In the other direction, any interval $(s, t]$ with $s \geq t_0$ is obviously $\mathcal{B}([t_0, \infty))$-measurable. Thus, the σ-algebras indeed coincide. This completes the first part of the proof.

Next, we show that π_2 is Σ_P-F_{t_0}-measurable. Let $S \in F_{t_0}$. We have

$$\pi_2^{-1}(S) = [t_0, \infty) \times S$$

The assumptions already imply $(t_0, \infty) = (\bigcup \mathcal{I})$. Furthermore $S \in F_t$ for all $t \geq t_0$. Hence, we have

$$(t_0, \infty) \times S = (\bigcup \mathcal{I}) \times S \in \Sigma_P \text{ and } \{t_0\} \times S \in \Sigma_P$$

which together imply $[t_0, \infty) \times S \in \Sigma_P$. □

Remark. Our formal proof for the Σ_P-Borel-measurability of π_1 follows an alternative path to the one given here. The lemma `borel_Ioi` establishes that the Borel σ-algebra $\mathcal{B}$

on the entire space is generated by open rays. We then consider the restricted σ-algebra $\mathcal{B}([t_0, \infty))$, which is defined in Isabelle as

$$\mathcal{B}([t_0, \infty)) = \sigma\left(\{[t_0, \infty) \cap A \mid A \in \mathcal{B}\}\right)$$

Together with `borel_Ioi` this yields

$$\mathcal{B}([t_0, \infty)) = \sigma\left(\{[t_0, \infty) \cap A \mid A \in \sigma(\{(s, \infty) \mid s \in (-\infty, \infty)\})\}\right)$$

In our formalization, we show that this σ-algebra on the right hand side is equal to $\sigma(\{(s, t] \mid t_0 \le s \wedge s < t\})$.

On a different note, we believe strongly that π_2 is not Σ_P-F_t-measurable for $t > t_0$ in general, or at least our condition is not sufficient to show this. This stems from the fact that $\pi_2^{-1}(S) = [t_0, \infty) \times S$ and the element t_0 can only originate from some set in $\{\{t_0\} \times A \mid A \in F_{t_0}\}$. However, in general $F_t \not\subseteq F_{t_0}$.

In the discrete-time case, the family $\mathcal{I} = \{\{n+1\}\}_{n \in \mathbb{N}}$ fulfills this condition. Similarly, in the continous-time case we can use $\mathcal{I} = \{(0, n+1]\}\}_{n \in \mathbb{N}}$. The following lemmas in Isabelle reflect this.

Lemma 4.4.5

```
lemma (in nat_filtered_measure) measurable_predictable_sigma_snd:
  shows "snd ∈ Σ_P →_M F 0"
  by (intro measurable_predictable_sigma_snd[of "range (λx. {Suc x})"])
     (force | simp add: greaterThan_0)+

lemma (in nat_filtered_measure) measurable_predictable_sigma_fst:
  shows "fst ∈ Σ_P →_M borel"
  by (intro measurable_predictable_sigma_fst[of "range (λx. {Suc x})"])
     (force | simp add: greaterThan_0)+
```

Lemma 4.4.6

```
lemma (in real_filtered_measure) measurable_predictable_sigma_snd:
  shows "snd ∈ Σ_P →_M F 0"
  using real_arch_simple
  by (intro measurable_predictable_sigma_snd[of "range (λx :: nat. {0<..real (Suc x)})"])
     (fastforce intro: add_increasing)+

lemma (in real_filtered_measure) measurable_predictable_sigma_fst:
  shows "fst ∈ Σ_P →_M borel"
  using real_arch_simple
  by (intro measurable_predictable_sigma_fst[of "range (λx :: nat. {0<..real (Suc x)})"])
     (fastforce intro: add_increasing)+
```

These measurability results concerning projections are necessary to show the following statements about "constant" processes being predictable.

Lemma 4.4.7

```
lemma (in filtered_measure) predictable_process_const_fun:
  assumes "snd ∈ Σ_P →_M F t₀" "f ∈ borel_measurable (F t₀)"
  shows "predictable_process M F t₀ (λ_. f)"
  using measurable_compose_rev[OF assms(2)] assms(1)
  by (unfold_locales) (auto simp add: measurable_split_conv)
```

Lemma 4.4.8

```
lemma (in filtered_measure) predictable_process_const:
  assumes "fst ∈ borel_measurable Σ_P" "c ∈ borel_measurable borel"
  shows "predictable_process M F t₀ (λi_. c i)"
  using assms by (unfold_locales) (simp add: measurable_split_conv)
```

We will now show that a predictable process is necessarily progressively measurable.

Lemma 4.4.9. *A predictable process $(X_t)_{t\in[t_0,\infty)}$ is also progressively measurable.*

Proof. Let $i \geq t_0$. Let ι denote the identity function, restricted to the domain $[t_0, i] \times \Omega$, i.e. $\iota = \mathrm{id}|_{[t_0,t]}$. We aim to show that ι is $(\mathcal{B}([t_0, i]) \otimes F_i)$-$\Sigma_P$-measurable. The statement follows simply from definitions of predictability and progressive measurability.

For any S in the generating set of Σ_P, we will show that $\iota^{-1}(S) \in \mathcal{B}([t_0, i]) \otimes F_i$. This is enough to show the required measurability.

First, let $S = \{t_0\} \times A$ for some $A \in F_{t_0}$. Then

$$\iota^{-1}(S) = \{t_0\} \times A \in \mathcal{B}([t_0, i]) \otimes F_i$$

since $\{t_0\}$ is closed and $F_{t_0} \subseteq F_i$.

Next, let $S = (s, t] \times A$ for some $A \in F_s$ and $s, t \geq t_0$ with $s < t$. Then

$$\iota^{-1}(S) = (s, \min(i, t)] \times A$$

Assume $s \leq i$. Then $A \in F_i$. Furthermore, $(s, \min(i, t)] \in \mathcal{B}([t_0, i])$ since the σ-algebra $\mathcal{B}([t_0, i])$ is generated by half-open intervals. Hence, $\iota^{-1}(S) \in \mathcal{B}([t_0, i]) \otimes F_i$.

Assume $s > i$. Then $\iota^{-1}(S) = \varnothing \in \mathcal{B}([t_0, i]) \otimes F_i$. This covers all cases and the proof is complete. $\square$

We formalize this fact as a sublocale relation.

Lemma 4.4.10

```
sublocale predictable_process ⊆ progressive_process
  ...
```

In the scope of our thesis, we will only use results concerning discrete-time predictable processes. We will now show that the above definitions for predictable processes coincide when the index set is $\mathbb{N}$. First we show the following lemma.

Theorem 4.4.11. *Let* $(\bigcup_{i\in\mathbb{N}} \{i\} \times A_i) \in \Sigma_P$ *for some collection of sets* $(A_i)_{i\in\mathbb{N}}$*. Then* $A_0 \in F_0$ *and* $A_{i+1} \in F_i$ *for all* $i \in \mathbb{N}$*.*

Proof. Consider the set

$$\mathcal{D} = \{S \in \Sigma_P \mid \forall (A_i)_{i\in\mathbb{N}}.\ S = (\bigcup_{i\in\mathbb{N}} \{i\} \times A_i) \Longrightarrow A_{i+1} \in F_i \ \wedge \ A_0 \in F_0\}$$

We will show that $\mathcal{D}$ constitutes a σ-algebra. Obviously $\varnothing \in \mathcal{D}$.

Assume $S \in \mathcal{D}$.

Let $(A_i)_{i\in\mathbb{N}}$ be a family of sets with $(\mathbb{N} \times \Omega) \setminus S = (\bigcup_{i\in\mathbb{N}}\{i\} \times A_i)$. Then

$$\begin{aligned} S &= (\mathbb{N} \times \Omega) \setminus (\bigcup_{i\in\mathbb{N}} \{i\} \times A_i) \\ &= (\bigcup_{i\in\mathbb{N}} \{i\} \times \Omega) \setminus (\bigcup_{i\in\mathbb{N}} \{i\} \times A_i) \\ &= (\bigcup_{i\in\mathbb{N}} \{i\} \times (\Omega \setminus A_i)) \end{aligned}$$

Hence, we know $\Omega \setminus A_{i+1} \in F_i$ and $\Omega \setminus A_0 \in F_0$. Therefore, $A_{i+1} \in F_i$ and $A_0 \in F_0$. We have $(\mathbb{N} \times \Omega) \setminus S \in \mathcal{D}$.

Assume $S_i \in \mathcal{D}$ for $i \in \mathbb{N}$.

Let $(A_i)_{i\in\mathbb{N}}$ be a family of sets with $(\bigcup_{i\in\mathbb{N}} S_i) = (\bigcup_{i\in\mathbb{N}}\{i\} \times A_i)$. For each S_i, we need to find some family of sets $(B_j(i))_{j\in\mathbb{N}}$, such that $S_i = (\bigcup_{j\in\mathbb{N}}\{j\} \times B_j(i))$. Define

$$B_j(i) = \pi_2(S_i \cap \{j\} \times \Omega)$$

The intuition is as follows. We first select only those pairs in S_i, with the first component equal to j. Then we project onto the second component. Hence, we have the equality

$$\{j\} \times B_j(i) = S_i \cap \{j\} \times \Omega$$

Therefore $S_i = (\bigcup_{j\in\mathbb{N}}\{j\} \times B_j(i))$. We have $B_{j+1}(i) \in F_j$ and $B_0(i) \in F_0$. Furthermore, we know

$$\begin{aligned} A_i &= \pi_2\left((\bigcup_{j\in\mathbb{N}} S_j) \cap \{i\} \times \Omega \right) \\ &= \bigcup_{j\in\mathbb{N}} \pi_2(S_j \cap \{i\} \times \Omega) \\ &= \bigcup_{j\in\mathbb{N}} B_i(j) \end{aligned}$$

Hence, $A_{i+1} \in F_i$ and $A_0 \in F_0$. Thus $\mathcal{D}$ is indeed a σ-algebra. Now we show

$$\{\{s+1,\ldots,t\} \times A \mid A \in F_s \ \wedge\ s < t\} \cup \{\{0\} \times A \mid A \in F_0\} \subseteq \mathcal{D}$$

Let $S \in \{\{0\} \times A \mid A \in F_0\}$. Then $S = (\bigcup_{i\in\mathbb{N}}\{i\} \times A_i)$ implies $A_0 \in F_0$ and $A_i = \varnothing$ for $i > 0$. Hence $S \in \mathcal{D}$.

Let $S = \{s+1,\ldots,t\} \times B$ with $s < t$ and $B \in F_s$ for some s, t and B. Then, $S = (\bigcup_{i\in\mathbb{N}}\{i\} \times A_i)$ implies $A_i = B$ for $i \in \{s+1,\ldots,t\}$ and $A_i = \varnothing$ otherwise. Thus, $A_0 = \varnothing \in F_0$. Moreover, $A_{i+1} = B \in F_i$ if $i \in \{s,\ldots,t-1\}$ since the subalgebras F_i are nested, and $A_{i+1} = \varnothing \in F_i$ for $i \notin \{s,\ldots,t-1\}$. Together with our previous result, this implies $\Sigma_P \subseteq \mathcal{D}$, which completes the proof. □

Remark. For the proof of this lemma in Isabelle, we have used the induction scheme `sigma_sets.induct`, since the generated σ-algebra $\sigma(\cdot)$ is defined as an inductive set in Isabelle. The proof above demonstrates that this induction scheme is equivalent to the principle of "good-sets" which we have utilized.

We can now characterize predictability in the discrete-setting as follows.

Theorem 4.4.12. *A stochastic process $(X_n)_{n\in\mathbb{N}}$ is predictable, if and only if $(X_{n+1})_{n\in\mathbb{N}}$ is adapted to the filtration $(F_n)_{n\in\mathbb{N}}$ and X_0 is F_0-measurable.*

Proof. Assume $(X_n)_{n\in\mathbb{N}}$ is predictable. Since predictable processes are also adapted, X_0 is F_0-measurable.

Let $n \in \mathbb{N}$ and let S be an open set. Consider the map ψ defined by $\psi(i,x) = X_i(x)$. We have

$$\psi^{-1}(S) \cap (\{n+1\} \times \Omega) = \psi^{-1}(S) \cap ((n, n+1] \times \Omega) \in \Sigma_P$$

On the other hand

$$\psi^{-1}(S) \cap (\{n+1\} \times \Omega) = \{n+1\} \times X_{n+1}^{-1}(S)$$

Applying the previous lemma for

$$A_i = \begin{cases} X_{n+1}^{-1}(S) & \text{if } i = n+1 \\ \varnothing & \text{otherwise} \end{cases}$$

we get $X_{n+1}^{-1}(S) \in F_n$. Hence $(X_{n+1})_{n\in\mathbb{N}}$ is adapted to the filtration $(F_n)_{n\in\mathbb{N}}$.

For the other direction, assume $(X_{n+1})_{n\in\mathbb{N}}$ is adapted to the filtration $(F_n)_{n\in\mathbb{N}}$ and X_0 is F_0-measurable.

Let S be an open set. We have

$$\{0\} \times X_0^{-1}(S) \in \Sigma_P$$

using the definition of Σ_P and the fact that X_0 is F_0-measurable. Similarly, for $n \in \mathbb{N}$ we have $X_{n+1}^{-1}(S) \in F_n$. Hence

$$\{n+1\} \times X_{n+1}^{-1}(S) = (n, n+1] \times X_{n+1}^{-1}(S) \in \Sigma_P$$

Putting it all together, we have

$$\psi^{-1}(S) = \left(\bigcup_{i\in\mathbb{N}} \{i\} \times X_i^{-1}(S) \right) \in \Sigma_P$$

since Σ_P is closed under countable unions. Thus $(X_n)_{n\in\mathbb{N}}$ is predictable. □

This finalizes our formalization of various types of stochastic processes in Isabelle.

5 Martingales

In this section we will introduce and discuss martingales, the namesake of our thesis. Originally referring to a system of betting strategies, martingales have evolved far beyond their gambling origins and have found profound applications in various fields, including finance, probability theory, and statistical analysis. Our formalization aims for a high level of generality while maintaining clarity and simplicity, making it easier for future formalization efforts to build upon our foundation.

5.1 Fundamentals

Definition 5.1.1. Let $(F_t)_{t\in[t_0,\infty)}$ be a filtration of the measure space M. A stochastic process $(X_t)_{t\in[t_0,\infty)}$ taking values in a Banach space $(E, \|\cdot\|)$ is a *martingale with respect to the filtration* $(F_t)_{t\in[t_0,\infty)}$ if the following conditions hold

1. $(X_t)_{t\in[t_0,\infty)}$ is adapted to the filtration $(F_t)_{[t_0,\infty)}$,
2. $X_t \in L^1(E)$ for all $t \in [t_0, \infty)$,
3. $X_s = \mathbb{E}(X_t \mid F_s)$ μ-a.e. for all $s, t \in [t_0, \infty)$ with $s \leq t$.

Replacing "$=$" in the third condition with "$\leq$" or "$\geq$" gives rise to the definition of a sub- or supermartingale, respectively.

Using the results we have formalized in the previous chapters, we define the following locales.

Definition 5.1.2

```
locale martingale = sigma_finite_adapted_process +
  assumes integrable: "⋀i. t₀ ≤ i ⟹ integrable M (X i)"
      and martingale_property: "⋀i j. t₀ ≤ i ⟹ i ≤ j
     ⟹ AE x in M. X i x = cond_exp M (F i) (X j) x"
```

Remark. In addition to what we've discussed in the last chapter, we have introduced the locale `sigma_finite_adapted_process` which combines the locale `adapted_process` with the locale `sigma_finite_filtered_measure`. Without this additional restriction, we can't use the operator `cond_exp`. Similary, the locale `sigma_finite_adapted_process_order` places a restriction on the Banach space $(E, \|\cdot\|)$, asserting the existence of an ordering compatible with scalar multiplication. Finally, the locale `sigma_finite_adapted_process_linorder` further mandates that this ordering be total. We have also introduced locales for discrete-time and continuous-time counterparts.

Any stochastic process that is both a submartingale and a supermartingale is a martingale. Conversely, every martingale is also a submartingale and a supermartingale if there exists an ordering on the Banach space E. In anticipation of this result, we introduce the following locales.

Definition 5.1.3

```
locale martingale_order = martingale M F t₀ X for M F t₀
    and X :: "_ ⇒ _ ⇒ _ :: {order_topology, ordered_real_vector}"
locale martingale_linorder = martingale M F t₀ X for M F t₀
    and X :: "_ ⇒ _ ⇒ _ :: {linorder_topology, ordered_real_vector}"
sublocale martingale_linorder ⊆ martingale_order ..
```

Locales for submartingales and supermartingales are introduced similarly.

Definition 5.1.4

```
locale submartingale = sigma_finite_adapted_process_order +
  assumes integrable: "⋀i. t₀ ≤ i ⟹ integrable M (X i)"
      and submartingale_property: "⋀i j. t₀ ≤ i ⟹ i ≤ j
       ⟹ AE x in M. X i x ≤ cond_exp M (F i) (X j) x"

locale submartingale_linorder = submartingale M F t₀ X for M F t₀
    and X :: "_ ⇒ _ ⇒ _ :: {linorder_topology}"

sublocale martingale_order ⊆ submartingale using martingale_property
  by (unfold_locales) (force simp add: integrable)+
sublocale martingale_linorder ⊆ submartingale_linorder ..
```

Definition 5.1.5

```
locale supermartingale = sigma_finite_adapted_process_order +
  assumes integrable: "⋀i. t₀ ≤ i ⟹ integrable M (X i)"
      and supermartingale_property: "⋀i j. t₀ ≤ i ⟹ i ≤ j
       ⟹ AE x in M. X i x ≥ cond_exp M (F i) (X j) x"

locale supermartingale_linorder = supermartingale M F t₀ X for M F t₀
    and X :: "_ ⇒ _ ⇒ _ :: {linorder_topology}"

sublocale martingale_order ⊆ supermartingale using martingale_property
```

```
  by (unfold_locales) (force simp add: integrable)+
sublocale martingale_linorder ⊆ supermartingale_linorder ..
```

As noted before, a stochastic process taking values on an ordered Banach space is a martingale, if and only if it is both a submartingale and a supermartingale. The following lemma formalizes this fact.

Lemma 5.1.6

```
lemma martingale_iff:
  shows "martingale M F t0 X ⟷ submartingale M F t0 X ∧ supermartingale M F t0 X"
proof (rule iffI)
  assume asm: "martingale M F t0 X"
  interpret martingale_order M F t0 X by (intro martingale_order.intro asm)
  show "submartingale M F t0 X ∧ supermartingale M F t0 X"
    using submartingale_axioms supermartingale_axioms by blast
next
  assume asm: "submartingale M F t0 X ∧ supermartingale M F t0 X"
  interpret submartingale M F t0 X by (simp add: asm)
  interpret supermartingale M F t0 X by (simp add: asm)
  show "martingale M F t0 X"
    using submartingale_property supermartingale_property
    by (unfold_locales) (intro integrable, blast, force)
qed
```

Additionally, we have included lemmas for introducing martingales in simple cases. For $f \in L^1(E)$ and F_{t_0}-measurable, the constant stochastic process defined by $X_t = f$ is a martingale. The following lemma reflects this.

Lemma 5.1.7

```
lemma (in sigma_finite_filtered_measure) martingale_const_fun[intro]:
  assumes "integrable M f" "f ∈ borel_measurable (F t0)"
  shows "martingale M F t0 (λ_. f)"
  using assms sigma_finite_subalgebra.cond_exp_F_meas[OF _ assms(1), THEN AE_symmetric]
        borel_measurable_mono
  by (unfold_locales) blast+
```

The statements below follow directly.

Corollary 5.1.8

```
corollary (in sigma_finite_filtered_measure) martingale_zero[intro]:
    "martingale M F t0 (λ_ _. 0)" by fastforce

corollary (in finite_filtered_measure) martingale_const[intro]:
    "martingale M F t0 (λ_ _. c)" by fastforce
```

The stochastic process defined by $X_t = \mathbb{E}(f \mid F_t)$ for $f \in L^1(E)$ is also a martingale. This follows from the tower property of the conditional expectation.

Lemma 5.1.9

```
lemma (in sigma_finite_filtered_measure) martingale_cond_exp[intro]:
  assumes "integrable M f"
  shows "martingale M F t₀ (λi. cond_exp M (F i) f)"
  using sigma_finite_subalgebra.borel_measurable_cond_exp' borel_measurable_cond_exp
  by (unfold_locales) (auto intro: sigma_finite_subalgebra.cond_exp_nested_subalg[OF _ assms]
                             simp add: subalgebra_F subalgebra)
```

5.2 First Consequences, Basic Operations and Sufficient Conditions

First and foremost, we will discuss elementary properties of martingales, submartingales and supermartingales. Let $(X_t)_{t\in[t_0,\infty)}$ be a martingale with respect to the filtration $(F_t)_{t\in[t_0,\infty)}$. Using the martingale property and the characterization of the conditional expectation, we have the following lemma.

Lemma 5.2.1

```
lemma (in martingale) set_integral_eq:
  assumes "A ∈ F i" "t₀ ≤ i" "i ≤ j"
  shows "set_lebesgue_integral M A (X i) = set_lebesgue_integral M A (X j)"
```

This lemma already shows us the intuition behind the definition of a martingale. Let A be a set which is measurable at time i, i.e. some property of the process which we can inspect at time i. The average value that the process has on this set at time i is equal to the average value it will have on the same set at a future time j. Essentially, this is the reason why martingales are employed for modeling fair games that incorporate an element of chance. Similarly, for sub- and supermartingales we have the following lemmas.

Lemma 5.2.2

```
lemma (in submartingale) set_integral_le:
  assumes "A ∈ F i" "t₀ ≤ i" "i ≤ j"
  shows "set_lebesgue_integral M A (X i) ≤ set_lebesgue_integral M A (X j)"
```

Lemma 5.2.3

```
lemma (in supermartingale) set_integral_ge:
  assumes "A ∈ F i" "t₀ ≤ i" "i ≤ j"
  shows "set_lebesgue_integral M A (X i) ≥ set_lebesgue_integral M A (X j)"
```

In this case, the intuition is similar. The average value of a submartingale on a set which is measurable at time i is less than or equal to the average value it will take on the same set at a future time j. The case for a supermartingale is analogous. Here is a simple example illustrating this concept.

Example. Consider a coin-tossing game, where the coin lands on heads with probability $p \in [0, 1]$. Assume that the gambler wins a fixed amount $c > 0$ on a heads outcome and loses the same amount c on a tails outcome. Let $(X_n)_{n \in \mathbb{N}}$ be a stochastic process, where X_n denotes the gambler's fortune after the n-th coin toss. Then, we have the following three cases.

1. If $p = \frac{1}{2}$, it means the coin is fair and has an equal chance of landing heads or tails. In this case, the gambler, on average, neither wins nor loses money over time. The expected value of the gambler's fortune stays the same over time. Therefore, $(X_n)_{n \in \mathbb{N}}$ is a martingale.

2. If $p \leq \frac{1}{2}$, it means the coin is biased in favor of tails. In this scenario, the gambler is more likely to lose money on each bet because the probability of getting heads is lower than $\frac{1}{2}$. Over time, the gambler's fortune decreases on average. Therefore, $(X_n)_{n \in \mathbb{N}}$ is a supermartingale.

3. If $p \geq \frac{1}{2}$, it means the coin is biased in favor of heads. In this case, the gambler is more likely to win money on each bet. Over time, the gambler's fortune tends to increase on average. Therefore, $(X_n)_{n \in \mathbb{N}}$ is a submartingale.

5.3 Discrete-Time Martingales

6 Discussion

6.1 Formalization Approach

A convex function of a martingale is a submartingale, by Jensen's inequality. For example, the square of the gambler's fortune in the fair coin game is a submartingale (which also follows from the fact that Xn2 - n is a martingale). Similarly, a concave function of a martingale is a supermartingale. Stopped Processes... Wiener Processes

6.2 Comparison with Existing Formalizations

The following tables provide a list of the entries in the mathlib formalization of martingales, all of which have counterparts in our formalization.

Lean	Isabelle
`martingale`	`martingale (locale)`
`martingale.adapted`	`adapted_process.adapted`
`martingale.add`	`martingale.add`
`martingale.condexp_ae_eq`	`martingale.martingale_property`
`martingale.eq_zero_of_predictable`	`martingale.predictable_eq_zero`
`martingale.integrable`	`martingale.integrable`
`martingale.neg`	`martingale.uminus`
`martingale.set_integral_eq`	`martingale.set_integral_eq`
`martingale.smul`	`martingale.scaleR`
`martingale.strongly_measurable`	`stochastic_process.random_variable`
`martingale.sub`	`martingale.diff`
`martingale.submartingale`	via sublocale relation
`martingale.supermartingale`	via sublocale relation
`martingale_condexp`	`sigma_finite_filtered_measure.martingale_cond_exp`
`martingale_const`	`finite_filtered_measure.martingale_const`
`martingale_const_fun`	`sigma_finite_filtered_measure.martingale_const`
`martingale_iff`	`martingale_iff`
`martingale_nat`	`nat_sigma_finite_adapted_process.martingale_nat`
`martingale_of_condexp_sub_eq_zero_nat`	`nat_sigma_finite_adapted_process.martingale_of _cond_exp_diff_Suc_eq_zero`

martingale_of_set_integral_eq_succ	nat_sigma_finite_adapted_process.martingale_of_set_integral_eq_Suc
martingale_zero	sigma_finite_filtered_measure.martingale_zero

Table 6.1: Lookup table for martingale lemmas and definitions

Lean	Isabelle
submartingale	submartingale (locale)
submartingale.adapted	adapted_process.adapted
submartingale.add	submartingale.add
submartingale.add_martingale	submartingale.add
submartingale.ae_le_condexp	submartingale_property
submartingale.condexp_sub_nonneg	submartingale.cond_exp_diff_nonneg
submartingale.integrable	submartingale.integrable
submartingale.neg	submartingale.uminus
submartingale.pos	submartingale.max_0
submartingale.set_integral_le	submartingale_linorder.set_integral_le
submartingale.smul_nonneg	submartingale.scaleR_nonneg
submartingale.smul_nonpos	submartingale.scaleR_nonpos
submartingale.strongly_measurable	stochastic_process.random_variable
submartingale.sub_martingale	submartingale.diff
submartingale.sub_supermartingale	submartingale.diff
submartingale.sum_mul_sub	nat_submartingale.partial_sum_scaleR
submartingale.sum_mul_sub’	nat_submartingale.partial_sum_scaleR’
submartingale.sup	submartingale_linorder.max
submartingale.zero_le_of_predictable	nat_submartingale.predictable_ge_bot
submartingale_nat	nat_sigma_finite_adapted_process_linorder.submartingale_nat
submartingale_of_condexp_sub_nonneg	sigma_finite_adapted_process_order.submartingale_of_cond_exp_diff_nonneg
submartingale_of_condexp_sub_nonneg_nat	nat_sigma_finite_adapted_process_linorder.submartingale_of_cond_exp_diff_Suc_nonneg
submartingale_of_set_integral_le	sigma_finite_adapted_process_linorder.submartingale_of_set_integral_le
submartingale_of_set_integral_le_succ	nat_sigma_finite_adapted_process_linorder.submartingale_of_set_integral_le_Suc

Table 6.2: Lookup table for submartingale lemmas and definitions

Lean	Isabelle
supermartingale	supermartingale (locale)
supermartingale.adapted	adapted_process.adapted

supermartingale.add	supermartingale.add
supermartingale.add_martingale	supermartingale.add
supermartingale.condexp_ae_le	supermartingale_property
supermartingale.integrable	supermartingale.integrable
supermartingale.le_zero_of_predictable	supermartingale.predictable_le_zero
supermartingale.neg	supermartingale.uminus
supermartingale.set_integral_le	supermartingale_linorder.set_integral_ge
supermartingale.smul_nonneg	supermartingale.scaleR_nonneg
supermartingale.smul_nonpos	supermartingale.scaleR_nonpos
supermartingale.strongly_measurable	stochastic_process.random_variable
supermartingale.sub_martingale	supermartingale.diff
supermartingale.sub_submartingale	supermartingale.diff
supermartingale_nat	nat_sigma_finite_adapted_process_linorder .supermartingale_nat
supermartingale_of_condexp_sub_nonneg_nat	nat_sigma_finite_adapted_process_linorder .supermartingale_of_cond_exp_diff_Suc_nonneg
supermartingale_of_set_integral_succ_le	nat_sigma_finite_adapted_process_linorder .supermartingale_of_set_integral_le_Suc

Table 6.3: Lookup table for supermartingale lemmas and definitions

6.3 Challenges and Limitations

6.4 Future Research

Semimartingales

Doob's Martingale Convergence

Fundamental Theorem of Arbitrage The fundamental theorem of asset pricing relates the concept of a fair market price for a financial asset to the notion of a risk-neutral measure.

It provides the necessary and sufficient conditions for a market to be arbitrage-free.

In this framework, the prices of financial assets can be treated as martingales, ensuring that there is no arbitrage opportunity.

7 Conclusion

Concluded.

Abbreviations

AFP Archive of Formal Proofs

List of Figures

List of Tables

Bibliography

[DY22a] R. Degenne and K. Ying. *A Formalization of Doob's Martingale Convergence Theorems in mathlib*. 2022. arXiv: `2212.05578 [cs.LO]`.

[DY22b] R. Degenne and K. Ying. *probability.martingale.basic - mathlib*. `https://leanprover-community.github.io/mathlib_docs/probability/martingale/basic.html`, Accessed: 25 Aug 2023. 2022.

[EC82] R. J. Elliott and S. L. Cohen. *Stochastic Calculus and Applications*. Springer, 1982.

[Ech18] M. Echenim. "Pricing in discrete financial models." In: *Archive of Formal Proofs* (July 2018). `https://isa-afp.org/entries/DiscretePricing.html`, Formal proof development. ISSN: 2150-914x.

[Eng89] R. Engelking. *General Topology*. Sigma series in pure mathematics. Heldermann, 1989. ISBN: 9783885380061.

[EP17] M. Echenim and N. Peltier. "The Binomial Pricing Model in Finance: A Formalization in Isabelle." In: *Automated Deduction – CADE 26*. Ed. by L. de Moura. Cham: Springer International Publishing, 2017, pp. 546–562. ISBN: 978-3-319-63046-5.

[Fam65] E. F. Fama. "The Behavior of Stock-Market Prices." In: *The Journal of Business* 38.1 (1965), pp. 34–105. ISSN: 00219398, 15375374.

[Gou15] S. Gouëzel. "Ergodic Theory." In: *Archive of Formal Proofs* (Dec. 2015). `https://isa-afp.org/entries/Ergodic_Theory.html`, Formal proof development. ISSN: 2150-914x.

[Gou16] S. Gouëzel. "Lp spaces." In: *Archive of Formal Proofs* (Oct. 2016). `https://isa-afp.org/entries/Lp.html`, Formal proof development. ISSN: 2150-914x.

[HH11] J. Hölzl and A. Heller. "Three Chapters of Measure Theory in Isabelle/HOL." In: *Interactive Theorem Proving (ITP 2011)*. Ed. by M. C. J. D. van Eekelen, H. Geuvers, J. Schmaltz, and F. Wiedijk. Vol. 6898. LNCS. 2011, pp. 135–151.

[Hyt+16] T. Hytönen, J. v. Neerven, M. Veraar, and L. Weis. *Analysis in Banach Spaces Volume I: Martingales and Littlewood-Paley theory*. Springer International Publishing, 2016.

[Lan93] S. Lang. *Real and functional analysis*. Springer, 1993.

[MR05] M. Musiela and M. Rutkowski. *Martingale methods in financial modelling*. Springer, 2005.

[MU05] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005. DOI: 10.1017/CBO9780511813603.

[NPW02] T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*. Vol. 2283. LNCS. Springer, 2002.

[Pas11] A. Pascucci. *PDE and Martingale Methods in Option Pricing*. Springer, 2011.

[Pis16] G. Pisier. *Martingales in Banach Spaces*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2016. DOI: 10.1017/CBO9781316480588.

[Rol+23] É. Roldán, I. Neri, R. Chetrite, S. Gupta, S. Pigolotti, F. Jülicher, and K. Sekimoto. *Martingales for Physicists*. 2023. arXiv: 2210.09983 [cond-mat.stat-mech].

[Sz-+10] B. Sz.-Nagy, C. Foias, H. Bercovici, and L. Kérchy. *Harmonic analysis of operators on Hilbert space. 2nd revised and enlarged ed*. Jan. 2010. ISBN: 978-1-4419-6093-1. DOI: 10.1007/978-1-4419-6094-8.

[Tho] G. Thowler.

[YB89] R. J. Yaes and A. S. Bechhoefer. "The Efficient Market Hypothesis." In: *Science* 244.4911 (1989), pp. 1424–1424. ISSN: 00368075, 10959203.